

3ª EDICIÓN

Elementos de
Bioestadística

Colección manuales uex - 79



Agustín
García Nogales

79

3ª EDICIÓN

AGUSTÍN GARCÍA NOGALES

ELEMENTOS DE
BIOESTADÍSTICA

UNIVERSIDAD  DE EXTREMADURA

2011

GARCÍA NOGALES, Agustín

Elementos de Bioestadística / Agustín García Nogales. — Cáceres :
Universidad de Extremadura, Servicio de Publicaciones, 2011
363 pp.; 17x24 cm. - (Manuales UEX, ISSN 1135-870-X; 79)
ISBN 978-84-694-9432-5

1. Biometría. 2. Estadística médica. I. Tít. II. Serie. III. Universidad de
Extremadura, Servicio de Publicaciones, ed.

519.22:61

61:519.22



© El autor

© Universidad de Extremadura para esta 3ª edición

Edita:

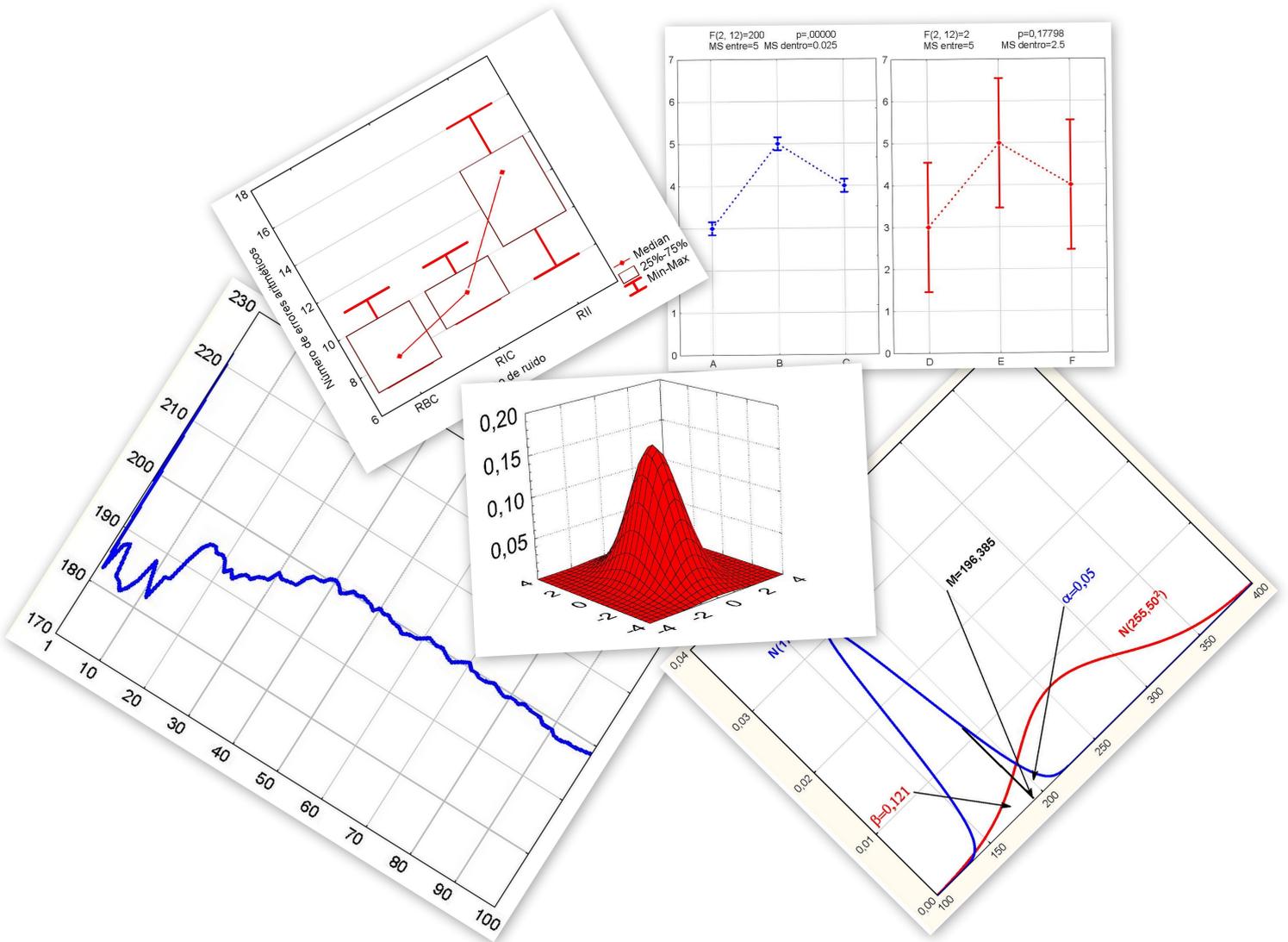
Universidad de Extremadura. Servicio de Publicaciones
C/ Caldereros, 2 - Planta 2ª. 10071 Cáceres (España)
Tel. 927 257 041 ; Fax 927 257 046
E-mail: publicac@unex.es
<http://www.unex.es/publicaciones>

ISSN 1135-870-X

ISBN 978-84-694-9432-5

Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra sólo puede ser realizada con la autorización de sus titulares, salvo excepción prevista por la ley. Diríjase a CEDRO (Centro Español de Derechos Reprográficos, www.cedro.org) si necesita fotocopiar o escanear algún fragmento de esta obra.

ELEMENTOS de BIOESTADÍSTICA



Agustín García Nogales

Departamento de Matemáticas

Facultad de Medicina

Universidad de Extremadura

A María José y Carlos

A mis padres

*Tal vez la realidad no sea más
que una utopía más*

Prólogo a esta edición

Tal vez, lo que llamamos realidad, sea lo que ésta sea, no pueda ser para el hombre más que una utopía, inalcanzable, por tanto, en buena medida, al menos, a su limitada, aunque siempre sorprendente, inteligencia, pero también, como cualquier utopía, estímulo último, y útil, que hace avanzar el conocimiento, en el sentido de que parece existir un amplio consenso en que los modelos que, como especie, vamos construyendo para explicarla se adaptan a ella cada vez mejor; el método científico, en general, y el estadístico, en particular, se han consolidado como herramientas eficientes para el diseño de esas aproximaciones a la realidad, y, en definitiva, para recorrer el agotador, pero apasionante, camino que, eso nos gustaría y nos motiva, conduce a ella.

Este libro contiene un buen puñado de técnicas estadísticas elementales apropiadas para la extracción de la información que contienen los datos obtenidos de experimentos diseñados con el propósito de descubrir nuevas (o de confirmar sospechadas) relaciones entre variables de interés en ciencias de la salud.

Sus contenidos han sido especialmente seleccionados para impartir un curso de sesenta horas de duración en el primer curso del Grado en Medicina por la Universidad de Extremadura. Entre ellas, cuarenta y cinco corresponden a clases teóricas, nueve a la resolución de ejercicios similares a los que se incluyen en este manual y las seis restantes a clases prácticas, en las que el alumno hace uso de las técnicas estadísticas estudiadas con la ayuda de algún programa estadístico. Nos vemos, pues, obligados a prescindir de otros muchos métodos estadísticos habitualmente utilizados en la investigación biomédica, como pueden ser las técnicas de regresión múltiple, regresión logística, análisis multivariante o análisis de supervivencia, por ejemplo, que el lector puede localizar, entre otros temas de interés, en la bibliografía recomendada. Sin embargo, esa limitación temporal nos ha permitido presentar un material didáctico que ha sido contrastado en clase ante alumnos brillantes y exigentes, y depurado de acuerdo con la lectura que, sobre su asimilación, hemos ido haciendo a lo largo de los últimos quince años. Esta obra se ha beneficiado también, en varios aspectos, de muchas conversaciones con varios compañeros del Departamento de Matemáticas y de

la Facultad de Medicina de la Universidad de Extremadura a quiénes, aunque sea así, de forma genérica, quiero mostrar mi agradecimiento.

Es obvio que es tanto más difícil aportar originalidad a un tema cuanto más básico es, y eso es lo que ocurre con este manual, cuyos contenidos aparecen de uno u otro modo en buena parte de los tratados más usados de bioestadística. Los conceptos básicos son introducidos con la ayuda de ejemplos concretos tomados de las ciencias de la salud y con un aparato matemático muy elemental, que sustituye las demostraciones de los resultados –lejos del alcance de este texto– por justificaciones intuitivas. No obstante, se ha procurado al mismo tiempo que sus definiciones aparezcan en la forma más próxima posible a la definición matemática original, con la intención de conservar al máximo su precisión, sacrificando sólo aquellas componentes que verdaderamente están fuera del alcance del alumno de primer curso del Grado en Medicina –como puede ser el concepto de σ -álgebra–. Un ejemplo llamativo de lo que decimos es el concepto de variable aleatoria, que es presentado normalmente en la literatura como algo diferente del concepto matemático de función, a pesar de que éste es bien conocido por el alumno que llega a esta titulación. Algo análogo ocurre con los conceptos de distribución de probabilidad o de muestra. Creemos que la concisión de las definiciones matemáticas –aliviadas de sus componentes más abstractos– facilita su asimilación por parte del alumno, procurando evitar que largas parrafadas aclaratorias –necesarias, a menudo, por otra parte– reemplacen a definiciones que no tienen otra aspiración que la de ser sencillas, inequívocas y matemáticamente rigurosas en lo posible.

Este manual viene a sustituir a otro anterior, del que aparecieron dos ediciones bajo la denominación “Bioestadística Básica”, que fue diseñado para la asignatura de Bioestadística de primer curso de la Licenciatura en Medicina de la Universidad de Extremadura. La adaptación de los estudios de Medicina al Espacio Europeo de Educación Superior justifica la necesidad de adaptar los contenidos del texto original a las nuevas circunstancias. De una presencialidad de 75 horas se pasa a una de 60 horas y eso obliga a recortar el programa de la asignatura y, por ende, a modificar los contenidos de la obra original. De hecho, se han modificado y añadido ejemplos y comentarios, y se han añadido incluso algunos nuevos métodos estadísticos. En realidad, más que recortar contenido, se han marcado con el

símbolo (★) aquellos epígrafes o ejercicios que quedan fuera del programa de la asignatura y cuyo estudio puede ser omitido por el alumno de la misma. No obstante, debido la magnitud de los cambios realizados, hemos considerado oportuno, además, cambiar el título de la obra.

El libro ha sido dividido en dos partes: la primera dedicada a la estadística descriptiva y al cálculo de probabilidades (capítulos 0, 1 y 2) y la segunda, a la inferencia estadística (capítulos 3, 4, 5 y 6).

Al final de cada capítulo, y de cada una de las dos partes, el lector puede encontrar colecciones de ejercicios, entre los que podemos distinguir dos tipos básicos: en los del primer tipo se hace una afirmación y el lector debe pronunciarse razonadamente sobre su certeza o falsedad, con el objetivo de que pueda valorar su grado de asimilación de los conceptos estudiados; los ejercicios del segundo tipo son problemas que, además y en general, pretenden adiestrar al lector en el uso de los métodos probabilísticos y estadísticos presentados.

Badajoz, otoño de 2011

Índice general

Prólogo a esta edición	VII
PRIMERA PARTE:	
Estadística Descriptiva y Cálculo de Probabilidades	9
0. Preliminares y Estadística Descriptiva	11
0.0. Introducción a la Estadística	
en las Ciencias de la Salud	12
0.0.1. Algunas definiciones e ideas básicas	12
0.0.2. Utilidad de la estadística en las ciencias de la salud	15
0.0.3. Breve repaso de la teoría de conjuntos	16
0.1. Estadística Descriptiva	22
0.1.1. Síntesis de datos: principales estadísticos descriptivos	22
0.1.2. (★) Resumen de datos agrupados en clases	26
0.1.3. Métodos gráficos	28
¿Verdadero o falso? Capítulo 0.	36
Problemas del Capítulo 0	38
1. Introducción al Cálculo de Probabilidades	41
1.1. Espacios de Probabilidad	42
1.1.1. Fenómenos aleatorios	42

1.1.2.	Espacio muestral. Sucesos	43
1.1.3.	Definición de probabilidad. Propiedades	45
1.2.	Dependencia e Independencia de Sucesos.	
	El Teorema de Bayes:	
	Aplicaciones al Diagnóstico Clínico	51
1.2.1.	Probabilidad condicionada	51
1.2.2.	Independencia de sucesos	52
1.2.3.	Teorema de Bayes. Aplicación al diagnóstico clínico	54
	¿Verdadero o falso? Capítulo 1.	61
	Problemas del Capítulo 1	63
2.	Variables Aleatorias. Distribución de Probabilidad	65
2.1.	Definición de Variable Aleatoria	66
2.1.1.	Variable aleatoria. Distribución de probabilidad	66
2.1.2.	Independencia de variables aleatorias	68
2.2.	Variables Aleatorias Discretas:	
	La Distribución Binomial	71
2.2.1.	Variable aleatoria discreta. Función de probabilidad	71
2.2.2.	Distribuciones de Bernoulli, binomial y uniforme discreta	72
2.2.3.	Momentos de una distribución discreta	73
2.3.	Variables Aleatorias Continuas:	
	La Distribución Normal	79
2.3.1.	Variable aleatoria continua: Función de densidad. Momentos	79
2.3.2.	La distribución normal. Propiedades	81
2.3.3.	Aplicación al diagnóstico clínico: intervalos de normalidad	87
2.4.	Aproximación Probabilística al Concepto de Muestra	90

2.4.1. Muestreo aleatorio simple	90
2.4.2. Concepto de muestra	92
2.4.3. Momentos muestrales: sus distribuciones	94
2.4.4. Teorema del límite central	98
2.4.5. Aproximación de la distribución binomial por la normal	99
2.4.6. Muestreo estratificado	103
¿Verdadero o falso? Capítulo 2.	105
Problemas del Capítulo 2	108
¿Verdadero o falso? PRIMERA PARTE	114
PROBLEMAS DE LA PRIMERA PARTE	120
SEGUNDA PARTE:	
Inferencia Estadística.	128
3. Introducción a la Inferencia Estadística	129
3.1. Cálculo de Probabilidades	
e Inferencia Estadística	130
3.1.1. Distinción entre Probabilidad y Estadística:	
Parámetros de una distribución	130
3.1.2. Estructura estadística. Estadísticos	132
3.1.3. Los dos grandes problemas de la Inferencia Estadística	133
3.2. Estimación	135
3.2.1. Estimación puntual. Estimador	135
3.2.2. Estimación de los parámetros de una distribución	
normal y de una proporción	136
3.2.3. Estimación conjuntista. Intervalos de confianza	139
3.2.4. Intervalo de confianza para la media de una	
distribución normal de varianza conocida	140

3.2.5.	Determinación del tamaño de muestra necesario para conseguir una cierta precisión	143
3.2.6.	Cotas superior e inferior de confianza	144
3.3.	Contraste de Hipótesis	146
3.3.1.	Introducción	146
3.3.2.	Cuatro etapas en la ejecución de un test de hipótesis	147
3.3.3.	Valor P o nivel mínimo de significación de un test	153
3.3.4.	Tests unilaterales	154
4.	Problemas de Inferencia Estadística	
	sobre una o dos Muestras	161
4.1.	Problemas de Inferencia Estadística sobre una Muestra	162
4.1.1.	Inferencia sobre la media de una distribución normal	162
4.1.2.	Inferencia sobre la varianza de una distribución normal	166
4.1.3.	(★) Comprobando la normalidad: Test de normalidad D'Agostino	167
4.1.4.	Inferencia sobre una proporción	168
4.2.	Comparación de dos Muestras	172
4.2.1.	Introducción: muestras independientes y muestras relacionadas	172
4.2.2.	Comparación de dos varianzas	174
4.2.3.	Test t de Student de comparación de dos medias: Muestras independientes y varianzas iguales	177
4.2.4.	Test de Welch de comparación de dos medias: Muestras independientes y varianzas distintas	178
4.2.5.	(★) Test no paramétrico de suma de rangos de Mann-Whitney-Wilcoxon de comparación de dos muestras independientes	180
4.2.6.	Test t de Student de comparación de dos medias en el caso de muestras relacionadas	183

4.2.7. (★) Test no paramétrico de rangos con signo de Wilcoxon de comparación de dos muestras relacionadas	184
4.2.8. Comparación de dos proporciones: muestras independientes	186
4.2.9. Test de McNemar de comparación de dos proporciones: muestras rela- cionadas	188
¿Verdadero o falso? Capítulos 3 y 4	191
Problemas de los Capítulos 3 y 4	194
5. Comparación de Varias Muestras	201
5.1. Comparación de Varias Muestras Cuantitativas: Modelo de Clasificación Simple en Análisis de la Varianza	202
5.1.1. Comparación de varias medias: muestras independientes y varianzas iguales	202
5.1.2. Comparaciones múltiples	210
5.1.3. (★) Solución no paramétrica al problema de comparación de k muestras: Test de Kruskal-Wallis	214
5.1.4. (★) Modelo de clasificación doble en análisis de la varianza sin inter- acciones y una observación por celda o diseño de bloques al azar: test F (paramétrico) y test de Friedman (no paramétrico)	217
5.2. Comparación de Varias Muestras Cualitativas	220
5.2.1. Tablas de contingencias	220
5.2.2. Comparación de varias muestras cualitativas	221
5.2.3. Comparación de varias proporciones: muestras independientes	223
5.2.4. (★) Comparación de varias proporciones: muestras relacionadas	224

¿Verdadero o falso? Capítulo 5.	226
Problemas del Capítulo 5	227
6. Relación entre Variables	231
6.1. Relación entre dos Variables Cuantitativas:	
Regresión Lineal	232
6.1.1. El problema de regresión	232
6.1.2. Regresión lineal	233
6.1.3. Estimación puntual en el modelo de regresión lineal	237
6.1.4. Contraste de hipótesis e intervalos de confianza sobre la pendiente de regresión	240
6.1.5. Inferencia sobre la media de Y dado un valor de X y sobre la ordenada en el origen	243
6.1.6. Intervalos de predicción	244
6.2. Relación entre dos Variables Cuantitativas:	
Correlación	247
6.2.1. Coeficiente de correlación de Pearson y coeficiente de determinación	247
6.2.2. (★) Correlación no paramétrica: Coeficiente de correlación de Spearman	251
6.3. Relación entre dos Variables Cualitativas	253
6.3.1. Test de independencia de dos variables cualitativas	253
6.3.2. Medidas de asociación de dos variables cualitativas	258
6.3.3. Otras medidas de asociación en el caso 2×2 : Riesgo relativo y razón del producto cruzado	259
6.3.4. Inferencia sobre el riesgo relativo y la razón del producto cruzado	260
¿Verdadero o falso? Capítulo 6.	264
Problemas del Capítulo 6	267

¿Verdadero o falso? SEGUNDA PARTE	271
PROBLEMAS DE LA SEGUNDA PARTE	277
Epílogo: ¿Qué hacer con los datos?	287
Primera etapa: entrada de datos	287
Segunda etapa: estadística descriptiva	288
Tercera etapa: inferencia estadística	291
Bibliografía	299
Apéndice: Tablas estadísticas.	300
Tabla I: 3000 dígitos aleatorios	301
Tabla II: Distribución binomial $B(n, p)$	302
Tabla III.1: Distribución $N(0, 1)$. Cuantiles	309
Tabla III.2: Distribución $N(0, 1)$. Función de distribución	310
Tabla IV: Distribución t de Student	313
Tabla V: Distribución chi-cuadrado $\chi^2(n)$	315
Tabla VI: Distribución $F(m, n)$ de Fisher	317
Tabla VII: Test de normalidad de D'Agostino	330
Tabla VIII: Test de Mann-Whitney-Wilcoxon: región de aceptación	331
Tabla IX: Distribución T de Wilcoxon para muestras relacionadas	332
Tabla X: Distribución q de Tukey para comparaciones múltiples	333
Tabla X: Distribución Q para comparaciones múltiples no paramétricas	336
Tabla XII: Distribución de Bonferroni	337
Tabla XIII: Valores críticos para el coeficiente de correlación de Spearman	339
Tabla XIV: Factores K para límites de tolerancia bilaterales para distribuciones normales	340
Soluciones para algunas de las cuestiones ¿Verdadero o Falso?	341
Soluciones para algunos de los problemas	342

Índice Alfabético

343

PRIMERA PARTE

Estadística Descriptiva y Cálculo de Probabilidades

Capítulo **0**

Preliminares y Estadística

Descriptiva

0.0. Introducción a la Estadística en las Ciencias de la Salud

0.0.1. Algunas definiciones e ideas básicas

Id	N	S	Id	N	S	Id	N	S									
1	9,0	0	22	4,1	1	43	6,7	0	64	1,0	0	85	5,5	0	106	5,6	0
2	9,0	0	23	7,2	0	44	5,8	0	65	6,0	0	86	8,5	1	107	1,9	1
3	8,3	1	24	8,4	1	45	3,4	1	66	6,3	0	87	7,1	0	108	7,0	0
4	7,2	0	25	7,1	0	46	10	1	67	7,7	0	88	8,3	0	109	5,0	0
5	7,2	1	26	7,2	0	47	6,6	1	68	3,5	0	89	8,3	1	110	10	0
6	8,2	0	27	9,6	0	48	9,1	1	69	7,1	1	90	7,7	1	111	9,0	0
7	8,7	1	28	7,4	1	49	7,0	0	70	7,0	0	91	6,4	0	112	8,5	0
8	3,9	0	29	5,9	1	50	7,5	0	71	7,4	1	92	5,7	0	113	6,4	1
9	1,3	1	30	7,9	0	51	7,2	0	72	7,8	0	93	8,7	0	114	5,8	0
10	8,2	0	31	4,1	0	52	7,8	0	73	8,0	0	94	6,7	0	115	8,2	0
11	2,7	0	32	9,0	0	53	9,0	1	74	5,4	0	95	6,7	1	116	1,2	1
12	5,5	0	33	7,2	0	54	5,5	0	75	6,5	1	96	6,4	0	117	8,4	0
13	7,5	0	34	3,1	0	55	8,6	0	76	7,7	0	97	7,5	0	118	7,6	1
14	9,0	0	35	8,0	0	56	9,0	1	77	3,7	1	98	7,7	1	119	9,1	0
15	7,9	0	36	2,7	0	57	10	0	78	8,2	1	99	8,5	0	120	1,4	1
16	8,7	1	37	9,0	0	58	5,4	0	79	8,3	0	100	6,2	1	121	9,7	1
17	6,7	0	38	10	0	59	7,8	0	80	6,0	1	101	6,5	0	122	9,1	0
18	8,8	1	39	9,4	0	60	9,0	1	81	8,6	0	102	10	0	123	5,9	0
19	9,5	1	40	9,2	0	61	7,6	1	82	8,7	1	103	3,8	1	124	7,7	0
20	6,4	0	41	7,3	0	62	3,2	0	83	6,5	0	104	7,5	0	125	7,7	0
21	8,0	0	42	8,6	0	63	3,1	0	84	10	0	105	10	0	126	10	0

Calificaciones de un examen final de Bioestadística

Los ingredientes básicos de un problema de estadística aplicada son los datos, y hemos

querido reservar para un conjunto de datos reales las primeras líneas de este libro. La columna N recoge las calificaciones de un examen final de la asignatura de bioestadística celebrado hace algunos años en la Facultad de Medicina de la Universidad de Extremadura, mientras que las columnas Id y S recogen el número de identificación del alumno y el sexo, respectivamente, donde 0 =mujer y 1 =hombre. Este conjunto de datos, aunque relativamente pequeño, ya muestra la dificultad que entraña el análisis de datos. Cabría preguntarse, por ejemplo, si esas notas son razonablemente buenas o si existe relación entre las calificaciones y el sexo. La estadística ha diseñado determinadas técnicas que ayudan a interpretar los datos y a obtener conclusiones de ellos, y el objetivo de este libro es hacer una introducción a algunos de los métodos estadísticos más sencillos. Pero antes de eso, intentaremos, aunque sea de forma intuitiva, fijar una terminología que utilizaremos con frecuencia en lo sucesivo.

No parece sencillo definir en pocas palabras y con precisión cuáles son los objetivos de una determinada ciencia, en general, ni de la estadística, en particular; una primera aproximación aparece en la siguiente definición.

DEFINICIÓN 0.1. (Estadística) La estadística es una ciencia que se interesa por la recogida, presentación y resumen de datos y la obtención de información a partir de ellos con el propósito de estudiar posibles relaciones entre variables de interés para el hombre.

Una primera clasificación de la estadística, acorde con la definición anterior, distingue entre el diseño de experimentos, la estadística descriptiva y la inferencia estadística.

DEFINICIÓN 0.2. (Diseño de experimentos) El diseño de experimentos tiene por objeto la planificación de la recogida de los datos de forma que queden garantizadas las suposiciones que requieren los métodos estadísticos que vayamos a utilizar en el análisis de los mismos.

DEFINICIÓN 0.3. (Estadística descriptiva) La estadística descriptiva estudia diversas técnicas útiles en la presentación y el resumen de un conjunto de datos.

DEFINICIÓN 0.4. (Inferencia estadística) La inferencia estadística presenta ciertos métodos que nos permiten tomar decisiones sobre relaciones entre variables en un gran conjunto de datos a partir del estudio de una pequeña y apropiada parte de los mismos.

Una segunda clasificación de la estadística la considera dividida en dos grandes áreas: la estadística matemática y la estadística aplicada.

DEFINICIÓN 0.5. (Estadística matemática) La estadística matemática tiene por objeto el desarrollo de nuevos métodos estadísticos con la ayuda de, a menudo, complejas técnicas matemáticas.

DEFINICIÓN 0.6. (Estadística aplicada) La estadística aplicada consiste en la aplicación de los métodos de la estadística matemática en otras áreas como pueden ser la economía, psicología, sociología, biología o medicina.

DEFINICIÓN 0.7. (Bioestadística) La bioestadística puede ser definida como la rama de la estadística aplicada que corresponde a la aplicación de los métodos estadísticos en ciencias de la salud y en biología.

Observación 0.8. (Población) En términos intuitivos, se habla de población en estadística en un sentido amplio para referirnos, por ejemplo, a un conjunto de personas o animales o máquinas o, en general, al conjunto de todos los objetos a los que afecta nuestro estudio y sobre los que deseamos obtener alguna información. Sus elementos suelen llamarse individuos, también en un sentido amplio. ◀

Observación 0.9. (Muestra) Una primera aproximación al concepto de muestra nos la presenta como una parte de la población que podemos considerar representativa de ella. El objetivo de la inferencia estadística consiste, precisamente, en obtener información sobre toda la población a partir de los datos obtenidos para una muestra de la misma. ◀

Observación 0.10. (Variable) Podemos estar interesados en el estudio de una o varias características (numéricas o no) de los individuos de esa población, y se habla de variable para hacer alusión a la aplicación que asigna a cada individuo de la población el valor o los valores que de él nos interesan. Una variable que toma valores numéricos suele llamarse variable cuantitativa (ej.: el peso, el número de piezas dentales con caries, el nivel de colesterol, . . .); se habla de variable cualitativa cuando la característica que nos interesa de los individuos de la población no es descrita por un número en el mismo sentido que lo es el peso o la edad sino que, de acuerdo con esa característica, el individuo es clasificado en una de varias posibles categorías (ej.: clase social a la que pertenece, partido político por el que votó en las últimas elecciones, . . .). Entre las variables cuantitativas suelen distinguirse a veces dos tipos principales: las discretas y las continuas; las primeras son aquellas que toman una cantidad finita de valores numéricos; las variables continuas son variables cuantitativas que, en principio, pueden tomar cualquier valor dentro de un cierto intervalo. Si las variables cuantitativas discretas se usan habitualmente para “contar” (ej.: el número de admisiones diarias en un cierto hospital, el número de piezas dentales con caries en los alumnos de un

cierto colegio, . . .), las continuas se suelen usar para “medir” (ej.: el peso o la altura de un individuo, su nivel de colesterol, . . .) A veces, cuando sólo nos interesa el estudio de una cierta característica en los individuos de una población, se identifica la población con el conjunto de los valores posibles de la variable que describe esa característica. <

0.0.2. Utilidad de la estadística en las ciencias de la salud

Durante las últimas décadas, las ciencias de la salud han experimentado un importante proceso de cuantificación: además del uso tradicional de información cualitativa, como puede ser el aspecto de una herida, o el estado general del enfermo, se ha aprovechado el desarrollo de la tecnología para la determinación de ciertas cantidades numéricas que pudieran tener alguna relación con la salud del paciente, como pueden ser la presión sanguínea, el nivel de glucosa en suero, etcétera. Si, salvo anomalías, algunas de esas cantidades permanecen invariantes de un individuo a otro (número de ojos, o de brazos, por ejemplo), otras reflejan lo que se suele conocer como variabilidad biológica, es decir, el hecho de que una persona es siempre diferente de otra persona (e, incluso, un ser humano es diferente a sí mismo en distintas etapas de su vida); como consecuencia, dos personas diferentes tendrán, en general, diferentes niveles de glucosa (o de colesterol, . . .) en suero sanguíneo. La estadística posee un conjunto de técnicas que nos permiten profundizar en el estudio de la variabilidad que pueda presentar un conjunto de datos de esta naturaleza, y distinguir, por ejemplo, si el valor de una cierta variable para un individuo concreto se puede considerar normal o, por el contrario, podría ser un indicio de la presencia de una cierta enfermedad. Conscientes de ese hecho, las revistas de investigación en el campo biomédico exigen un tratamiento estadístico riguroso de los datos. Algo análogo se puede decir de la investigación en biología.

Otro campo de aplicación de la estadística en las ciencias de la salud tiene que ver con su dimensión social, pues las autoridades sanitarias de un país necesitan tener una idea clara de las características de la población a la que se quiere aplicar una cierta política sanitaria y, siendo usualmente imposible o muy costoso estudiar a todos y cada uno de los individuos de la población, se puede hacer uso de la estadística para conseguir esa información a partir de una muestra representativa de la población.

Debido a ello, el método estadístico se hace cada día más necesario para el profesional de la salud, tanto en su dimensión clínica, como en la administrativa o la investigadora.

Observación 0.11. Una situación típica en la que resultan útiles los métodos estadísticos es en la comparación de dos o más tratamientos, o la comparación de dos o más grupos o poblaciones con relación a una característica de interés; por ejemplo, podemos estar interesados en comparar las presiones sanguíneas sistólicas (PSS) en un grupo de vegetarianos y en un grupo de individuos que siguen una dieta normal, con el fin de decidir si esos niveles son los mismos en ambos grupos o si, por el contrario, existen diferencias significativas entre las PSS de ambos grupos. Puede ocurrir perfectamente que elegido un par de personas de forma que una sea vegetariana y la otra con una dieta normal se obtenga una PSS mayor en la primera que en la segunda, mientras que para otro par elegido en las mismas condiciones se obtenga justo lo contrario. No parece pues apropiado basar nuestra decisión en un sólo par de personas, y la manera razonable de llegar a una decisión final es haciendo uso del método estadístico a partir de muestras suficientemente amplias y convenientemente seleccionadas en las dos poblaciones que deseamos comparar. \triangleleft

0.0.3. Breve repaso de la teoría de conjuntos

Con el fin de fijar las notaciones, repasamos a continuación algunos conceptos básicos de la teoría de conjuntos. Esta teoría comienza con los términos primitivos⁽¹⁾ de conjunto, elemento y la idea de pertenencia. Si a es un elemento de un conjunto A escribiremos $a \in A$ y leeremos a pertenece al conjunto A . Denotaremos por \mathbb{N} el conjunto de los números naturales, es decir,

$$\mathbb{N} = \{1, 2, 3, \dots\}$$

y por \mathbb{R} el conjunto de todos los números reales (rationales e irracionales), que identificaremos en la forma usual con los puntos de una recta. Si $a < b$ son números reales, los intervalos cerrado y abierto con extremos a y b se definen como sigue:

$$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$$

$$]a, b[= \{x \in \mathbb{R} : a < x < b\}$$

¹ Un término primitivo en una teoría es un concepto que no se define; resultan necesarios desde un punto de vista formal pues, siendo finito el lenguaje, es imposible definir todos los conceptos sin terminar cayendo en el error de “utilizar lo definido en la definición”.

Recordemos que se suele denotar por e el número 2.71728..., la base de los logaritmos neperianos, y por π el número 3.141592..., razón entre la longitud de la circunferencia y su diámetro. ⁽²⁾

Los conjuntos se pueden describir de dos modos: encerrando entre llaves todos y cada uno de sus elementos separados por comas ($A = \{1, 2, 3\}$) o escribiendo entre llaves una proposición verificada por todos y cada uno de los elementos de ese conjunto y sólo por ellos (por ejemplo, el conjunto de los números pares se puede escribir en la forma $\{n \in \mathbb{N} : n = 2m \text{ para algún } m \in \mathbb{N}\}$, que se lee así: conjunto de los números naturales n tales que $n = 2m$ para algún otro número natural m).

DEFINICIÓN 0.12. (Inclusión) Dados dos conjuntos A y B , diremos que A está incluido en B y escribiremos $A \subset B$ si cada elemento de A es también un elemento de B ; diremos también que A es subconjunto de B .

DEFINICIÓN 0.13. (Conjunto vacío) El conjunto vacío (que no posee ningún elemento) se denotará por \emptyset .

DEFINICIÓN 0.14. (Unión de conjuntos) La unión de dos subconjuntos A y B de un conjunto Ω es un nuevo conjunto que se denotará por $A \cup B$ y que está formado por todos los elementos de A y de B : $A \cup B = \{x \in \Omega : x \in A \text{ o } x \in B\}$.

DEFINICIÓN 0.15. (Intersección de conjuntos) La intersección de dos subconjuntos A y B de un conjunto Ω es un nuevo conjunto que se denotará por $A \cap B$ y que está formado por todos los elementos que están simultáneamente en A y en B : $A \cap B = \{x \in \Omega : x \in A \text{ y } x \in B\}$.

DEFINICIÓN 0.16. (Complementario de un conjunto) El complementario (en Ω) de un

² Redondeo: Si, operando con números reales, se desea trabajar con una precisión de, por ejemplo, dos cifras decimales, aproximaremos el número por el que forman su parte entera y las dos primeras cifras decimales cuando la tercera cifra decimal es 0, 1, 2, 3 o 4, y por el que forman su parte entera, la primera cifra decimal, y la segunda aumentada en una unidad cuando la tercera cifra decimal es 5 o superior. Por ejemplo, aproximaciones del número π con una, dos, tres y cuatro cifras decimales son 3.1, 3.14, 3.142 y 3.1416.

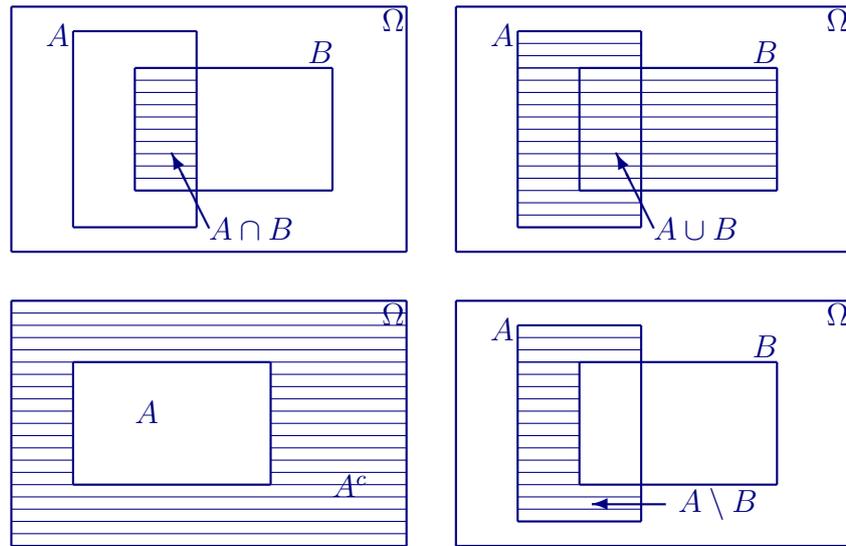


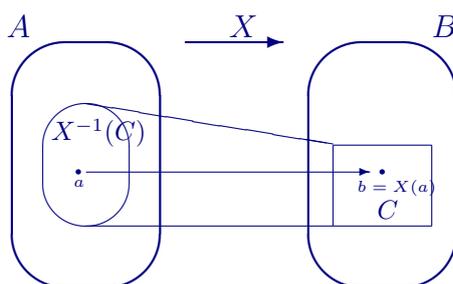
Figura 1: Intersección, unión, complementario y diferencia de conjuntos

subconjunto A de Ω se define como el conjunto A^c formado por todos los puntos de Ω que no están en A : $A^c = \{x \in \Omega: x \notin A\}$.

DEFINICIÓN 0.17. (Diferencia de dos conjuntos) La diferencia de dos conjuntos A y B se denota por $A \setminus B$ y se define como el conjunto formado por todos los elementos que están en A y no en B , es decir, $A \setminus B = A \cap B^c$.

DEFINICIÓN 0.18. (Conjunto producto o producto cartesiano de dos o más conjuntos) Dados dos conjuntos no vacíos A_1 y A_2 se define el conjunto producto $A_1 \times A_2$ como el conjunto de pares (a_1, a_2) tales que $a_1 \in A_1$ y $a_2 \in A_2$: $A_1 \times A_2 = \{(a_1, a_2): a_1 \in A_1, a_2 \in A_2\}$. Dado un conjunto A , $A \times A$ se escribe también A^2 . Análogamente, dados n conjuntos no vacíos A_1, \dots, A_n , se define el conjunto producto $A_1 \times A_2 \times \dots \times A_n$ como el conjunto de n -uplas (a_1, a_2, \dots, a_n) tales que $a_1 \in A_1, a_2 \in A_2, \dots, a_n \in A_n$: $A_1 \times A_2 \times \dots \times A_n = \{(a_1, a_2, \dots, a_n): a_1 \in A_1, a_2 \in A_2, \dots, a_n \in A_n\}$.

DEFINICIÓN 0.19. (Correspondencia entre dos conjuntos) Una correspondencia X entre los conjuntos no vacíos A y B es un subconjunto del conjunto producto $A \times B$ que se denota normalmente en la forma $X: A \rightarrow B$; si $(a, b) \in X$ suele escribirse $b = X(a)$ y decirse que b es una imagen de a por la correspondencia X y que a es una contraimagen de b .

Figura 2: Contraimagen de un conjunto C

DEFINICIÓN 0.20. (Aplicación o función) Una aplicación o función entre dos conjuntos A y B es una correspondencia $X : A \rightarrow B$ en la que cada elemento del conjunto origen A tiene una y sólo una imagen en el conjunto de llegada B . Si, además, cada elemento de B es imagen de uno y sólo un elemento de A , diremos que la aplicación es biyectiva.

DEFINICIÓN 0.21. (Contraimagen de un conjunto por una aplicación) Si $X : A \rightarrow B$ es una aplicación y C es un subconjunto de B , denotaremos por $X^{-1}(C)$ la contraimagen de C , es decir, el subconjunto de A formado por todos los elementos cuya imagen está en C :

$$X^{-1}(C) = \{a \in A : X(a) \in C\}$$

A menudo se escribe también $\{X \in C\}$ en lugar de $X^{-1}(C)$.

Observación 0.22. (Representaciones gráficas de correspondencias y aplicaciones) Una correspondencia X (y una aplicación, en particular) de un conjunto finito A en un conjunto B se puede representar gráficamente utilizando un diagrama de Venn para el conjunto A (es decir, un recinto cerrado en cuyo interior se dibujan tantos puntos como elementos tiene el conjunto A con el símbolo utilizado para cada elemento al lado de uno de los puntos) y otro para el conjunto B a la derecha de aquel, y uniendo mediante una flecha cada elemento $a \in A$ con cada una de sus imágenes en el conjunto B . Otra representación de esa correspondencia utiliza un diagrama cartesiano con una línea horizontal (llamada eje de abscisas) en el que se sitúan los elementos del conjunto A y otra vertical (llamada eje de ordenadas) en el que se sitúan los elementos del conjunto B ; a continuación se traza una línea vertical sobre cada elemento de A y una línea horizontal sobre cada elemento de B y se marcan los puntos de la malla así formada que corresponden a pares (a, b) tales que $b = X(a)$. El gráfico anterior muestra la contraimagen de un subconjunto C de B por la correspondencia X . \triangleleft

Un conjunto se dice infinito si existe una aplicación biyectiva entre él y algún subconjunto propio suyo (un subconjunto de un conjunto A se dice subconjunto propio si es no vacío y

distinto de A). Por ejemplo, \mathbb{N} es un conjunto infinito pues la aplicación $n \in \mathbb{N} \longrightarrow 2n \in \mathbb{P}$ es una biyección entre \mathbb{N} y el subconjunto \mathbb{P} de \mathbb{N} formado por los números pares. Un conjunto se dice finito si no es infinito. Un conjunto infinito A se dice numerable si existe una biyección entre \mathbb{N} y A ; por ejemplo, \mathbb{P} y el conjunto \mathbb{Q} de todos los números racionales son conjuntos numerables. Pero el conjunto \mathbb{R} de todos los números reales es un conjunto infinito no numerable.

Recordemos ahora la definición de número combinatorio. Si $n \in \mathbb{N}$ y $0 \leq k \leq n$ se define el número combinatorio n sobre k mediante:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

donde $n! = n \cdot (n-1) \dots 3 \cdot 2 \cdot 1$ (se lee n factorial y es el número de permutaciones en el conjunto $\{1, \dots, n\}$, es decir, el número de aplicaciones biyectivas de $\{1, \dots, n\}$ en sí mismo). Por convenio, $0! = 1! = 1$ y $\binom{n}{0} = 1$. El número combinatorio $\binom{n}{k}$ es el número de combinaciones de n elementos tomados de k en k , siendo dos combinaciones diferentes cuando una de ellas contiene un elemento que no contiene la otra, es decir, $\binom{n}{k}$ es el número de subconjuntos de $\{1, \dots, n\}$ que tienen k elementos. Es obvio que $\binom{n}{k} = \binom{n}{n-k}$.

Ejemplo 0.23. (★) El número de aplicaciones biyectivas (o permutaciones) del conjunto $\{1, 2, 3, 4\}$ en sí mismo es igual a $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$. Un ejemplo de permutación en ese conjunto es $(3, 1, 4, 2)$, forma abreviada ésta de describir la aplicación biyectiva que lleva el 1 en el 3 (que aparece en primer lugar en la permutación), el 2 en el 1 (que aparece en segundo lugar en la permutación), el 3 en el 4 y el 4 en el 2. Esa permutación se puede escribir también en la forma 3142, identificándose así con un número de cuatro cifras. Podemos, por tanto, afirmar que con las cifras 1, 2, 3 y 4 se pueden escribir $4! = 24$ números de cuatro cifras diferentes. Como vemos, en cada permutación en un conjunto se utiliza cada elemento de éste una y sólo una vez, y una permutación se diferencia de otra en el orden en que aparecen colocados los elementos. Por otra parte, el conjunto $\{1, 2, 3, 4\}$ tiene $\binom{4}{2} = 6$ subconjuntos con 2 elementos: $\{1, 2\}$, $\{1, 3\}$, $\{1, 4\}$, $\{2, 3\}$, $\{2, 4\}$ y $\{3, 4\}$. Se dice que hay 6 combinaciones de 4 elementos tomados de 2 en 2. Por ejemplo, hay 6 formas distintas de asignar dos manzanas idénticas a 4 niños si las manzanas no se pueden partir y cada niño recibe a lo sumo una manzana. Así, en una combinación en un conjunto no pueden aparecer elementos repetidos, y el orden no influye en la combinación, de tal suerte que dos combinaciones son diferentes cuando una de ellas contiene un elemento que no contiene la otra. Veamos un ejemplo más. En un ensayo clínico para verificar la eficacia de un tratamiento contra una enfermedad, podemos seleccionar 20 individuos enfermos que dividiremos en dos grupos de 10: los individuos de uno de los grupos (grupo de tratamiento) recibirán el tratamiento y

los del otro no (grupo de control); para evitar preferencias por parte del investigador, la asignación de 10 individuos al grupo de tratamiento se hará al azar (por sorteo), pues los 10 restantes irán al grupo de control. Existen $\binom{20}{10}$ formas posibles de asignar 10 de los 20 individuos al grupo de tratamiento. \triangleleft

Si x_1, \dots, x_n son números reales, la suma $x_1 + \dots + x_n$ se denota también por

$$\sum_{i=1}^n x_i$$

El símbolo \sum se conoce con el nombre de sumatorio. En ocasiones un conjunto finito de números reales aparece descrito con la ayuda de dos subíndices i y j ; en ese caso, la suma de todos ellos se escribe con la ayuda de un doble sumatorio: si x_{ij} , $1 \leq i \leq m$, $1 \leq j \leq n$, son números reales (exactamente, $m \cdot n$ números), su suma se denota por

$$\sum_{i=1}^m \sum_{j=1}^n x_{ij}$$

0.1. Estadística Descriptiva

0.1.1. Síntesis de datos: principales estadísticos descriptivos

Como ya hemos comentado, la estadística descriptiva diseña técnicas que puedan resultar útiles en la presentación y el resumen de un conjunto de datos. Los datos que encontramos en la práctica se suelen dividir en dos tipos, que se corresponden con los dos tipos de variables consideradas anteriormente: cualitativos (también llamados categóricos) y cuantitativos (que pueden ser, a su vez, discretos o continuos).

Ejemplo 0.24. Volvamos a los datos de la tabla 0.1. Recordemos que, junto con el número de identificación Id (dato cuantitativo discreto) y el sexo S (dato cualitativo), siendo 1=hombre y 0=mujer, se recogen las notas N (dato cuantitativo continuo) en un examen final de la asignatura de bioestadística. Podemos considerar Id , N y S como tres variables que, a cada individuo de la población formada por los alumnos que realizaron el examen final de bioestadística, asignan su número de identificación, su calificación y su sexo, respectivamente.

◁

A la hora de resumir un conjunto de datos cuantitativos x_1, \dots, x_n se suelen utilizar diversas cantidades, que llamaremos estadísticos descriptivos,⁽³⁾ que nos proporcionan información parcial sobre ellos, acerca de, por ejemplo, la “posición” de esos datos en la recta real, o la “dispersión” o “variabilidad” de los mismos.

DEFINICIÓN 0.25. (Media de un conjunto de datos cuantitativos) La media del conjunto de datos x_1, \dots, x_n es una medida de posición (la más utilizada) que denotaremos por \bar{x} y se define como la media aritmética de los datos:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

³ En este contexto, entiéndase el término “estadístico” como una función de los datos. En la Segunda Parte volveremos sobre este concepto.

Observación 0.26. La media se considera también como una medida de tendencia central de ese conjunto de datos, en el sentido de que proporciona una definición intuitivamente razonable de “centro” del mismo; de hecho, si imaginamos un cuerpo de masa 1 dividido en n masas puntuales iguales (por tanto, con masa $1/n$ cada una de ellas) situadas en los puntos de abscisas x_1, \dots, x_n , la media señala el lugar que ocupa el centro de gravedad de ese cuerpo. Se verifica además que

$$\sum_{i=1}^n (x_i - \bar{x})^2 \leq \sum_{i=1}^n (x_i - a)^2$$

para cualquier número a ; se dice por ello que la media es el valor que mejor aproxima a ese conjunto de datos en el sentido de los mínimos cuadrados. \triangleleft

La alta sensibilidad de la media a valores extremos (valores excesivamente grandes o excesivamente pequeños) puede hacer aconsejable, en ocasiones, el uso de otras medidas de tendencia central, entre las que debemos destacar la mediana, que se define a continuación.

DEFINICIÓN 0.27. (Mediana de un conjunto de datos cuantitativos) La mediana de un conjunto de datos cuantitativos se define como un valor tal que no más de la mitad de los datos son menores que él y no más de la mitad de los datos son mayores que él. Para calcular la mediana se considera la sucesión

$$x_{(1)} \leq \dots \leq x_{(n)}$$

obtenida al ordenar de menor a mayor los datos x_1, \dots, x_n . Si n es impar, la mediana es

$$x_{(\frac{1}{2}(n+1))}$$

Si n es par, cualquier valor entre los dos datos centrales $x_{(\frac{1}{2}n)}$ y $x_{(\frac{1}{2}n+1)}$ es una mediana de ese conjunto de datos; por conveniencia, en ese caso, se suele llamar mediana al valor medio de los dos datos centrales.

Observación 0.28. (★) La mediana m de ese conjunto de datos verifica que

$$\sum_{i=1}^n |x_i - m| \leq \sum_{i=1}^n |x_i - a|$$

para cualquier otro valor a . \triangleleft

Junto a la media y la mediana, otras medidas de posición que se utilizan con frecuencia son los cuantiles.

DEFINICIÓN 0.29. (Cuantiles, percentiles, cuartiles) Si α es un número entre 0 y 1, el cuantil de orden α es el dato tal que no más del $\alpha \cdot 100\%$ de los datos son menores que él y no más del $100 \cdot (1 - \alpha)\%$ de los datos son mayores que él.⁽⁴⁾ El cuantil de orden α se llama también percentil $\alpha \cdot 100$. Los percentiles 25 y 75 se suelen llamar primer y tercer cuartil, respectivamente, o, también, cuartiles inferior y superior.

Observación 0.30. La mediana es el percentil 50. \triangleleft

Observación 0.31. Entre los cuartiles inferior y superior se sitúa el 50% central del conjunto de datos. \triangleleft

Ejemplo 0.32. (★) Consideremos el siguiente conjunto de datos ordenados de menor a mayor: 1, 3, 5, 5, 6, 7, 8. La media y la mediana son iguales a 5, el percentil 25 (o cuantil de orden 0.25) es 3 (pues el 25% de 7 es 1.75 y el 75% de 7 es 5.25, y hay un sólo dato < 3 y 5 datos > 3). El percentil 10 es 1. Nótese como el hecho de reemplazar el valor 8 por 15 no tiene efecto sobre la mediana, mientras que la media, más sensible a valores extremos, pasa a ser igual a 6. \triangleleft

La medida de dispersión para un conjunto de datos más conocida es la varianza.

DEFINICIÓN 0.33. (Varianza de un conjunto de datos cuantitativos) La varianza s^2 de un conjunto de datos x_1, \dots, x_n se define por

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right)$$

es decir, es la suma de los cuadrados de las desviaciones a la media de cada uno de los datos dividida por $n - 1$. Su raíz cuadrada s se conoce como desviación típica.

Observación 0.34. La desviación típica presenta sobre la varianza la ventaja de estar expresada en la misma unidad de medida que los datos. \triangleleft

Observación 0.35. La varianza, como medida de la variabilidad de un conjunto de datos, debe ser pequeña (respectivamente, grande) si los datos están próximos a (respectivamente, lejos de) la media. Algunos autores prefieren definir la varianza muestral como $\frac{n-1}{n}s^2$, expresión que coincide con la varianza, salvo que la suma de las desviaciones cuadráticas respecto a la media se divide por n en lugar de dividir por $n - 1$, con lo cual se convierte

⁴A pesar de que hemos insistido en que el cuantil es uno de los datos, como en el caso de la mediana, si la definición es verificada por dos datos, se suele tomar como cuantil el valor medio de ambos. Intuitivamente, pensaremos que un $\alpha \cdot 100\%$ de los datos se sitúa por debajo del cuantil de orden α y que un $(1 - \alpha) \cdot 100\%$ de los datos son mayores que él.

en la desviación cuadrática media de los datos respecto a la media; la justificación de la definición presentada quedará clara en los temas séptimo (ver observación en la página 94) y noveno (véase la primera observación en la página 137). Aquí denominaremos quasi-varianza muestral a $\frac{n-1}{n}s^2$. ◁

Observación 0.36. (★) Se puede probar que, cualesquiera que sean los datos, a una distancia de dos desviaciones típicas de la media se encuentran, como mínimo, el 75 % de los datos; a tres desviaciones típicas de la media podemos encontrar al menos el 88.88 % de los datos. ◁

Observación 0.37. (★) Si el análogo físico de la media de un conjunto de datos x_1, \dots, x_n se corresponde con el centro de gravedad de un cuerpo de masa 1 dividido n masas puntuales situadas en los puntos de abscisa x_1, \dots, x_n , la quasi-varianza muestral $\frac{n-1}{n}s^2$ se corresponde con el momento de inercia de ese mismo cuerpo alrededor de su centro de gravedad, y se interpreta como una medida de la resistencia de ese cuerpo a girar entorno a un eje situado en el centro de gravedad. ◁

Observación 0.38. (★) Si μ y m son la media y la mediana de un conjunto de datos y s es su desviación típica, se verifica que $|m - \mu| \leq s$. ◁

Ejemplo 0.39. (Ejemplo 0.24, continuación) Los estadísticos descriptivos que acabamos de introducir nos permiten resumir el conjunto de datos de la variable N del siguiente modo:

Datos válidos	Media	Mediana	Mín.	Máx.
100	7.1	7.5	1	10
Varianza	Desv. Típ.	Cuartil Inf.	Cuartil Sup.	Rango
4.4	2.1	6.2	8.6	9

Para la variable S basta notar que hay 39 chicos y 87 chicas. También puede resultar de interés conocer los estadísticos descriptivos de la variable N para cada uno de los valores de la variable S (es decir, categorizados por los valores de S). La tabla siguiente contiene algunos de esos estadísticos:

	Datos válidos	Media	Desv. Típ.	Mediana	Mín.	Máx.
$S = 0$	87	7.2	1.9	7.5	1	10
$S = 1$	39	6.8	2.5	7.6	1.2	10

Observación 0.40. En inferencia estadística, los datos x_1, \dots, x_n se suponen a menudo una realización concreta de una muestra de tamaño n , es decir, valores tomados por n variables aleatorias independientes e idénticamente distribuidas (véanse las definiciones de estos conceptos en el Capítulo 2). En ese caso, los estadísticos descriptivos definidos anteriormente (media, mediana, varianza, etc...) se suelen acompañar del calificativo “muestral”. Los datos podrían ser, por ejemplo, los resultados de n mediciones de una misma cantidad, quedando explicada la variabilidad observada en los mismos por la imprecisión del aparato de medida utilizado; la variabilidad presente en un experimento como ese se atribuye a lo que genéricamente se denomina “azar”, aludiendo con ello al elevado número de factores incontrolables que dan lugar a diferentes resultados para un experimento que se ha procurado realizar bajo idénticas condiciones. ◁

Observación 0.41. Como hemos dicho antes, una medida de posición como la media y una medida de dispersión como la varianza contienen información parcial sobre un conjunto de datos (evidentemente, sólo en casos triviales es posible recuperar el conjunto de datos original una vez conocida su media y su varianza), pero constituyen una buena primera síntesis de los mismos en la medida en que nos permiten situar su centro y conocer su variabilidad alrededor del mismo:

- (a) Los conjuntos de datos 3,4,5 y 2,4,9 tienen distintas medias (4 y 5, respectivamente) y distintas varianzas (1 y 13, respectivamente).
- (b) Los conjuntos de datos 3,4,5 y 2,4,6 tienen la misma media (4) y distintas varianzas (1 y 4, respectivamente).
- (c) Los conjuntos de datos 3,4,5 y 7,8,9 tienen distintas medias (4 y 8, respectivamente) y la misma varianza (1).
- (d) Los conjuntos de datos 3,4,5 y $\frac{15}{4}(1 + \sqrt{13})$, $\frac{15}{4}(1 - \sqrt{13})$, 4,5 tienen la misma media (4) y la misma varianza (1). ◁

Proposición 0.42. *Si a todos los datos le sumamos una misma cantidad, la media queda aumentada en esa misma cantidad, pero la varianza permanece inalterable. Si todos los datos se multiplican por una misma cantidad a , la media queda multiplicada por a y la varianza queda multiplicada por a^2 .*

0.1.2. (★) Resumen de datos agrupados en clases

En ocasiones, el tamaño n del conjunto de datos es excesivamente grande y conviene agrupar los datos en tablas de frecuencia que nos permiten una mejor visualización de los mismos.

DEFINICIÓN 0.43. (Tabla de frecuencias, clases, frecuencias absoluta y relativa de una clase) Una tabla de frecuencias es un despliegue ordenado de los valores que aparecen en un conjunto de datos junto con las frecuencias de esos valores. Considerando los datos como valores concretos tomados por una cierta variable, el conjunto de valores posibles de esa variable se divide en un cierto número de clases; en la tabla de frecuencias aparecen cada una de esas clases junto con su frecuencia absoluta (número de datos que están en esa clase) y su frecuencia relativa (que se define como la frecuencia absoluta de la clase dividida

por el número total n de datos); se pueden añadir también las frecuencias (absoluta y relativa) acumuladas de cada clase, que se obtienen sumando las frecuencias de todas las clases anteriores, incluida ella misma. Si la variable es cualitativa o cuantitativa discreta, es habitual considerar como clase cada valor posible de la variable, aunque también se pueden agrupar varios valores en una sola clase cuando hay demasiados valores posibles. Para variables cuantitativas continuas, las clases suelen ser intervalos de la recta que llamaremos intervalos de clase, y el punto medio del intervalo se denomina marca de la clase.

Ejemplo 0.44. (Ejemplo 0.24, continuación) La primera de las tablas de frecuencias siguientes (divididas en 4 clases cada una, clases que han sido elegidas debido a la naturaleza de los datos) corresponden al conjunto de todos los datos de la variable N , mientras que las dos siguientes corresponden a los datos de la variable N para cada uno de los dos valores posibles de la variable S , es decir, son tablas de frecuencias de calificaciones categorizadas por sexos.

Todos	Frec. Abs.	Frec. Abs. Acum.	Frec. Rel.	Frec. Rel. Ac.
$0 \leq N < 5$	9	9	0.13	0.13
$5 \leq N < 7$	21	30	0.22	0.35
$7 \leq N < 9$	41	71	0.47	0.82
$9 \leq N \leq 10$	16	87	0.18	1

$S = 0$	Frec. Abs.	Frec. Abs. Acum.	Frec. Rel.	Frec. Rel. Ac.
$0 \leq N < 5$	17	17	0.10	0.10
$5 \leq N < 7$	28	45	0.24	0.34
$7 \leq N < 9$	59	104	0.47	0.81
$9 \leq N \leq 10$	22	126	0.19	1

$S = 1$	Frec. Abs.	Frec. Abs. Acum.	Frec. Rel.	Frec. Rel. Ac.
$0 \leq N < 5$	8	8	0.21	0.21
$5 \leq N < 7$	7	15	0.18	0.39
$7 \leq N < 9$	18	33	0.46	0.85
$9 \leq N \leq 10$	6	39	0.15	1

Observación 0.45. (★) Un problema importante en la elaboración de una tabla de frecuencias para datos cuantitativos es la elección del número de intervalos de clase y la longitud de los mismos. Como regla general, se elegirá un número de clases no menor que 5 (un número menor de clases supondría una excesiva pérdida de la información contenida en el conjunto original de datos) ni mayor que 20 (un número mayor de clases podría oscurecer la visualización de los datos). A falta de un mejor criterio, Sturges propuso en 1926 que, para un número de datos n superior a 50, se utilice un número de clases igual al entero más próximo a $1 + 3.322 \log_{10} n$. En ocasiones, se aconseja también el entero más próximo a \sqrt{n} . Una vez elegido el número k de intervalos de clases, se suelen tomar intervalos de

clase contiguos y disjuntos con la misma longitud (igual al rango dividido por k); a veces, el extremo inferior del primer intervalo de clase se reemplaza por $-\infty$ y el extremo superior del último intervalo de clase se reemplaza por $+\infty$. \triangleleft

Observación 0.46. (★) Una vez construida la tabla de frecuencias, los datos originales han desaparecido y sólo se conservan las clases y el número de datos de cada clase; a todos los efectos, los datos dentro de un intervalo de clase se identifican con lo que llamaremos marca de clase, que se define como el punto medio del intervalo correspondiente. Puesto que no conservamos los datos originales, no es posible tampoco calcular a partir de los datos agrupados estadísticos descriptivos como la media, la mediana o la varianza. En ese caso, la media se aproxima por $\frac{1}{n} \sum_{i=1}^k f_i x_i$, donde x_i denota la marca de clase del intervalo de clase i -ésimo y f_i la frecuencia absoluta del mismo. En nuestro Ejemplo 0.24, para el conjunto de todos los datos la media es igual a 7.094. A partir de la primera tabla de frecuencias ese valor se aproxima por 7.075. La varianza se aproxima por $\frac{1}{n-1} \sum_{i=1}^k f_i (x_i - \bar{x})^2$. \triangleleft

Observación 0.47. (★) De acuerdo con lo dicho en la observación anterior, la mediana debería aproximarse por la marca de la clase en la que por primera vez la frecuencia relativa acumulada supera 0.5, que en el ejemplo sería 8; no obstante, es preferible hacer una interpolación lineal del siguiente modo: si la frecuencia relativa acumulada hasta 7 es 0.35 y hasta 9 es 0.82, asumiendo un comportamiento lineal dentro de ese intervalo, la frecuencia relativa acumulada sería igual a 0.5 en el punto $7 + \frac{9-7}{0.82-0.35}(0.5 - 0.35) = 7.64$. Véase una descripción más detallada del método de interpolación en la página 97. \triangleleft

0.1.3. Métodos gráficos

La estadística descriptiva también tiene por objeto el diseño de ciertos gráficos que nos permitan obtener una rápida impresión del conjunto de datos. Describimos a continuación algunos gráficos que se utilizan con frecuencia en la práctica.

DEFINICIÓN 0.48. (Diagrama de barras) El diagrama de barras se utiliza preferentemente para datos categóricos o para datos numéricos no agrupados. Para cada valor posible de la variable cualitativa o cada grupo se construye un rectángulo de altura proporcional a la frecuencia; los rectángulos, generalmente, no son contiguos, tienen la misma base y están igualmente separados unos de otros.

Para datos numéricos agrupados se suele utilizar el histograma.

DEFINICIÓN 0.49. (Histograma) Un histograma es una representación gráfica de datos numéricos agrupados en el que sobre cada intervalo de clase se construye un rectángulo

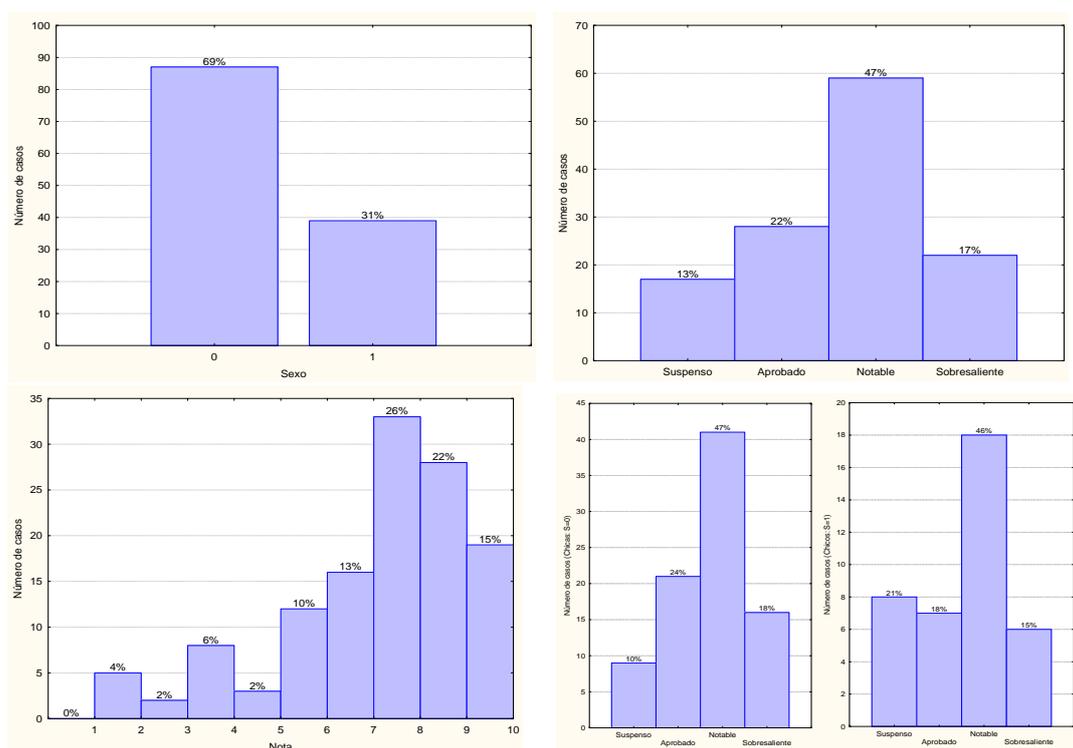


Figura 3: Diagrama de barras e histogramas

cuya base es ese intervalo y cuyo área (no necesariamente la altura) es proporcional a la frecuencia del intervalo de clase.

Para los datos de la tabla 0.1., la Figura 3 contiene un diagrama de barras para la variable Sexo, dos histogramas para la variable Nota con diferentes números de clases (obsérvese cómo influye el número de clases influye en la impresión visual que produce el histograma) y un histograma para la variable Nota categorizada por la variable Sexo.

Un tipo de gráfico, introducido por Tukey en el desarrollo de lo que se ha dado en llamar análisis exploratorio de datos, que es cada vez más utilizado, es el diagrama de caja (en inglés, box plot).

DEFINICIÓN 0.50. (Diagrama de caja) El diagrama de caja de un conjunto de datos numéricos consta de un rectángulo vertical que va desde el cuartil inferior al cuartil superior (el ancho del rectángulo no tiene significado alguno) que se atraviesa con una raya horizontal

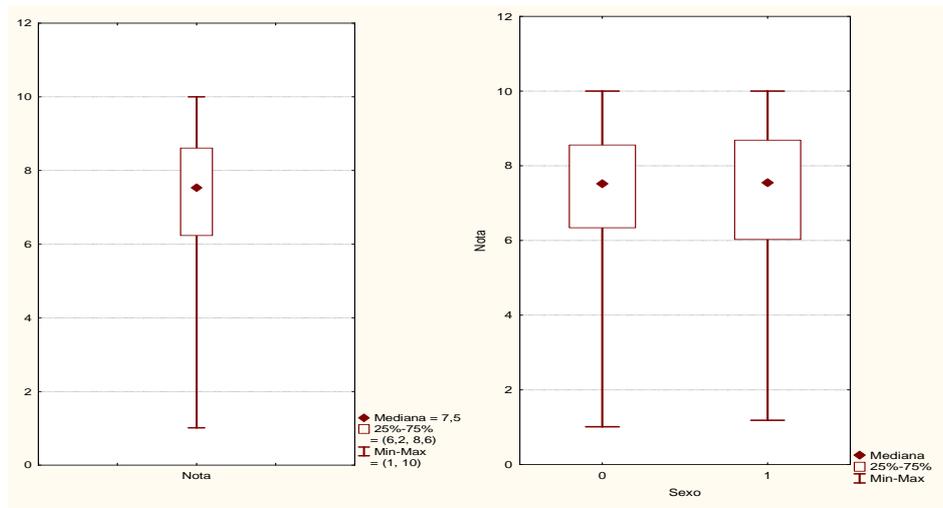


Figura 4: Diagrama de caja

(u otro tipo de marca) a la altura de la mediana y que se completa añadiendo una línea que une el centro de la base del rectángulo con un punto a la altura del menor de los datos y con otra línea análoga en la parte superior del rectángulo.

Para la variable N , los diagramas de caja, incluyendo categorizaciones por sexo, son los siguientes:

Observación 0.51. El diagrama de caja nos proporciona una primera imagen sobre la distribución del conjunto de datos, la simetría de la misma y su rango, pues contiene los cuatro puntos correspondientes a los valores mínimo y máximo (ver también observación siguiente sobre datos anómalos) y los cuartiles inferior y superior, y entre cada dos consecutivos de ellos se sitúa el 25 % de los datos. En el ejemplo de las notas de bioestadística, nótese como el diagrama de caja de todas las calificaciones presenta una ligera asimetría, pues el 50 % de los datos que son menores que la mediana está más disperso que el 50 % de datos mayores que la mediana; incluso, el primer y segundo 25 % de los datos está más disperso que el cuarto y el tercero, respectivamente. Nótese también que esa asimetría es fundamentalmente debida a la distribución de las calificaciones en el grupo $S = 0$, pues la distribución de las notas en el grupo de las chicas es prácticamente simétrica. \triangleleft

Observación 0.52. (★) En la definición original de box plot dada por Tukey, en lugar de los cuartiles inferior y superior se utilizan las que podemos traducir del inglés como bisagras inferior y superior, pero, para nuestras necesidades, es más sencilla y fácil de interpretar la definición que hemos dado, no habiendo, en cualquier caso mucha diferencia entre ambas.

\triangleleft

Observación 0.53. (Datos anómalos y datos anómalos extremos) En ocasiones, un dato puede parecer demasiado discrepante del resto. Podemos calificar como anómalo cualquier dato mayor que el cuartil superior más 1,5 veces la diferencia entre los cuartiles superior e inferior o menor que el cuartil inferior menos 1,5 veces la diferencia entre los cuartiles superior e inferior. Si se reemplaza el factor 1,5 por 3 se habla a veces de dato anómalo extremo. Un diagrama de caja más completo se obtiene añadiendo una línea que une el centro de la base del rectángulo con un punto a la altura del menor de los datos no anómalos y se traza una línea análoga en la parte superior del rectángulo, y se identifican los datos anómalos con un 0 y los datos anómalos extremos con un *. Conviene señalar que diferentes autores proponen diferentes definiciones de datos anómalos (u observaciones extremas). Antes de depurar nuestros datos eliminando datos anómalos, conviene notar que las consecuencias pueden resultar decisivas (sobre todo cuando hacemos inferencia estadística a partir de ellos, es decir, cuando extrapolamos conclusiones extraídas de los datos a una población mucho más grande que contiene a los individuos que los proporcionaron). Los datos anómalos se rechazarán sólo cuando existan fundadas sospechas de que no corresponden a un valor de la variable que estudiamos. <

Observación 0.54. (★) (Coeficiente de asimetría) Como medida de la posible asimetría del conjunto de datos se utiliza a veces la cantidad $m_3 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3$; con la idea de usar una cantidad “sin unidades”, es preferible utilizar el coeficiente de asimetría m_3/s^3 , donde s es la desviación típica del conjunto de datos. La asimetría del conjunto de datos es tanto más grande cuanto más distinto de cero es dicho coeficiente, lo cual se utiliza a veces como indicio de la no “normalidad” de los datos, según comentaremos en el tema 9. Si el coeficiente de asimetría es nulo, los datos se distribuyen simétricamente entorno a la media, que, en este caso, coincide con la mediana. Si el coeficiente de asimetría es menor que cero se dice que el conjunto de datos es sesgado hacia la izquierda o negativamente sesgado (es decir, el 50 % de los datos situado a la izquierda de la mediana presenta una mayor dispersión que el 50 % de los datos mayores que la mediana); en ese caso, la media es menor que la mediana. Imaginemos una barra de hierro homogénea con una sola pesa; si la pesa se sitúa en el punto medio de la barra, la masa total se distribuye simétricamente y el centro de gravedad μ -la media- coincide con el punto m que deja a su izquierda la misma masa que a la derecha -la mediana-; si la pesa se desplaza hacia la izquierda, tanto μ como m se desplazan hacia la izquierda, pero ocurre que $m < \mu$. Si el coeficiente de asimetría es mayor que cero se dice que el conjunto de datos es sesgado hacia la derecha o positivamente sesgado (es decir, el 50 % de los datos menores que la mediana está menos disperso que el 50 % de los datos mayores que la mediana); en ese caso, la media es mayor que la mediana. El gráfico precedente, en el que el histograma ha sido aproximado por el área comprendida entre el eje X y las curvas, nos muestra ambas situaciones; en el Ejemplo 5, el histograma de la variable N indica un ligero sesgo hacia la izquierda (el coeficiente de asimetría es igual a -0.212) debido especialmente al grupo $S = 0$ (el coeficiente de asimetría en ese grupo es igual a -0.48). <

Observación 0.55. (★) (Curtosis) Otra medida de forma que se utiliza también con el objetivo de verificar la posible normalidad de los datos es el coeficiente de curtosis que se

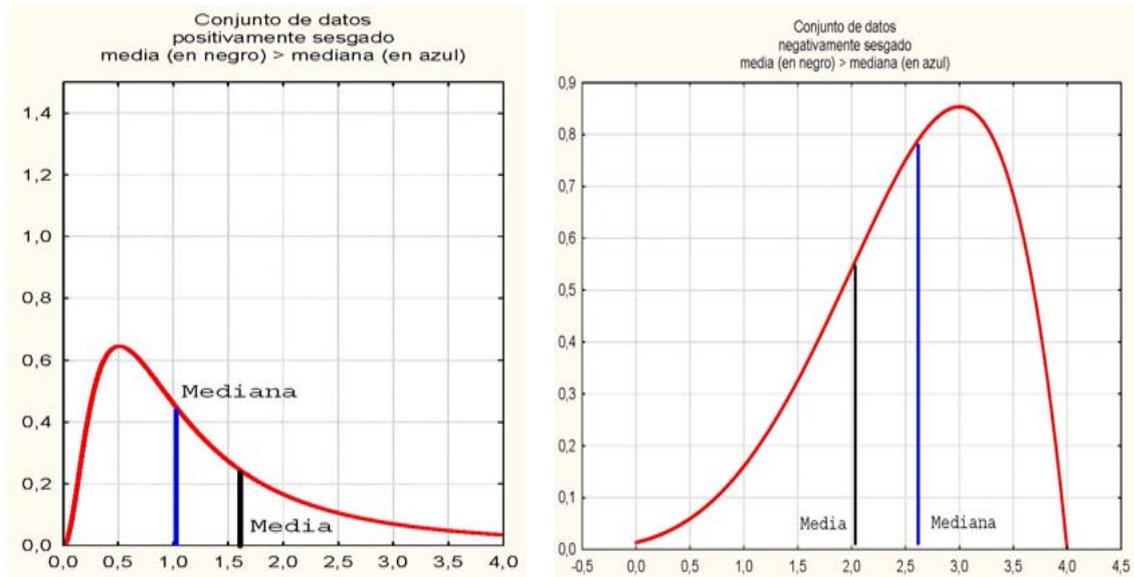


Figura 5: Asimetría de un conjunto de datos

define a partir de las cantidades $m_4 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4$ y s^4 mediante una expresión que, por compleja, no recogemos aquí. Este coeficiente es nulo para un conjunto de datos procedentes de una distribución normal (conjunto que se dice mesocúrtico), distribución para la que un 68 % aproximadamente de los valores se encuentran a menos de una desviación típica de la media (véase la página 84). Un conjunto de datos que a menos de una desviación típica de la media contiene menos valores que la normal tendrá un coeficiente de curtosis menor que cero y se dirá platicúrtico. Un conjunto de datos que a menos de una desviación típica de la media concentre más del 68 % de los mismos tendrá un coeficiente de curtosis mayor que cero y se dirá leptocúrtico. Así pues, un coeficiente de curtosis muy distinto de cero debe interpretarse como un indicio de no normalidad de los datos. <

Observación 0.56. (★) (Diagrama de tallos y hojas) Un método pseudográfico para datos numéricos alternativo al histograma es el llamado diagrama de tallos y hojas que describimos a continuación en varios pasos: a) Cada dato es dividido en dos partes: el último dígito a la derecha, llamado hoja del dato, y el resto, llamado tallo (así, el tallo de 546 es 54 y su hoja es 6). b) Se determinan los tallos menor t_1 y mayor, y en una columna se escriben de arriba a abajo el tallo menor t_1 y los números $t_1 + 1, t_1 + 2, \dots$ hasta alcanzar el tallo mayor. c) Se traza una línea vertical justo a la derecha de la columna de tallos y a la derecha de ésta se van escribiendo todas y cada una de las hojas en la fila de su tallo correspondiente y unas a continuación de otras ordenadas de menor a mayor en cada una de las filas.

Por ejemplo, el conjunto de datos

78, 23, 21, 44, 48, 49, 57, 20, 23, 78, 29, 77

se representaría del siguiente modo:

2	0	1	3	3	9
3					
4	4	8	9		
5	7				
6					
7	7	8	8		

Presenta este diagrama frente al histograma la ventaja de ser fácilmente construible y de conservar el valor de todos y cada uno de los datos, al tiempo que nos permite una rápida visualización de la distribución de los datos, y el inconveniente de ser menos flexible que aquel. Para otros conjuntos de datos puede ser preferible utilizar otras variaciones de este diagrama: (a) Si los datos presentan una gran variabilidad, podemos permitir que las hojas tengan dos o más dígitos, en cuyo caso los dígitos de cada hoja aparecerán subrayados: por ejemplo, para el conjunto de datos

2105, 3223, 3280, 3254, 2327, 3111, 2999, 2114, 2600

obtendríamos el siguiente diagrama de tallos y hojas:

21	<u>05</u>	<u>14</u>
22		
23	<u>27</u>	
24		
25		
26	<u>00</u>	
27		
28		
29	<u>99</u>	
30		
31	<u>11</u>	
32	<u>23</u>	<u>54</u> <u>80</u>

(b) Cuando hay muchas hojas por tallo, podemos dividir cada tallo en dos filas consecutivas, que podemos marcar con dos signos distintos (un * y un ●, por ejemplo): en la superior se colocan las hojas 0,1,2,3,4 y en la inferior se colocan las hojas 5,6,7,8,9. Así, podemos representar el conjunto de datos

23, 21, 23, 22, 28, 32, 33, 38, 25, 27, 26, 25, 35, 33, 30, 31

mediante el siguiente diagrama

2*	1	2	3	3	
2●	5	5	6	7	8
3*	0	1	2	3	3
3●	5	5	8		

(c) Para el conjunto de notas de la tabla 0.1, suponiendo que sólo conservamos una cifra decimal de cada nota, el tallo sería la parte entera de la calificación y la hoja sería la parte decimal. Se obtiene así un diagrama como el siguiente,

1	1 3 3 3	4	4
2	1 5 9	3	7
3	0 1 2 3 4 5 7 7 9 9	10	17
4	0 0 1 3 3 5 5 5 5 7	10	27
5	1 1 2 3 3 4 5 5 5 5 6 6 7 7 8 8	16	43
6	0 0 0 0 0 0 1 1 2 4 4 5 5 5 6 7 7 8	18	61
7	0 0 0 0 0 0 0 1 1 1 2 3 4 5 5 7 7 8 9	19	80
8	0 0 2 2 4 5 6 8	8	88
9	0 2 3 4 5 5 6 6	8	96
10	0 0 0 0	4	100

en el que hemos añadido dos columnas a la derecha que contienen la frecuencia absoluta y la frecuencia absoluta acumulada de cada tallo, respectivamente. De esta forma es sencillo determinar la mediana (6.1) como el valor medio de los datos que ocupan las posiciones 50 y 51 (aparecen en negrita en el diagrama). ◀

Observación 0.57. (★) Se utilizan en la práctica otros muchos tipos de gráficos, como pueden ser los diagramas de sectores (un círculo es dividido en tantos sectores como clases tenemos y el área de aquel es proporcional a la frecuencia de esta) o el polígono de frecuencias (a cada clase se le asigna un punto del plano cuya abscisa es la marca de la clase y cuya ordenada es la frecuencia -absoluta o relativa- de la misma, y los puntos correspondientes a dos clases consecutivas se unen mediante una recta). ◀

Observación 0.58. (Gráficos engañosos) Cambios de escala en alguno de los ejes coordenados (en una observación de la página 84 veremos un ejemplo), o el hecho de que los ejes se corten en un punto distinto del $(0, 0)$ tienen una importante repercusión en la impresión que el gráfico produce en el lector. La Figura 6 contiene dos diagramas de barras, que, aunque no lo parezca (la diferencia entre los números de chicas y chicos parece mucho mayor en el segundo gráfico que en el primero), representan el mismo conjunto de datos (véanse además las figuras 6.2 y 6.3, correspondientes también a un mismo conjunto de datos).

Son éstos trucos que se utilizan indebidamente con la intención de transmitir una impresión equivocada del conjunto de datos. La prensa diaria está plagada de gráficos engañosos (evolución en el tiempo de las cifras de paro, o del gasto farmacéutico, etc), acorde probablemente con los argumentos utilizados por el autor del artículo. No queremos decir que lo correcto sea siempre utilizar la misma longitud para el intervalo unidad en ambos ejes, pues el simple hecho de cambiar la unidad de medida (de centímetros a metros, por ejemplo) en uno de los ejes provoca un cambio de escala y un gráfico aparentemente diferente; más bien queremos llamar la atención sobre este asunto para que el gráfico que finalmente se presente permita una adecuada visualización de los datos y una impresión “honesta” en el lector. ◀

Observación 0.59. (★) Otras medidas de centralización menos utilizadas son la media geométrica de un conjunto de datos positivos x_1, \dots, x_n , que se define como $(x_1 \cdots x_n)^{\frac{1}{n}}$,

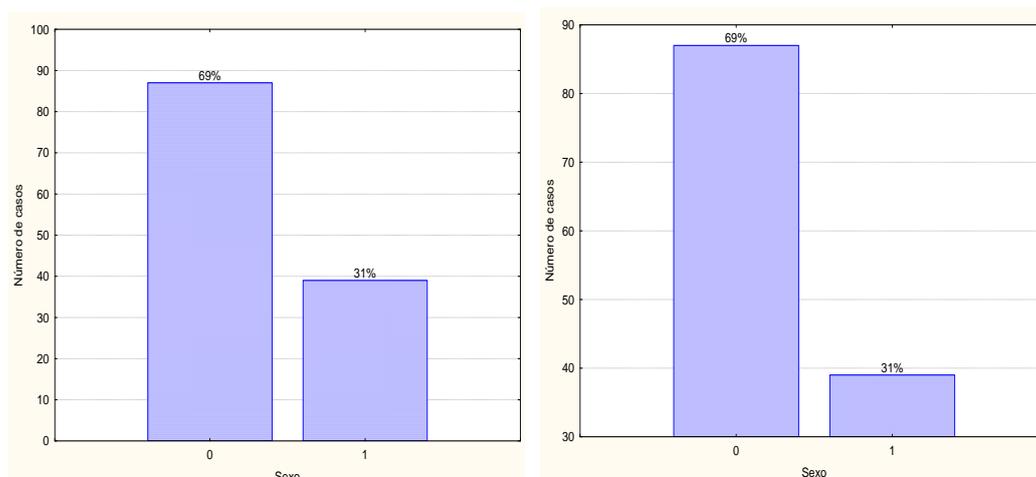


Figura 6: Gráficos engañosos

y la media armónica, que se define como el inverso de la media de los inversos de los datos. Cuando, a la hora de evaluar la media aritmética, no concedemos la misma relevancia a todos los datos, podemos pensar en utilizar una media ponderada de la forma $(\sum_{i=1}^n p_i x_i) / \sum_{i=1}^n p_i$ donde los coeficientes p_1, \dots, p_n son números positivos llamados pesos. \triangleleft

Observación 0.60. (★) Otras medidas de dispersión que se utilizan a veces son el rango de un conjunto de datos (diferencia entre el mayor y el menor de ellos), el rango intercuartílico (diferencia entre los cuartiles superior e inferior), la desviación media, que se define como la media de los valores absolutos de las desviaciones absolutas a la media de los datos, es decir,

$$\frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}|$$

y la meda, que se define como la mediana de las desviaciones absolutas de los datos respecto a su mediana. Para ciertas variables cabe esperar un aumento de la variabilidad cuando la media aumenta. Por ello y por que una desviación típica de 5 unidades se interpretará de distinta forma cuando la media es 10 que cuando es 1000, puede resultar útil como medida de dispersión relativa el coeficiente de variación

$$CV = 100 \cdot \frac{s}{\bar{x}}$$

que se expresa en porcentaje (por ejemplo, podemos decir que un conjunto de datos tiene un coeficiente de variación del 20% si $s/\bar{x} = 0.2$) y es adimensional, pues s y \bar{x} se expresan en las mismas unidades de medida. \triangleleft

Otros estadísticos descriptivos y métodos gráficos serán introducidos en temas posteriores en un contexto más apropiado.

¿Verdadero o Falso? Capítulo 0

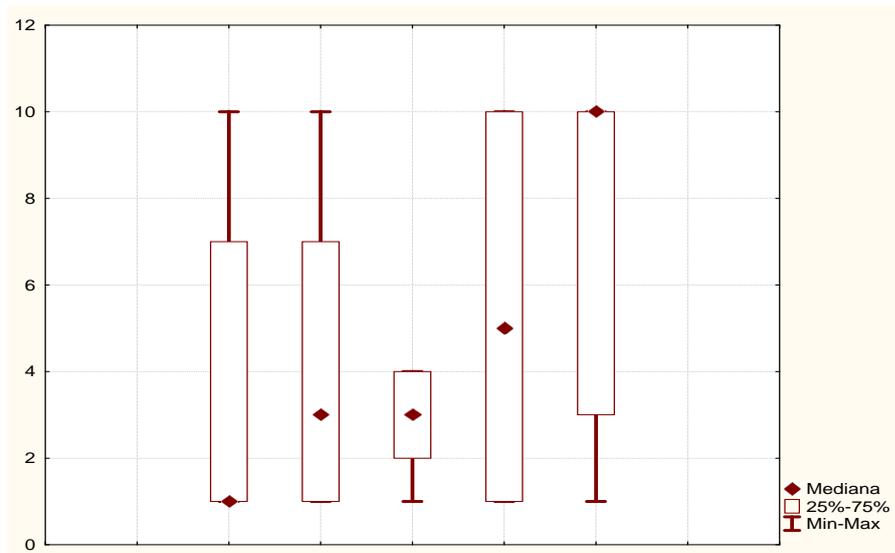
Decidir si cada una de las proposiciones siguientes es verdadera (V) o falsa (F), justificando la respuesta.

- 0.1** Si $x_{ij} = i + j$, $1 \leq i \leq 4$, $1 \leq j \leq 3$, entonces $\sum_{j=1}^3 x_{4j} = 18$.
- 0.2** Si $x, y \in \mathbb{R}$, $(x + y)^2 = x^2 + y^2$.
- 0.3** (★) Dos permutaciones distintas en el conjunto $\{a, b, c, d\}$ se diferencian solamente en el orden en que aparecen sus elementos. SOLUCIÓN: PÁG. 341
- 0.4** La negación de la proposición “este individuo es hipertenso y fumador” es “este individuo ni es hipertenso ni es fumador”.
- 0.5** (★) Si un conjunto de datos numéricos es agrupado en clases, su media coincide con la media calculada a partir de los datos agrupados.
- 0.6** El diagrama de barras resulta apropiado para la presentación gráfica de datos cualitativos. SOLUCIÓN: PÁG. 341
- 0.7** Si en un conjunto de datos numéricos, al menor de los datos se le resta una unidad, la varianza aumenta.
- 0.8** Si en un conjunto de tres o más datos numéricos, al mayor de los datos se le suma una unidad, la mediana aumenta.
- 0.9** La media de un conjunto de datos numéricos es más sensible a valores extremos que la mediana. SOLUCIÓN: PÁG. 341
- 0.10** La media de un conjunto de datos numéricos x_1, \dots, x_n es el número m tal que $(\sum_{i=1}^n x_i - m)^2 \leq (\sum_{i=1}^n x_i - a)^2$, para cada número real a .
- 0.11** El percentil 10 de un conjunto de 10 datos numéricos dos a dos distintos es el menor de esos datos.

- 0.12** Los conjuntos de datos x_1, \dots, x_n y $2x_1, \dots, 2x_n$ tienen la misma desviación típica. SOLUCIÓN: PÁG. 341
- 0.13** La mediana de un conjunto de datos numéricos proporciona una buena medida de la dispersión del mismo.
- 0.14** De un grupo de 10 voluntarios para participar en un ensayo clínico para valorar la eficacia de un cierto tratamiento médico, se desean elegir 5 para el grupo de tratamiento (los 5 restantes constituirán el grupo de control). Hay 252 formas de hacerlo.

Problemas del Capítulo 0

- 0.1** Consideremos los conjuntos $A = \{1, 2, 4, 8, 16, 32\}$, $B = \{n \in \mathbb{N}: 1 \leq n \leq 10\}$, $C = \{a, b, c\}$ y $D = \{n \in \mathbb{N}: 2n^2 \in A\}$. Describe explícitamente los conjuntos $A \cap B$, $A \cup B$, $A \times C$, $B \setminus A$, C^2 y D .
- 0.2** Un hospital tiene siete plantas con 2 alas cada una y 10 habitaciones para acoger enfermos en cada ala; en cada habitación hay una cama al lado de la puerta y otra al lado de la ventana. a) ¿Cuántas camas hay en el hospital? b) Inspirándote en la idea de conjunto producto, diseña una forma de etiquetar las camas del hospital que nos permita ubicar su posición inmediatamente.
- 0.3** (★) En un cierto hospital existen 3 consultas externas de traumatología y en plantilla existen 7 traumatólogos.
- a) ¿Cuántas planillas diferentes pueden hacerse para traumatología si las consultas están igualmente equipadas?
- b) ¿Cuántas planillas diferentes podrían hacerse si las consultas fuesen 2 a 2 distintas?
- c) ¿Cuántas planillas diferentes podrían hacerse si dos de las consultas fuesen idénticas y diferentes de la tercera?
- d) ¿Cuántas planillas diferentes podrían hacerse si además de atender las 3 consultas, hay que atender 2 alas de la planta de traumatología, cada una de las cuales necesita un traumatólogo? Suponemos que a los traumatólogos les da igual ocupar una consulta que otra, o atender un ala u otra de la planta, pero no les da igual ir a consulta o a planta. **SOLUCIÓN: PÁG. 342**
- 0.4** (★) El gráfico siguiente contiene los diagramas de caja de 5 conjuntos de datos de 10 datos cada uno. a) ¿Qué información puedes extraer de cada uno de esos diagramas sobre los conjuntos de datos respectivos?
- b) Construye 5 conjuntos de datos con 10 datos cada uno cuyos diagramas de caja correspondan con los de la siguiente figura.



Capítulo **1**

Introducción al Cálculo de Probabilidades

1.1. Espacios de Probabilidad

1.1.1. Fenómenos aleatorios

Como advertíamos en el tema anterior si, tras la obtención de los datos, el investigador sólo está interesado en resumirlos y presentarlos de forma adecuada, las técnicas de la estadística descriptiva son más que suficientes. Si, en cambio, pretende obtener de esos datos conclusiones generales sobre una clase amplia de individuos del mismo tipo de los que han sido estudiados, entonces las técnicas descriptivas sólo son el comienzo del análisis. Para obtener conclusiones válidas y hacer predicciones correctas sobre una población a partir de la observación de una parte de la misma necesitaremos los métodos de la inferencia estadística. La herramienta básica de la inferencia estadística es el cálculo de probabilidades que introducimos en este tema.

Debe tenerse presente que la inferencia estadística parte de la información proporcionada por los datos de la muestra para tomar una decisión sobre toda la población. Puesto que no hemos estudiado toda la población, nunca tendremos la certeza absoluta sobre si la decisión finalmente adoptada es o no correcta; así pues, en un problema de inferencia estadística ni siquiera tiene sentido preguntarse sobre si hemos cometido o no un error, y lo más que podemos preguntarnos es por la probabilidad de cometer un error. La mayor parte de los métodos estadísticos que presentaremos a lo largo del libro han sido diseñados con el objetivo de que la probabilidad de cometer un error sea lo más pequeña posible. Posteriormente volveremos sobre esta cuestión, pero de las últimas afirmaciones (en las que se habla de probabilidad en términos intuitivos) se desprende la importancia del concepto de probabilidad en estadística.

Tendremos también ocasión de comprobar que el cálculo de probabilidades, además de una herramienta fundamental en la inferencia estadística, tiene sus propias aplicaciones en las ciencias de la salud.

En la naturaleza podemos considerar fenómenos que, realizados en las mismas condiciones, dan lugar a idénticos resultados. Por ejemplo, en el movimiento uniforme, todo aumento de la velocidad (v) de un cuerpo da lugar a un aumento del espacio (e) recorrido por el mismo en un mismo intervalo de tiempo (t) de acuerdo con la ecuación $e = v \cdot t$. La repetición del experimento en diferentes ocasiones siempre conduce al mismo resultado. Son los llamados fenómenos deterministas.

Existen otros fenómenos en la naturaleza en los que no se puede predecir el resultado final, incluso cuando se hace todo lo posible para que las condiciones en que se realiza el experimento sean las mismas, como puede ocurrir en el lanzamiento al aire de una moneda, o en la curación de un enfermo tras la aplicación de un cierto tratamiento. Son los llamados fenómenos aleatorios. A pesar de lo dicho, los fenómenos deterministas tienen también una componente aleatoria debida a la imprecisión de los aparatos de medida utilizados en la experiencia.

1.1.2. Espacio muestral. Sucesos

Si queremos proporcionar una base matemática para el estudio de los fenómenos aleatorios, tendremos que apoyarnos en la teoría de conjuntos como se explica a continuación.

DEFINICIÓN 1.1. (Espacio muestral de un fenómeno aleatorio, sucesos elementales) El conjunto de todos los resultados posibles de un fenómeno aleatorio se llama espacio muestral o espacio de las observaciones, y se denotará por Ω . Cada elemento de Ω se llamará suceso elemental.⁽¹⁾

Ejemplo 1.2. En el lanzamiento de una moneda, el espacio muestral sería $\Omega = \{0, 1\}$ (supongamos que 0=cruz y 1=cara, y eliminemos la posibilidad de que la moneda caiga de canto); en el lanzamiento de un dado, el espacio de las observaciones es $\Omega = \{1, 2, 3, 4, 5, 6\}$ (eliminando la posibilidad de que el dado quede apoyado sobre una de sus aristas o sus vértices); en el experimento aleatorio que consiste en lanzar dos veces al aire una moneda tendríamos $\Omega = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$; si consideramos el fenómeno aleatorio que consiste en la determinación de la cantidad de pacientes que demandan los servicios de urgencia de un hospital en un día, podemos considerar como espacio muestral el conjunto

¹ Se procurará, al menos, que el espacio muestral Ω contenga todos los resultados posibles del experimento.

$\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$, pues esa cantidad será un número natural indeterminado de antemano; si el experimento consiste en la medición del nivel de colesterol de los individuos de una población, podemos considerar como espacio muestral el conjunto \mathbb{R} de todos los números reales.

DEFINICIÓN 1.3. (Suceso) Un suceso es un subconjunto del espacio muestral Ω .

DEFINICIÓN 1.4. (Suceso seguro) A Ω le llamaremos suceso seguro.

DEFINICIÓN 1.5. (Suceso imposible) A \emptyset le llamaremos suceso imposible.

DEFINICIÓN 1.6. (Sucesos incompatibles o disjuntos) Dos sucesos $A, B \subset \Omega$ se dicen incompatibles o disjuntos si $A \cap B = \emptyset$.

Observación 1.7. En un experimento concreto, podemos estar interesados simplemente en un aspecto parcial de la experiencia, lo que nos conduce a la consideración de otros sucesos distintos a los elementales; en el caso del dado, por ejemplo, podríamos estar interesados exclusivamente en si el resultado va a ser un número par, es decir, en el suceso “obtener un número par”; en el caso del servicio de urgencias del hospital, podríamos estar interesados en si la demanda va a superar un cierto número N que se considera como tope para que el servicio sea de calidad, es decir, en el suceso “la demanda ha superado el tope N ”. En el caso del dado sólo nos interesa la ocurrencia del suceso $\{2, 4, 6\}$ (se dice que ha ocurrido un suceso cuando el resultado del experimento es uno de los sucesos elementales que constituyen el suceso), y en el caso del hospital sólo estamos interesados en la ocurrencia del suceso $\{N + 1, N + 2, \dots\}$. Parece razonable identificar un suceso con un subconjunto del espacio muestral Ω . No obstante, llamar suceso a cualquier subconjunto de Ω es matemáticamente apropiado en ciertas ocasiones e inapropiado en otras por razones que quedan muy lejos del alcance de esta obra; a efectos prácticos, debe quedarnos la tranquilidad de que cualquier suceso cuya probabilidad pueda interesarnos queda perfectamente considerado en la teoría matemática que se ha desarrollado para el cálculo de probabilidades. \triangleleft

Observación 1.8. (Operaciones con sucesos) Puesto que, sea cual sea el resultado del experimento, Ω siempre ocurre, parece lógico llamar a Ω suceso seguro. Es lógico también llamar suceso imposible a \emptyset ya que, sea cual sea el resultado del experimento, ese suceso nunca ocurre. Las operaciones conjuntistas habituales de unión, intersección, paso al complementario, \dots , tienen su contrapartida en términos de sucesos. Así, si A y B son dos sucesos, decir que han ocurrido los sucesos A y B simultáneamente es tanto como decir que ha ocurrido el suceso $A \cap B = \{x \in \Omega: x \in A \text{ y } x \in B\}$; decir que ha ocurrido el suceso A o el suceso B es lo mismo que decir que ha ocurrido el suceso $A \cup B = \{x \in \Omega: x \in A \text{ o } x \in B\}$; decir que no ha ocurrido el suceso A es tanto como decir que ha ocurrido el suceso $A^c = \{x \in \Omega: x \notin A\}$ complementario de A . Decir que ocurre A pero no B es tanto como decir que ha ocurrido el suceso diferencia $A \setminus B = A \cap B^c$. Si $A \subset B$ (se dice que A está incluido en B) entonces el suceso B ocurre siempre que ocurra el suceso A . Que dos sucesos A y B sean incompatibles (o disjuntos) significa que no pueden ocurrir simultáneamente. \triangleleft

1.1.3. Definición de probabilidad. Propiedades

En un experimento aleatorio concreto podemos estar interesados en calcular la probabilidad de que ocurra un cierto suceso. Podemos pensar, de momento, en la probabilidad de un suceso como una medida del grado de incertidumbre que poseemos sobre la ocurrencia del mismo en una realización futura del experimento; esa medida debe satisfacer ciertas propiedades que veremos posteriormente y que son análogas a las que verifica la medida del área de figuras en el plano (no en vano, el matemático ruso A.N. Kolmogorov fijó la misma base matemática para el cálculo de probabilidades que para la parte de la matemática que llamamos Teoría de la Medida, entre cuyos objetivos se encuentra el cálculo de longitudes, áreas, volúmenes, etc).

Empíricamente, el concepto de probabilidad surge como una abstracción de la idea de frecuencia relativa de un suceso en una larga serie de pruebas independientes del experimento. Por ejemplo, si lanzamos al aire una moneda un cierto número n de veces, la frecuencia relativa del suceso cara en esos n lanzamientos (es decir, el cociente entre el número $c(n)$ de caras obtenidas y el número n de pruebas realizadas) nos da una idea de la probabilidad de sacar cara y puede constatarse empíricamente que, cuando n se hace más y más grande, la frecuencia relativa $c(n)/n$ del suceso “cara” tiende a estabilizarse alrededor de un número que conocemos como probabilidad del suceso cara. En la figura siguiente, los tres primeros gráficos de la Figura 1.1 muestran la evolución de la frecuencia relativa del suceso cara en lanzamientos independientes sucesivos (en realidad, con la ayuda de un ordenador, se ha simulado el lanzamiento de una moneda perfecta 1000 veces y esos tres primeros gráficos muestran la evolución de la frecuencia relativa del suceso cara en los 10, 100 y 1000 primeros lanzamientos, respectivamente) de una moneda perfecta, mientras que el cuarto gráfico recoge la frecuencia relativa en 1000 nuevas realizaciones del experimento; nótese cómo la frecuencia relativa del suceso cara, cuando el número de lanzamientos aumenta, tiende a estabilizarse en torno al valor $1/2$, probabilidad de sacar cara.

En general, ocurre que, para cualquier fenómeno aleatorio y cualquier suceso A , la frecuencia relativa del suceso A tiende a un cierto número $P(A)$ cuando el número de pruebas del experimento tiende a infinito; el límite $P(A)$ se llama probabilidad del suceso

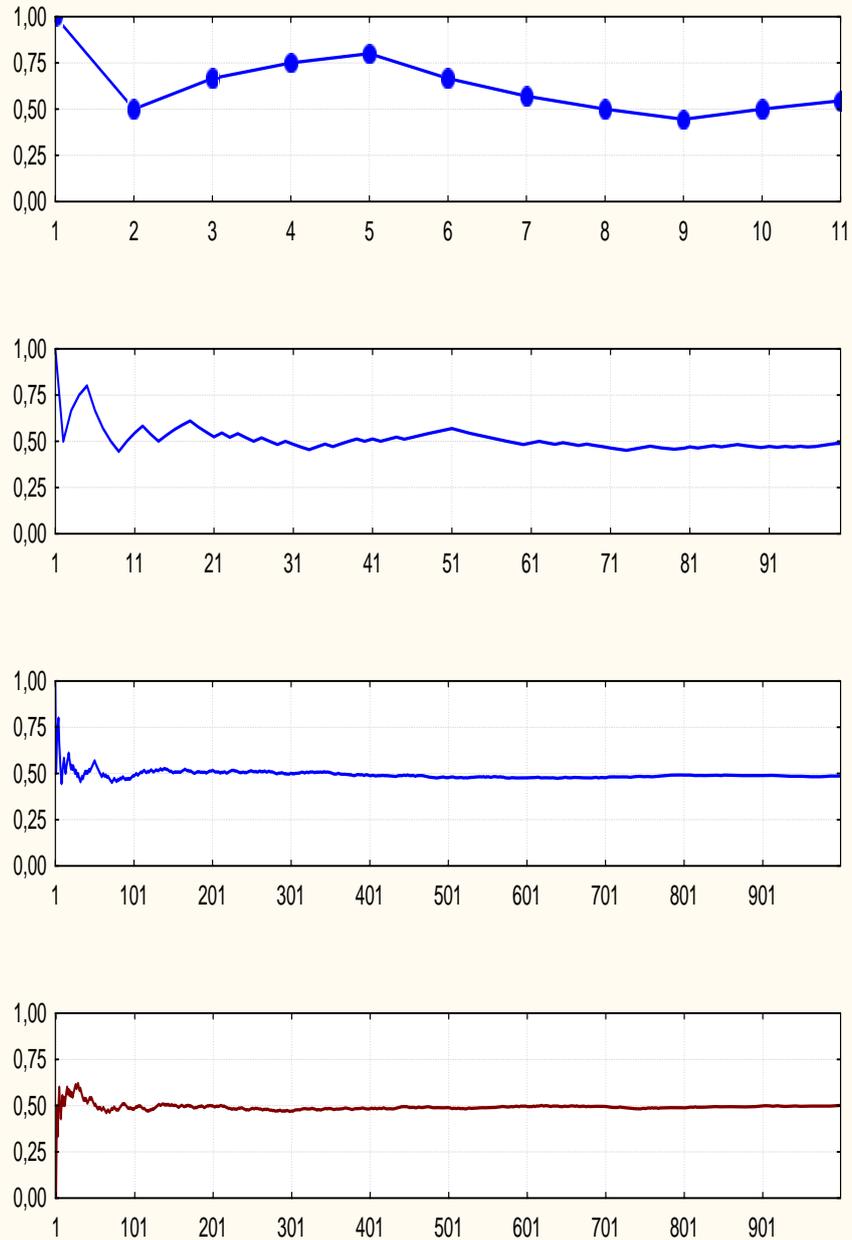


Figura 1.1: Frecuencia relativa suceso cara en función del número de lanzamientos

A. Ésta es la llamada definición frecuentista de la probabilidad. Más que una definición rigurosa, que, desde un punto de vista formal, presentaría ciertos inconvenientes que no vamos a comentar, debe considerarse como un hecho comprobable experimentalmente y, al mismo tiempo, como un resultado (consecuencia de la llamada ley de los grandes números) de la teoría matemática desarrollada por Kolmogorov para el cálculo de probabilidades.

Esta aproximación frecuentista de la probabilidad puede servir de punto de partida intuitivo para una aproximación matemáticamente más apropiada como es la llamada definición axiomática de probabilidad debida a Kolmogorov que veremos posteriormente.

DEFINICIÓN 1.9. (Definición axiomática de probabilidad) Una probabilidad en un conjunto Ω es una aplicación P que asigna a cada suceso $A \subset \Omega$ un número $P(A)$ de forma que se verifiquen los tres axiomas siguientes:

I. $P(A) \geq 0$.

II. $P(\Omega) = 1$.

III. (Aditividad) Si A y B son sucesos incompatibles (o disjuntos) entonces

$$P(A \cup B) = P(A) + P(B)$$

Observación 1.10. quede claro que esos tres axiomas son propiedades que exigimos a una aplicación o función de sucesos como P para que sea una probabilidad; como hemos dicho anteriormente quedan intuitivamente justificados por la aproximación frecuentista a la probabilidad del siguiente modo: el axioma I queda justificado por el hecho de que la frecuencia relativa es siempre un número positivo; el segundo se justifica porque el suceso seguro ocurrirá tantas veces como realicemos el experimento y, por tanto, la frecuencia relativa del suceso seguro es siempre 1; el tercer axioma queda justificado porque, según se prueba sin dificultad, la frecuencia relativa de la unión de dos sucesos incompatibles es la suma de las frecuencias relativas de cada uno de ellos. Nótese que, salvo en el axioma II, la definición anterior puede servir también para el cálculo del área de figuras planas (o de volúmenes de cuerpos en el espacio), ya que el área es siempre positiva y el área de la unión de dos figuras planas disjuntas es la suma de las áreas de cada una de ellas. \triangleleft

Observación 1.11. Si consideramos, como habíamos dicho, la probabilidad de un suceso como una medida del grado de incertidumbre que tenemos sobre la ocurrencia del mismo en una realización del experimento, entonces esa medida es un número entre 0 y 1 de forma que sucesos con probabilidad próxima a cero se consideran poco probables mientras que sucesos con probabilidad próxima a 1 se consideran como muy probables, y ante sucesos con probabilidad $1/2$ consideraremos que es igualmente probable que ocurra como que no ocurra. \triangleleft

Observación 1.12. En realidad, razones matemáticas convierten el tercer axioma de la definición de probabilidad en algo más complejo que el que hemos considerado, pues exige la propiedad de aditividad no sólo para el caso de dos sucesos incompatibles, sino también para una cantidad numerable de sucesos incompatibles. La distinción entre ambos axiomas queda nuevamente lejos de los propósitos de este libro. \triangleleft

Haciendo uso de esos tres axiomas podemos deducir una serie de propiedades de la probabilidad que recogemos a continuación (para las que también podemos usar la interpretación frecuentista y establecer su analogía con el cálculo de áreas):

Proposición 1.13. (P1) *Para cualquier suceso A ,*

$$P(A^c) = 1 - P(A)$$

(P2) *La probabilidad del suceso imposible \emptyset es 0.*

(P3) *Si $A \subset B$, $P(A) \leq P(B)$.*

(P4) *Si A_1, \dots, A_n son sucesos dos a dos incompatibles entonces*

$$P(A_1 \cup \dots \cup A_n) = \sum_{i=1}^n P(A_i)$$

(P5) *Para dos sucesos cualesquiera A y B se verifica*

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

(P6) *La probabilidad de cualquier suceso es un número entre 0 y 1.*

(P7) $P(A \cup B) = 1 - P(A^c \cap B^c)$.

(P8) (★) $P(A_1 \cup \dots \cup A_n) = 1 - P(A_1^c \cap \dots \cap A_n^c)$.

(P9) (★) (Desigualdad de Bonferroni) *Si A_1, \dots, A_n son sucesos arbitrarios, entonces*

$$P(A_1 \cup \dots \cup A_n) \leq \sum_{i=1}^n P(A_i)$$

Ejemplo 1.14. (Definición de probabilidad de Laplace) La definición clásica de probabilidad debida a Laplace considera únicamente el caso en que el espacio muestral es finito (supongamos que $\Omega = \{\omega_1, \dots, \omega_n\}$) y en el que todos los sucesos elementales son equiprobables, es decir, $P(\omega_i) = 1/n$, para cada $1 \leq i \leq n$. La probabilidad de un suceso en este caso se obtiene mediante la fórmula de Laplace, es decir,

$$\frac{\text{número de casos favorables}}{\text{número de casos posibles}}$$

donde se consideran favorables los elementos de ese suceso y posibles todos los elementos de Ω ; eso es consecuencia inmediata de la propiedad (P4). En el caso del lanzamiento de un dado, admitiremos que éste es perfecto si todos los sucesos elementales tienen la misma probabilidad y, puesto que esas probabilidades tienen que sumar 1, todas ellas son iguales a $1/6$. La probabilidad de sacar un número par es entonces $3/6=1/2$ y la de sacar un dos o un 5 es $2/6=1/3$. Si ese dado se lanza dos veces, existen 36 resultados posibles $((1,1), \dots, (1,6), \dots, (6,1), \dots, (6,6))$ todos ellos equiprobables (con probabilidad $1/36$), y la probabilidad de que la suma de ambos resultados sea igual a 7 es $6/36=1/6$, mientras que la probabilidad de que la suma sea igual a 6 es $5/36$. En el caso de dos lanzamientos de una moneda perfecta (en la que la probabilidad de sacar cara es igual a la de sacar cruz e igual a $1/2$) existen cuatro resultados posibles $((0,0), (0,1), (1,0)$ y $(1,1))$ equiprobables (cada uno con probabilidad $1/4$), y la probabilidad de que se obtenga una y sólo una cara es $2/4=1/2$, mientras que la probabilidad de que se obtengan dos cruces es $1/4$. Esas probabilidades se calculan sin dificultad haciendo uso de la fórmula de Laplace; nótese que la hipótesis de equiprobabilidad es fundamental en esos ejemplos y que esa hipótesis no se puede admitir sin una justificación apropiada: por ejemplo, elegido un individuo al azar de una población puede ocurrir que tenga una cierta enfermedad o no la tenga pero la hipótesis de equiprobabilidad (la mitad de los individuos de la población están enfermos) en este caso no parece razonable. Veamos un ejemplo más concreto: supongamos conocido que en España el 41 % de la población son del grupo sanguíneo A, el 9 % son del grupo B, el 4 % son del grupo AB y el 46 % restante son del grupo O; si pretendemos calcular la probabilidad de que un individuo elegido al azar en España sea de alguno de los tipos A, B o AB, el espacio muestral correspondiente sería $\Omega = \{A, B, AB, O\}$ en el que los sucesos elementales no son equiprobables (lo que nos impide usar la definición de Laplace); para calcular la probabilidad buscada basta hacer uso de la propiedad (P4) para obtener $P(\{A, B, AB\}) = 0.41 + 0.09 + 0.04 = 0.54$. \triangleleft

DEFINICIÓN 1.15. (Espacio de probabilidad) Un espacio de probabilidad es un par (Ω, P) formado por un conjunto no vacío Ω y una probabilidad P sobre él.

Observación 1.16. Así pues, a la hora de calcular probabilidades conviene tener claro cuál es el espacio muestral e identificar perfectamente el suceso que nos interesa con un subconjunto del mismo (posteriormente veremos algunos modelos probabilísticos teóricos que se utilizan como modelos matemáticos para el estudio de ciertos fenómenos aleatorios, lo que nos permitirá calcular probabilidades de sucesos relacionados con ese fenómeno). Basta ya conocer las probabilidades de los sucesos elementales si el espacio muestral es finito o numerable (usando la definición de Laplace en el caso finito con sucesos elementales equiprobables)

o conocer el modelo de probabilidad teórico que describe el fenómeno. Haciendo uso de las propiedades de la probabilidad podremos calcular probabilidades de ciertos sucesos a partir de las probabilidades conocidas de otros sucesos. Debe quedar claro en cualquier caso que, para calcular probabilidades, necesitamos un espacio muestral Ω y una probabilidad P sobre él; así pues, desde un punto de vista matemático, el par (Ω, P) contiene todos los ingredientes necesarios para calcular probabilidades de sucesos en Ω . \triangleleft

Observación 1.17. Las siguientes afirmaciones han sido realizadas en el ejemplo anterior: (a) “la probabilidad de sacar cara en un lanzamiento de una moneda perfecta es igual a $1/2$ ”, (b) “la probabilidad de sacar una cara en dos lanzamientos de esa misma moneda es igual a $1/2$ ” y (c) “la probabilidad de que un español elegido al azar sea del grupo sanguíneo A es igual a 0.41”. Estas expresiones, y otras similares, pueden llevarnos erróneamente a pensar que existe una probabilidad universal que lo rige todo y que nos permite calcular probabilidades de que ocurran ciertos sucesos en cualquier fenómeno aleatorio que se nos ocurra; el hecho de utilizar la misma letra P para designar la probabilidad en distintos fenómenos aleatorios puede también contribuir a ello. No obstante, conviene tener claro que a fenómenos aleatorios distintos se deben asignar espacios de probabilidad distintos, en el sentido de que, o bien los espacios muestrales asociados son distintos (en cuyo caso las probabilidades asociadas también lo son necesariamente), o bien los espacios muestrales coinciden pero las probabilidades correspondientes son distintas. Y esa distinción debe quedar reflejada en la notación cuando haya lugar a confusión. Por ejemplo, en el caso de la afirmación (a) estamos considerando el espacio de probabilidad (Ω_1, P_1) , donde $\Omega_1 = \{0, 1\}$ y $P_1(0) = P_1(1) = 1/2$, pero si la moneda no es perfecta y la probabilidad de sacar cara es, digamos, $1/3$, el espacio de probabilidad sería (Ω_1, P_2) , donde $P_2(0) = 2/3$ y $P_2(1) = 1/3$; en la afirmación (b) estamos considerando el espacio de probabilidad (Ω_3, P_3) , donde $\Omega_3 = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$ y $P_3(i, j) = 1/4$, $i, j = 0, 1$; el espacio de probabilidad correspondiente a la afirmación (c) es $(\{A, B, AB, O\}, P_4)$, donde $P_4(A) = 0.41$, $P_4(B) = 0.09$, $P_4(AB) = 0.04$ y $P_4(O) = 0.46$. \triangleleft

Observación 1.18. (★) En textos con un contenido matemático más riguroso, el espacio de probabilidad contiene un tercer ingrediente: la familia de los subconjuntos de Ω que serán considerados sucesos. Recordemos que nosotros hemos decidido llamar suceso a cualquier subconjunto de Ω , aun a sabiendas de que eso no es siempre matemáticamente adecuado. \triangleleft

1.2. Dependencia e Independencia de Sucesos.

El Teorema de Bayes:

Aplicaciones al Diagnóstico Clínico

1.2.1. Probabilidad condicionada

Ahora consideramos dos sucesos A y B y pretendemos definir la probabilidad del suceso A condicionada por el suceso B o probabilidad condicionada de A dado B , que denotaremos por $P(A|B)$. Se busca un número entre 0 y 1 que mida la probabilidad del suceso A supuesto conocido que ha ocurrido el suceso B .

DEFINICIÓN 1.19. (Probabilidad condicionada de un suceso respecto a otro) Dados dos sucesos A y B tales que $P(B) > 0$, se define la probabilidad condicionada de A respecto a B por

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Ejemplo 1.20. Consideremos una población formada por 1000 personas y, en ella, el suceso A formado por aquellas que son del grupo sanguíneo A y el suceso V formado por todos los varones de esa población. Supongamos que hay 600 varones (suceso V) y 400 mujeres (suceso M), y que 200 varones y 180 mujeres son del grupo sanguíneo A. Mientras que la probabilidad de que una persona elegida al azar de esa población sea del grupo A es $P(A) = (200 + 180)/1000 = 0.38$, la probabilidad de que un varón elegido al azar sea de ese grupo es $P(A|V) = 200/600 = 1/3$. Análogamente, $P(A|M) = 180/400 = 0.45$. ◀

Se sigue inmediatamente de la definición anterior la siguiente propiedad:

Proposición 1.21. (P10) $P(A \cap B) = P(B) \cdot P(A|B) = P(A) \cdot P(B|A)$.

Ejemplo 1.22. Consideremos el experimento aleatorio que consiste en la extracción de dos cartas de una baraja española,⁽²⁾ y consideremos los sucesos A : “la primera carta es un as” y

² 40 cartas divididas en cuatro palos -oros, copas, espadas y bastos- con 10 cartas cada uno -as, dos, tres, . . . , siete, sota, caballo y rey-.

B : “la segunda carta es un as”. El experimento se puede realizar con o sin reemplazamiento; en el primer caso, se extrae la primera carta, se mira y se devuelve a la baraja, de tal suerte que puede volver a ser extraída en una segunda ocasión, y en el caso sin reemplazamiento la primera carta no se devuelve a la baraja. En el caso con reemplazamiento, se verifica que

$$P(A \cap B) = P(A) \cdot P(B|A) = (4/40) \cdot (4/40) = 1/100$$

en el caso sin reemplazamiento se tiene que

$$P(A \cap B) = P(A) \cdot P(B|A) = (4/40) \cdot (3/39) = 12/1560$$

En realidad se trata de dos fenómenos aleatorios diferentes con espacios de probabilidad diferentes. Descríbanse como ejercicio esos dos espacios de probabilidad. \triangleleft

1.2.2. Independencia de sucesos

Relacionado con la definición de probabilidad condicional, aparece el concepto de independencia de dos sucesos, que presentamos a continuación.

DEFINICIÓN 1.23. (Independencia de sucesos) Dos sucesos A y B se dicen independientes si

$$P(A \cap B) = P(A) \cdot P(B)$$

Si A y B son independientes y $P(B) \neq 0$ entonces se verifica que $P(A|B) = P(A)$.

Ejemplo 1.24. Continuando con un ejemplo anterior, hemos afirmado que, en general, la proporción de hombres que son del grupo sanguíneo A no tiene por qué ser la misma que la de personas que son de ese grupo sanguíneo. Eso ocurrirá cuando y sólo cuando los sucesos “ser hombre” y “tener el grupo sanguíneo A” sean independientes en esa población de 1000 personas. Dos sucesos no independientes se dicen dependientes, pues a la hora de calcular la probabilidad de A no da lo mismo condicionar o no condicionar respecto al suceso B . \triangleleft

Ejemplo 1.25. Supongamos conocido que, entre los individuos adultos de una cierta población, el 45 % son fumadores (suceso A), el 20 % son hipertensos (suceso B) y el 50 % son fumadores o hipertensos (suceso $A \cup B$). Nos preguntamos si ser fumador es independiente de ser hipertenso en esa población. Puesto que $P(A \cap B) = P(A) + P(B) - P(A \cup B) = 0.45 + 0.20 - 0.50 = 0.15$, se sigue que $P(B|A) = 0.1/0.4 = 0.33$. Siendo $P(B) = 0.20$ se tiene que $P(B|A) \neq P(B)$ y A y B no son independientes. De hecho se ha visto que en la población total hay un 20 % de hipertensos mientras que entre los fumadores hay un 33 % de hipertensos, con lo cual existe una dependencia entre ambos sucesos (el hecho de que una persona sea fumadora aumenta la probabilidad de tener hipertensión). \triangleleft

Ejemplo 1.26. En el ejemplo del lanzamiento de un dado perfecto se comprueba fácilmente que los sucesos $\{1, 2\}$ y $\{1, 3\}$ no son independientes, pero sí lo son los sucesos $\{1, 2\}$ y $\{1, 3, 4\}$. \triangleleft

Ejemplo 1.27. En el experimento aleatorio que consiste en lanzar una moneda perfecta al aire dos veces, los sucesos $\{(0, 0), (0, 1)\}$ y $\{(0, 0), (1, 0)\}$ son independientes. Es decir, los sucesos “sacar cruz en el primer lanzamiento” y “sacar cruz en el segundo lanzamiento” son independientes. \triangleleft

Ejemplo 1.28. El concepto de independencia jugará un papel determinante en el uso de los métodos estadísticos que estudiaremos en esta obra. De los comentarios precedentes se deduce que dos sucesos son independientes cuando la ocurrencia de uno de ellos no condiciona en modo alguno la ocurrencia del otro en lo que al cálculo de sus probabilidades se refiere. Una situación típica en la que se tiene independencia es la que describe el ejemplo siguiente: supongamos que el experimento aleatorio consiste en dos lanzamientos al aire de un dado y consideremos dos sucesos A y B de forma que A sólo depende del resultado de la primera tirada (por ejemplo: el resultado del primer lanzamiento es un 1 o un 2) mientras que B sólo depende del resultado de la segunda tirada (por ejemplo: en el segundo lanzamiento se obtendrá un uno o un tres); puesto que el resultado que se obtenga en la primera tirada no condiciona el que se obtenga en la segunda, los sucesos A y B son independientes; nótese, en efecto, que para el experimento considerado se tiene

$$\begin{aligned}\Omega &= \{(1, 1), \dots, (1, 6), \dots, (6, 1), \dots, (6, 6)\}, \\ A &= \{(1, 1), \dots, (1, 6), (2, 1), \dots, (2, 6)\}, \quad P(A) = 12/36 = 1/3 \\ B &= \{(1, 1), \dots, (6, 1), (1, 3), \dots, (6, 3)\}, \quad P(B) = 12/36 = 1/3 \\ A \cap B &= \{(1, 1), (1, 3), (2, 1), (2, 3)\}, \quad P(A \cap B) = 4/36 = P(A) \cdot P(B)\end{aligned}$$

que prueba que A y B son independientes, como habíamos afirmado. Hemos comprobado en el ejemplo 1.26 anterior que al lanzar una vez un dado perfecto los sucesos “sacar un 1 o un 2” y “sacar un 1 o un 3” no son independientes y acabamos de ver que al lanzar dos veces un dado perfecto los sucesos “sacar un 1 o un 2 en el primer lanzamiento” y “sacar un 1 o un 3 en el segundo lanzamiento” sí son independientes. \triangleleft

Observación 1.29. En el ejemplo 1.27 anterior uno de los sucesos sólo depende del primer lanzamiento de la moneda y el otro sólo depende del segundo lanzamiento, y los sucesos son independientes. Ésta es una situación más general de lo que en principio pueda parecer, pues, análogamente, se verifica que, si un mismo experimento aleatorio se realiza en varias ocasiones y consideramos dos sucesos que dependen de los resultados obtenidos en diferentes realizaciones del experimento, entonces esos dos sucesos son independientes. Por ejemplo, supongamos que elegimos al azar y con reemplazamiento dos individuos de una población; los sucesos “el primer individuo tiene nivel de colesterol mayor que 200” y “el segundo individuo tiene nivel de colesterol mayor que 200” son independientes, pues el primero de ellos sólo depende de la primera realización del experimento (extraer un individuo de la población) y el segundo sólo depende de la segunda. Sin embargo, los sucesos “el primer individuo tiene nivel de colesterol mayor que 200” y “el primer individuo tiene los ojos azules”, en

general, no serán independientes, pues la independencia en ese caso vendría a decir que las proporciones de individuos con nivel de colesterol mayor que 200 en toda la población y en la subpoblación formada por los individuos con ojos azules coinciden, cosa que, en general, no tiene por qué ser cierta. \triangleleft

1.2.3. Teorema de Bayes. Aplicación al diagnóstico clínico

Veamos a continuación el llamado teorema de Bayes, del que extraeremos una aplicación al diagnóstico clínico.

TEOREMA 1.30. (Teorema de Bayes) Sean A_1, \dots, A_n sucesos de probabilidad no nula, dos a dos disjuntos y que recubren Ω , es decir, $\Omega = A_1 \cup \dots \cup A_n$. Sea B otro suceso de probabilidad no nula. Entonces, si $1 \leq k \leq n$,

$$P(A_k|B) = \frac{P(A_k) \cdot P(B|A_k)}{\sum_{i=1}^n P(A_i) \cdot P(B|A_i)}$$

Este resultado es consecuencia inmediata del siguiente

TEOREMA 1.31. (Teorema de la probabilidad total) Sean A_1, \dots, A_n sucesos dos a dos disjuntos y que recubren Ω . Sea B otro suceso. Entonces,

$$P(B) = \sum_{i=1}^n P(A_i) \cdot P(B|A_i)$$

Ejemplo 1.32. Una ciudad está dividida en 2 zonas de salud A y B que atienden al 40% y el 60% de la población, respectivamente. Un estudio llevado a cabo en ambas zonas de salud revela que un 30% de los individuos de la zona A y un 40% de los de la zona B fueron infectados por el virus de la gripe estacional en 2010. Nos preguntamos por la probabilidad de que un individuo elegido al azar en esa ciudad hubiera tenido el virus de la gripe en 2010: denotemos por G el suceso correspondiente. De acuerdo con el teorema de la probabilidad total:

$$P(G) = P(G \cap A) + P(G \cap B) = P(A) \cdot P(G|A) + P(B) \cdot P(G|B) = 0'4 \cdot 0'3 + 0'6 \cdot 0'4 = 0'36,$$

es decir, un 36% de los habitantes de la ciudad fueron infectados por el virus de la gripe estacional en 2010. Nótese como la probabilidad buscada es una media ponderada de las probabilidades 0'3 y 0'4 de ser infectado por el virus en las zonas A y B , donde los pesos utilizados en la ponderación, 0'4 y 0'6, son las probabilidades de que el individuo pertenezca a cada una de las zonas de salud consideradas. Concretando algo más, si la ciudad tiene

100000 habitantes, 40000 pertenecen a la zona A y 60000 a la B ; 12000 individuos de la zona A (el 30%) y 24000 de la zona B (el 40%) fueron infectados por el virus de la gripe estacional en 2010. Por tanto, un total de 36000 habitantes de la ciudad (un 36%) fueron infectados por el virus en 2010.

Si se elige un individuo de esa ciudad que fue infectado por el virus de la gripe en 2010, el teorema de Bayes nos permite calcular la probabilidad de que proceda de la zona de salud A :

$$P(A|G) = \frac{P(A) \cdot P(G|A)}{P(A) \cdot P(G|A) + P(B) \cdot P(G|B)} = \frac{0'4 \cdot 0'3}{0'4 \cdot 0'3 + 0'6 \cdot 0'4} = \frac{1}{3}.$$

En la práctica, el médico se cuestiona si un determinado paciente tiene una cierta enfermedad. Llamemos E al suceso “el paciente tiene esa enfermedad” y E^c al suceso complementario “el paciente no tiene la enfermedad”. Antes de examinar al paciente el médico podrá afirmar que, por ejemplo, dicho paciente padece esa enfermedad con una probabilidad $P(E) = 0.01$, si es conocido por medio de ciertas estadísticas sanitarias que el 1% de los individuos de la población general tienen la enfermedad.

Posteriormente, el médico realiza al paciente una serie de pruebas o test de diagnóstico anotando si el test da positivo (suceso T) o negativo (suceso T^c). Tras examinar al paciente, el médico puede modificar su opinión a priori y se interesa por la probabilidad condicionada $P(E|T)$ de que el paciente posea la enfermedad una vez conocido que el test ha resultado positivo o por la probabilidad $P(E^c|T^c)$ de que no posea la enfermedad si el test ha resultado negativo (usualmente, si el conjunto de síntomas que define el test tiene algo que ver positivamente con la enfermedad, tendremos $P(E) < P(E|T)$). Para calcular esas probabilidades condicionadas, utilizaremos el teorema de Bayes.

Puesto que el método de diagnóstico no es normalmente perfecto, el test puede resultar positivo o negativo tanto en los individuos que poseen la enfermedad como en los que no la poseen. En ese sentido, se introducen las cantidades siguientes:

DEFINICIÓN 1.33. (a) A $P(E)$ se le llama prevalencia de la enfermedad.

(b) La probabilidad condicionada $P(T^c|E)$ se llama probabilidad de falso negativo, mientras que $P(T|E)$ se conoce como sensibilidad del test de diagnóstico.

(c) La probabilidad condicionada $P(T|E^c)$ se llama probabilidad de falso positivo mientras que $P(T^c|E^c)$ se conoce como especificidad del test de diagnóstico.

Observación 1.34. Nótese que

$$\begin{aligned} \text{sensibilidad} + \text{probabilidad de falso negativo} &= 1 \\ \text{especificidad} + \text{probabilidad de falso positivo} &= 1 \end{aligned}$$

Es claro que un test de diagnóstico con una sensibilidad alta (pocos falsos negativos) resulta útil para descartar la enfermedad, mientras que un test con una especificidad alta (pocos falsos positivos) resulta útil para confirmar la enfermedad. Así pues, en una primera etapa del proceso de diagnóstico de una enfermedad necesitaremos tests muy sensibles que nos permitan decidir con fiabilidad si podemos descartar la enfermedad; si no la podemos descartar, en una segunda etapa del proceso utilizaremos tests muy específicos que nos permitan confirmar con fiabilidad la enfermedad. No obstante, como veremos más adelante, en una población concreta, el conocimiento de la prevalencia de la enfermedad nos permitirá determinar los valores predictivos positivo y negativo que nos permitirán graduar la utilidad del test a la hora de confirmar o descartar la enfermedad en esa población. \triangleleft

La prevalencia de una enfermedad suele obtenerse de la literatura médica, mientras que la determinación de la especificidad y de la sensibilidad de un nuevo test de diagnóstico requieren un experimento a propósito. Estos datos resultan útiles al médico y al enfermo para calcular las probabilidades que realmente le interesan: si el test da positivo interesa conocer qué probabilidad existe de que el paciente tenga la enfermedad (es decir, interesa conocer la probabilidad condicionada $P(E|T)$ de que el individuo posea la enfermedad en ese caso), y si el test da negativo interesa conocer qué probabilidad hay de que el paciente esté realmente sano (es decir, interesa conocer la probabilidad condicionada $P(E^c|T^c)$). Estas probabilidades se obtienen a partir del teorema de Bayes como sigue:

Proposición 1.35.

$$\begin{aligned} P(E|T) &= \frac{P(E) \cdot P(T|E)}{P(E) \cdot P(T|E) + P(E^c) \cdot P(T|E^c)} \\ P(E^c|T^c) &= \frac{P(E^c) \cdot P(T^c|E^c)}{P(E^c) \cdot P(T^c|E^c) + P(E) \cdot P(T^c|E)} \end{aligned}$$

DEFINICIÓN 1.36. La probabilidad condicionada $P(E|T)$ se llama valor predictivo positivo del test (o valor predictivo de un test positivo). La probabilidad condicionada $P(E^c|T^c)$ se llama valor predictivo negativo del test (o valor predictivo de un test negativo). Las cantidades $P(E|T) - P(E)$ y $P(E^c|T^c) - P(E^c)$ suelen conocerse como ganancia del positivo y ganancia del negativo.

Observación 1.37. La ganancia del positivo (respectivamente, del negativo) mide el incremento de la probabilidad de padecer la enfermedad si el test dio positivo (respectivamente, el incremento de la probabilidad de no padecer la enfermedad si el test dio negativo) al realizar el test de diagnóstico. \triangleleft

Observación 1.38. (Errores de diagnóstico) Nótese que $P(E^c|T) = 1 - P(E|T)$ es la proporción de diagnósticos positivos erróneos y $P(E|T^c) = 1 - P(E^c|T^c)$ es la proporción de diagnósticos negativos erróneos. \triangleleft

Ejemplo 1.39. Un investigador desea evaluar un cierto test de diagnóstico que él mismo ha diseñado para la detección de una cierta enfermedad. El test ha sido aplicado a 400 pacientes de los que un estudio más profundo revela que son enfermos y a 600 personas no enfermas elegidas independientemente y se han obtenido los resultados que recoge la tabla siguiente:

	Positivo	Negativo	Total
Enfermo	390	10	400
No enfermo	20	580	600

Las columnas recogen el resultado del test y las filas el diagnóstico previo de las personas consideradas. A partir de estos resultados, la sensibilidad del test se estima por $0.975=390/400$, mientras que la especificidad se estima por $0.967=580/600$. El test aparece así como razonablemente sensible y específico, aunque lo de razonablemente depende del uso que se quiere dar al test. Para calcular el valor predictivo positivo del test $P(E|T)$ (la probabilidad de que una persona para la que el test da positivo esté realmente enferma) necesitamos calcular previamente la probabilidad $P(T|E)$ de obtener un resultado positivo en enfermos y la probabilidad $P(T|E^c)$ de obtener un resultado positivo en personas no enfermas:

$$P(T|E) = 390/400 = 0.975, \quad P(T|E^c) = 20/600 = 0.033$$

Entonces

$$P(E|T) = \frac{0.975 \cdot P(E)}{0.975 \cdot P(E) + 0.033 \cdot P(E^c)},$$

$$P(E^c|T^c) = \frac{0.967 \cdot P(E^c)}{0.967 \cdot P(E^c) + 0.025 \cdot P(E)}$$

En las fórmulas precedentes la prevalencia $P(E)$ de la enfermedad permanece desconocida pues no puede ser estimada a partir de los datos de que disponemos: nótese que la cantidad $400/(400+600)$ no es una estimación de esa probabilidad, pues los tamaños de muestra 400 y 600 han sido elegidos por el investigador y las muestras se han seleccionado en grupos de enfermos y no enfermos, respectivamente. Por ejemplo, si suponemos conocido que la prevalencia es $P(E) = 0.1$ entonces $P(E|T) = 0.765$ y $P(E^c|T^c) = 0.997$, es decir, si el test resulta positivo hay una probabilidad de 0.765 de que el individuo posea la enfermedad, y si el test es negativo, la probabilidad de que no la posea es 0.997. La ganancia del positivo para ese test es 0.665 y la ganancia del negativo es 0.097.

Más allá de la sensibilidad y la especificidad de un test de diagnóstico, su posible utilidad a la hora de confirmar o descartar una enfermedad en los individuos de una cierta población dependerá también de su prevalencia en esa población: si $P(E|T)$ y $P(E^c|T^c)$ representan probabilidades de acierto en el proceso de diagnóstico en la población considerada, $P(E^c|T)$ y $P(E|T^c)$ representan probabilidades de error del test en la misma. En el ejemplo considerado, $P(E^c|T) = 1 - 0.765 = 0.235$ (es decir, el 23.5% de los individuos que sean declarados enfermos en esa población estarán realmente sanos) y $P(E|T^c) = 1 - 0.997 = 0.003$ (es decir, el 0.3% de los individuos de esa población que son declarados sanos poseen realmente la enfermedad).⁽³⁾ En ese sentido, en la población concreta considerada, el test propuesto resulta especialmente útil para descartar la enfermedad (sólo un 0.3% de error cuando se descarta la enfermedad), y menos útil a la hora de confirmarla (un 23.5% de error en ese caso).

Quede claro, pues, que este criterio para valorar la utilidad de un test de diagnóstico depende de la población en que se aplique; concretamente, a través de la prevalencia de la enfermedad en la misma. La Figura 1.2 muestra cómo varían los valores predictivos (positivo y negativo) del test de diagnóstico considerado en función de la prevalencia; nótese, por ejemplo, como, a medida que aumenta la prevalencia de la enfermedad en la población, el mismo test de diagnóstico va siendo cada vez más útil para confirmarla y menos para descartarla.

Interesa también que esas ganancias sean grandes. La Figura 1.3 muestra el comportamiento de la ganancia del positivo (línea continua) y del negativo (línea de puntos) frente a la prevalencia. <

Veamos un ejemplo más de aplicación del teorema de Bayes.

Ejemplo 1.40. En un estudio sobre afecciones cardíacas en individuos entre 40 y 50 años de edad en una cierta región, se han considerado como factores de riesgo el hecho de ser fumador (suceso F) y un consumo excesivo de grasas saturadas en la dieta habitual (suceso G), habiéndose detectado que las probabilidades de los sucesos $F^c \cap G^c$, $F \cap G^c$, $F^c \cap G$ y $F \cap G$ son 0.05, 0.1, 0.35 y 0.5 entre los individuos con algún tipo de afección cardíaca (suceso E), y 0.5, 0.3, 0.15 y 0.05 entre los individuos que no poseen problemas cardíacos. Las tablas 2×2 siguientes resumen la distribución conjunta de los dos factores de riesgo en las subpoblaciones E y E^c :

E	F	F^c
G	0.5	0.35
G^c	0.1	0.05

E^c	F	F^c
G	0.05	0.15
G^c	0.3	0.5

La probabilidad de que un individuo con algún tipo de afección cardíaca sea fumador es

$$P(F|E) = P(F \cap G|E) + P(F \cap G^c|E) = 0.5 + 0.1 = 0.6$$

³ El error de declarar enfermo a un individuo sano puede acarrear trastornos derivados del seguimiento de un tratamiento innecesario; más grave parece el error de declarar sano a un individuo que realmente posea la enfermedad.

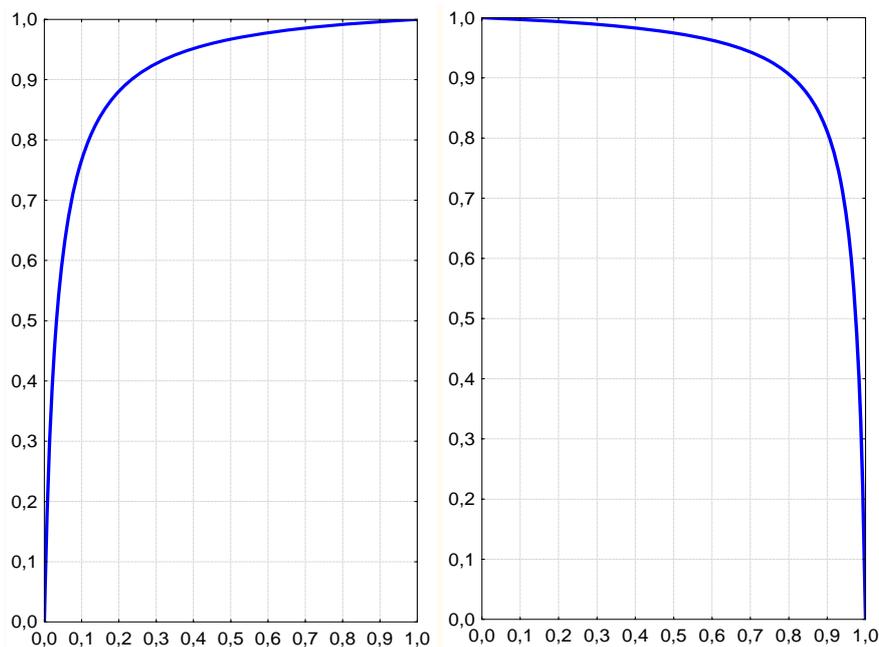


Figura 1.2: Valores predictivos positivo (izquierda) y negativo (derecha) en función de la prevalencia

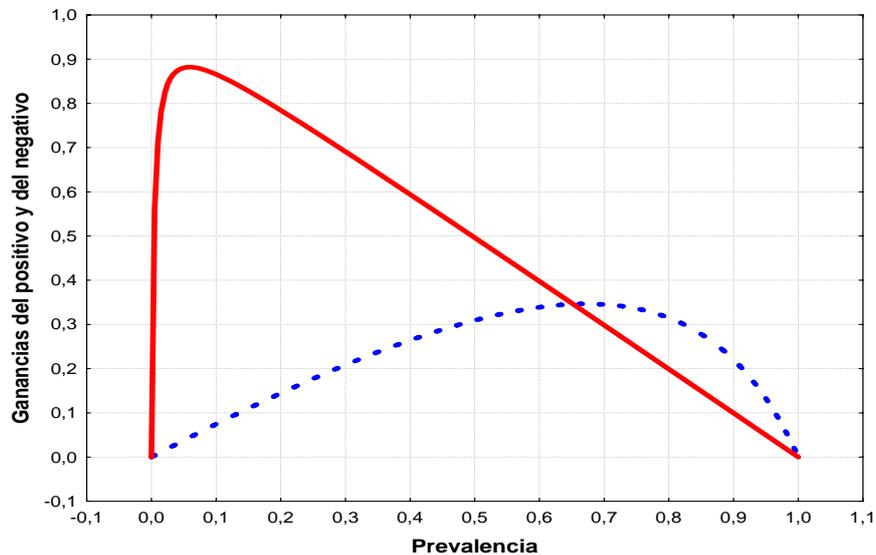


Figura 1.3: Ganancias del positivo (línea continua) y del negativo (línea punteada) en función de la prevalencia

Análogamente, $P(F|E^c) = 0.35$, $P(G|E) = 0.85$ y $P(G|E^c) = 0.2$.

Supongamos conocido que un 10% de los individuos entre 40 y 50 años de esa región posee algún tipo de afección cardíaca; es decir, $P(E) = 0.1$. Entonces, de acuerdo con el teorema de Bayes, la probabilidad de que un fumador tenga algún tipo de afección cardíaca es

$$P(E|F) = \frac{P(E) \cdot P(F|E)}{P(E) \cdot P(F|E) + P(E^c) \cdot P(F|E^c)} = \frac{0.1 \cdot 0.6}{0.1 \cdot 0.6 + 0.9 \cdot 0.35} = 0.16,$$

mientras que la probabilidad de que un consumidor habitual de grasas saturadas tenga ese tipo de enfermedad es

$$P(E|G) = \frac{P(E) \cdot P(G|E)}{P(E) \cdot P(G|E) + P(E^c) \cdot P(G|E^c)} = \frac{0.1 \cdot 0.85}{0.1 \cdot 0.85 + 0.9 \cdot 0.2} = 0.32.$$

Podemos concluir que, entre los individuos entre 40 y 50 años de esa región, consumir grasas saturadas es un factor de riesgo más importante que fumar a la hora de desarrollar algún tipo de afección cardíaca. (Ver Problema I.8) ◁

¿Verdadero o Falso? Capítulo 1

Decidir si cada una de las proposiciones siguientes es verdadera (V) o falsa (F), justificando la respuesta.

- 1.1 Si A y B son sucesos tales que $A \subset B^c$ entonces $P(A|B) = 0$.
- 1.2 En un test de diagnóstico cualquiera se verifica que la suma de su sensibilidad y su especificidad es igual a 1.
- 1.3 La probabilidad de falso negativo de un test de diagnóstico para una enfermedad en una población es la proporción de individuos que poseen la enfermedad para los que el test resulta negativo. **SOLUCIÓN: PÁG. 341**
- 1.4 En una ciudad A de 100000 habitantes, 12000 son hipertensos; un porcentaje de la población tiene factor Rh negativo y, entre éstos últimos, el 10% son hipertensos. En esa ciudad, los sucesos “tener factor Rh negativo” y “ser hipertenso” son independientes.
- 1.5 Cualesquiera que sean los sucesos A y B , se verifica que $P(A|A \cap B) = 1$.
- 1.6 El teorema de Bayes nos permite calcular el valor predictivo positivo de un test de diagnóstico a partir de la prevalencia, la sensibilidad y la especificidad del mismo. **SOLUCIÓN: PÁG. 341**
- 1.7 Si A^c es el complementario de un suceso A entonces $P(A^c) = 1/P(A)$.
- 1.8 La especificidad de un test de diagnóstico para una enfermedad en una población es la proporción de individuos con diagnóstico negativo entre aquellos que no poseen la enfermedad.
- 1.9 La sensibilidad de un test de diagnóstico para una enfermedad en una población es la probabilidad de que un individuo que posee la enfermedad reciba un diagnóstico positivo. **SOLUCIÓN: PÁG. 341**

- 1.10 Cualquier suceso A en un espacio de probabilidad (Ω, P) es independiente del suceso complementario A^c .
- 1.11 Supongamos que utilizamos “fumar” como test de diagnóstico para predecir el cáncer de pulmón en una determinada población. Si sabemos que el 90% de las personas con cáncer de pulmón y que el 55% de las personas sin cáncer de pulmón son fumadores, la especificidad del test es 0.45.
- 1.12 (★) El percentil 10 de un conjunto de 10 datos numéricos dos a dos distintos es el menor de esos datos. SOLUCIÓN: PÁG. 341
- 1.13 Los tests de diagnóstico con una especificidad alta resultan útiles a la hora de confirmar la enfermedad.
- 1.14 Si A y B son sucesos incompatibles entonces $P(A) + P(B) \leq 1$.
- 1.15 Si $A \subset B$, entonces $P(A|B) \geq P(A)$. SOLUCIÓN: PÁG. 341
- 1.16 Cualquier suceso A en el espacio de probabilidad (Ω, P) es independiente del suceso imposible \emptyset .
- 1.17 El valor predictivo negativo de un test de diagnóstico es la probabilidad de que, para un individuo sano, el test resulte positivo.
- 1.18 Si la prevalencia de la enfermedad de Alzheimer en personas mayores de 85 años en una población es del 30%, la probabilidad de que 3 personas de más de 85 años elegidas al azar tengan Alzheimer es 0,27. SOLUCIÓN: PÁG. 341
- 1.19 Si el valor predictivo positivo de un test de diagnóstico para una enfermedad es igual a 0,7, un 30% de los diagnósticos positivos son erróneos.

Problemas del Capítulo 1

- 1.1** Se sabe que entre los individuos que poseen una cierta enfermedad en una población, el 30 % poseen el síntoma S1, el 60 % posee el síntoma S2 y el 72 % posee al menos uno de los dos síntomas.

¿La presencia de uno de los dos síntomas es independiente de la del otro?

- 1.2** En una cierta ciudad hay 3 centros de salud A, B y C con 10000, 9000 y 11000 personas adscritas, respectivamente. Se sabe que son varones el 45 % en el A, el 50 % en el B y el 60 % en el C.

a) ¿Cuál es la probabilidad de que un individuo (adscrito a algún centro de salud) elegido al azar sea varón?

b) Si se elige al azar un individuo entre esos 3 centros de salud y resulta ser mujer, ¿cuál es la probabilidad de que proceda del centro de salud A?

c) Haz lo mismo que en el apartado b) para los centros de salud B y C y decide qué centro de salud es con mayor probabilidad el de procedencia de esa mujer.

d) Si se selecciona una persona al azar de cada centro de salud, ¿cuál es la probabilidad de que las 3 sean varones? ¿Y la de que exactamente dos de las 3 sean varones?

- 1.3** La tabla siguiente contiene las distribuciones categorizadas por sexo de los grupos sanguíneos para los individuos de una población (p. ej., la probabilidad de ser varón y del grupo A es 0.352):

	A	B	AB	O	Total
Varón	0.352	0.064	0.024	0.360	0.80
Mujer	0.054	0.048	0.008	0.090	0.20
Total	0.406	0.112	0.032	0.450	1.00

a) ¿Son independientes los sucesos “Grupo sanguíneo O” y “Varón”?

b) ¿Cuál es la probabilidad de que una mujer sea del grupo A? **SOLUCIÓN: PÁG. 342**

- 1.4** Supongamos que una máquina para medir la presión sanguínea clasifica como hipertensos al 84 % de los individuos hipertensos y al 23 % de los individuos con tensión normal.
- Determina la sensibilidad y la especificidad de la máquina.
 - Asumiendo que el 20 % de los individuos de la población son hipertensos, determina los valores predictivo positivo y predictivo negativo de la máquina.
- 1.5** Supongamos que en un hospital materno A nacen en un día 120 bebés, mientras que en otro hospital B nacen 12 bebés. Admitamos que en ambos la probabilidad de que un bebé sea niño es la misma que la de que sea niña.
- Calcula la probabilidad de que en el hospital A en ese día hayan nacido exactamente el doble de niños que de niñas.
 - Haz lo mismo para el hospital B.
 - Si desconoces el número de niños nacidos en cada uno de los dos hospitales, pero te aseguran que en uno de ellos han nacido exactamente el doble de niños que de niñas, ¿en cuál de los dos hospitales dirías que ha ocurrido eso? *Observación:* En la resolución del apartado (c) has utilizado el llamado principio de máxima verosimilitud en inferencia estadística que, ante la necesidad de tomar una decisión a partir de unos ciertos datos, sugiere tomar la opción que maximiza la probabilidad de obtener el resultado realmente obtenido.⁽⁴⁾
- 1.6** Un test de diagnóstico que ha sido diseñado para la detección del cáncer de colon en una cierta población posee una sensibilidad del 55 % y una especificidad del 80 %.
- Calcula los valores predictivos positivo y negativo del test si la prevalencia de la enfermedad es del 0.5 % en esa población. **SOLUCIÓN: PÁG. 342**

⁴ El principio de máxima verosimilitud que acabamos de describir en un caso especialmente sencillo es una de las justificaciones habituales de muchos métodos estadísticos –estimadores y test de hipótesis–, entre ellos algunos de los que estudiaremos en la segunda parte de este libro.

Capítulo **2**

**Variables Aleatorias. Distribución de
Probabilidad**

2.1. Definición de Variable Aleatoria

2.1.1. Variable aleatoria. Distribución de probabilidad

En el tema anterior hemos definido (axiomáticamente) una probabilidad como una aplicación del conjunto de los sucesos (recordar que hemos identificado un suceso con un subconjunto del espacio muestral Ω) a valores en el intervalo $[0, 1]$ que satisface ciertas propiedades. A una probabilidad en el espacio muestral Ω se le llama también distribución de probabilidad en Ω . Ha quedado claro también que, para calcular probabilidades, necesitamos un espacio muestral Ω y una probabilidad P sobre él, es decir, un espacio de probabilidad (Ω, P) .

En ocasiones, en lugar de interesarnos por el resultado concreto de un experimento aleatorio, sólo nos interesamos por una cierta función de ese resultado, lo que nos conduce al concepto de variable aleatoria.

DEFINICIÓN 2.1. (Variable aleatoria) Dado un espacio de probabilidad (Ω, P) , llamaremos variable aleatoria (brevemente, v.a.) sobre él a una aplicación $X : (\Omega, P) \longrightarrow \Omega'$ que asocia a cada suceso elemental ω un elemento $X(\omega) \in \Omega'$. Si $\Omega' = \mathbb{R}$ diremos que X es una variable aleatoria real (v.a.r.). Cuando $\Omega' = \mathbb{R}^n$ se suele hablar también de vector aleatorio o de variable aleatoria n -dimensional.

DEFINICIÓN 2.2. (Distribución de probabilidad de una v.a.) La distribución de probabilidad (o, simplemente, distribución) de una v.a. $X : (\Omega, P) \longrightarrow \Omega'$ es una nueva probabilidad P^X definida en Ω' sobre un suceso $A' \subset \Omega'$ mediante:

$$P^X(A') = P(X^{-1}(A'))$$

donde $X^{-1}(A') = \{\omega \in \Omega : X(\omega) \in A'\}$. Normalmente escribiremos $P(X \in A')$ en lugar de $P(X^{-1}(A'))$.

Observación 2.3. Muchos libros de bioestadística presentan el concepto de variable aleatoria de una forma confusa, atribuyendo al concepto de variable un significado aparentemente

distinto del concepto matemático de función, y sin distinguir el concepto de variable aleatoria de lo que es la distribución de probabilidad de esa variable. Conviene notar aquí también que la definición que hemos presentado de v.a. no es totalmente correcta desde un punto de vista estrictamente matemático, pero será más que suficiente para los objetivos de esta obra. \triangleleft

Observación 2.4. En la definición anterior hemos utilizado la letra P para la probabilidad considerada en Ω y, de acuerdo con la observación 1.16 de la página 50, una notación diferente P^X para la probabilidad en Ω' que hemos llamado distribución de probabilidad de la v.a. X . \triangleleft

Veamos algunos ejemplos de variables aleatorias y sus distribuciones.

Ejemplo 2.5. Consideremos, como ya hemos hecho anteriormente, una población Ω con N individuos en la que el 41% de ellos son del grupo sanguíneo A, el 9% del grupo B, el 4% del grupo AB y el 46% restante del grupo O. Si P es la probabilidad en Ω que atribuye a cada individuo probabilidad $1/N$, la v.a. $X : (\Omega, P) \rightarrow \{A, B, AB, O\}$ que asigna a cada individuo su grupo sanguíneo tiene distribución de probabilidad P^X definida por $P^X(A) = P(X = A) = 0.41$, $P^X(B) = P(X = B) = 0.09$, $P^X(AB) = P(X = AB) = 0.04$ y $P^X(O) = P(X = O) = 0.46$. Esa distribución de probabilidad P^X nos dice cómo se distribuye el grupo sanguíneo en esa población Ω . \triangleleft

Ejemplo 2.6. Consideremos una población Ω formada por 1000 individuos y consideremos el experimento aleatorio que consiste en elegir uno de ellos al azar; se entiende entonces que tomamos en Ω la probabilidad P para la cual todos los individuos tienen la misma probabilidad de salir elegidos (esa probabilidad será entonces igual a $1/1000$). Consideremos una v.a. $X : \Omega \rightarrow \{0, 1\}$ que asigna a cada individuo el valor 1 o el 0 según que posea o no una cierta enfermedad, respectivamente. Si suponemos conocido que existen 50 individuos enfermos en esa población es claro que la distribución P^X de X es una probabilidad en $\{0, 1\}$ que verifica $P^X(1) = P(X = 1) = 50/1000 = 0.05$ y $P^X(0) = P(X = 0) = 0.95$. De una v.a. como ésta que sólo toma dos valores posibles (0 y 1 que pueden representar cruz y cara, enfermo o sano, éxito o fracaso en un experimento, ...) y que toma el valor 1 con probabilidad un cierto número $p \in [0, 1]$ (y, entonces, el valor 0 con probabilidad $q = 1 - p$) se dice que tiene una distribución de Bernoulli de parámetro p . \triangleleft

Observación 2.7. Dado un espacio de probabilidad (Ω, P) como modelo para un experimento aleatorio, entendemos que Ω es el conjunto de resultados posibles del experimento y que P determina las probabilidades de que ocurra cualquier posible suceso de Ω . Al considerar una v.a. $X : (\Omega, P) \rightarrow \Omega'$ decidimos olvidarnos del resultado $\omega \in \Omega$ obtenido realmente en el experimento para conservar solamente el valor $X(\omega)$. De ese modo, el conjunto de resultados posibles pasa a ser el de los posibles valores de la v.a. X , es decir, Ω' . Así, considerar la v.a. X significa realmente utilizar el espacio de probabilidad (Ω', P^X) como modelo para el fenómeno considerado, en lugar del modelo original (Ω, P) . Los nuevos sucesos elementales serían ahora los elementos de Ω' . \triangleleft

2.1.2. Independencia de variables aleatorias

En el tema anterior hemos hablado de independencia de sucesos. A continuación hablaremos de independencia de v.a.

DEFINICIÓN 2.8. (Independencia de v.a.) Sean (Ω, P) un espacio de probabilidad y $X_i : (\Omega, P) \rightarrow \Omega'_i$, $1 \leq i \leq n$, v.a. Diremos que esas variables son independientes si cualesquiera que sean los sucesos $A'_i \subset \Omega'_i$, $1 \leq i \leq n$, se verifica que

$$P(\{X_1 \in A'_1\} \cap \dots \cap \{X_n \in A'_n\}) = P(X_1 \in A'_1) \dots P(X_n \in A'_n)$$

Observación 2.9. La hipótesis de independencia de v.a. juega un papel muy importante en el desarrollo del cálculo de probabilidades y de la inferencia estadística. Una situación típica en la que se asume la hipótesis de independencia es en la realización de pruebas sucesivas de un mismo experimento cuando a la hora de anotar el resultado de una de las pruebas no tenemos en cuenta los resultados de las pruebas precedentes (se habla en ese caso de realización de pruebas independientes). Concretando algo más, supongamos que Ω es un conjunto finito de personas (por ejemplo, de los alumnos de primer curso del Grado en Medicina) y que disponemos de un mecanismo para seleccionar alumnos al azar (es decir, garantizando que cada alumno tiene la misma probabilidad de ser elegido, y que cada par de alumnos, posiblemente repetidos, tiene la misma probabilidad de ser elegido que cualquier otro par); dos variables cualesquiera X e Y (por ejemplo, el grupo sanguíneo -A, B, AB o O- y el color de los ojos -claros u oscuros-) determinadas en un mismo individuo no serán, normalmente, independientes, por poco que parezca que tengan que ver una con la otra; pudieran serlo, pero es difícil que eso ocurra. Pero, elegidos dos individuos al azar (y con reemplazamiento), la variable X determinada para el primer individuo y la variable Y determinada para el segundo individuo, sean cuáles sean estas variables, son automáticamente independientes. Matemáticamente, dos v.a. $X : \omega \in \Omega \mapsto X(\omega) \in \Omega_1$ e $Y : \omega \in \Omega \mapsto Y(\omega) \in \Omega_2$ (en el ejemplo $\Omega_1 = \{A, B, AB, O\}$ y $\Omega_2 = \{\text{claro, oscuro}\}$) difícilmente serán independientes, pero sean cuáles sean X e Y , las v.a. $\hat{X} : (\omega, \omega') \in \Omega^2 \mapsto \hat{X}(\omega, \omega') = X(\omega) \in \Omega_1$ y $\hat{Y} : (\omega, \omega') \in \Omega^2 \mapsto \hat{Y}(\omega, \omega') = Y(\omega') \in \Omega_2$ son automáticamente independientes. \triangleleft

Veamos algunos ejemplos más:

Ejemplo 2.10. (3 lanzamientos independientes de una moneda) Supongamos ahora que disponemos de una moneda en la que la probabilidad de sacar cara (=1) es 0.05 y la de sacar cruz (=0) es 0.95. Si lanzamos al aire la moneda tres veces, los sucesos elementales son de la forma (i, j, k) , $i, j, k = 0, 1$, y el espacio muestral es $\Omega = \{0, 1\}^3$, es decir,

$$\Omega = \{(0, 0, 0), (0, 0, 1), (0, 1, 0), (1, 0, 0), (0, 1, 1), (1, 0, 1), (1, 1, 0), (1, 1, 1)\}$$

En este caso, puesto que la probabilidad de sacar cara no es la misma que la de sacar cruz, no podemos suponer que todos los sucesos elementales en Ω son equiprobables. Para calcular

las probabilidades de los sucesos elementales, denotemos por X, Y, Z los tres lanzamientos, es decir, son variables aleatorias definidas por $X(i, j, k) = i$, $Y(i, j, k) = j$ y $Z(i, j, k) = k$. Ya habíamos mencionado anteriormente que una situación típica en la que se tiene independencia es en la réplica de un mismo experimento un cierto número de veces como ocurre en el caso de los lanzamientos; de acuerdo con eso, parece razonable suponer que los lanzamientos X, Y y Z son independientes y, por ejemplo, la probabilidad P en Ω del suceso elemental $(0, 1, 0)$, es decir, la probabilidad de que en el primer lanzamiento se obtenga cruz, en el segundo cara y en el tercero otra vez cruz es el producto de las probabilidades correspondientes, es decir, $P(0, 1, 0) = 0.95 \cdot 0.05 \cdot 0.95$.

Análogamente, se tiene

$$\begin{aligned} P(0, 0, 0) &= 0.95 \cdot 0.95 \cdot 0.95 = 0.95^3 \\ P(0, 0, 1) &= 0.95 \cdot 0.95 \cdot 0.05 = 0.05 \cdot 0.95^2 \\ P(0, 1, 0) &= 0.95 \cdot 0.05 \cdot 0.95 = 0.05 \cdot 0.95^2 \\ P(1, 0, 0) &= 0.05 \cdot 0.95 \cdot 0.95 = 0.05 \cdot 0.95^2 \\ P(0, 1, 1) &= 0.95 \cdot 0.05 \cdot 0.05 = 0.05^2 \cdot 0.95 \\ P(1, 0, 1) &= 0.05 \cdot 0.95 \cdot 0.05 = 0.05^2 \cdot 0.95 \\ P(1, 1, 0) &= 0.05 \cdot 0.05 \cdot 0.95 = 0.05^2 \cdot 0.95 \\ P(1, 1, 1) &= 0.05 \cdot 0.05 \cdot 0.05 = 0.05^3 \end{aligned}$$

Se prueba sin dificultad que las distribuciones de probabilidad de X, Y y Z coinciden y $P^X(0) = P^Y(0) = P^Z(0) = 0.95$ y $P^X(1) = P^Y(1) = P^Z(1) = 0.05$.

Consideremos la v.a. S definida en Ω que a cada terna (i, j, k) le asigna la suma $i + j + k$; S cuenta, en realidad, el número de caras que se obtienen en los tres lanzamientos. Es claro que S toma valores en $\{0, 1, 2, 3\}$. Además

$$\begin{aligned} P^S(0) &= P(S = 0) = P(0, 0, 0) = 0.95^3 = \binom{3}{0} \cdot 0.05^0 \cdot 0.95^3 \\ P^S(1) &= P(\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}) = \binom{3}{1} \cdot 0.05^1 \cdot 0.95^2 \\ P^S(2) &= P(\{(1, 1, 0), (0, 1, 1), (1, 0, 1)\}) = \binom{3}{2} \cdot 0.05^2 \cdot 0.95^1 \\ P^S(3) &= P(1, 1, 1) = 0.05^3 = \binom{3}{3} \cdot 0.05^3 \cdot 0.95^0 \end{aligned}$$

En este ejemplo hemos considerado tres lanzamientos de una moneda para la que la probabilidad de sacar cara es 0.05, y hemos considerado la v.a. S que describe el número de caras obtenidas en esos lanzamientos. En probabilidad se habla frecuentemente de monedas, dados, . . . Ello es debido a que los orígenes del cálculo de probabilidades se encuentran en los juegos de azar (a raíz de ciertas cuestiones que el jugador profesional Caballero de Méré planteó al matemático Pascal); no obstante, estos ejemplos sobre monedas y dados constituyen además modelos fácilmente comprensibles para otros fenómenos aleatorios que sí tienen un interés práctico. Por ejemplo, el caso de la moneda sirve de modelo para el estudio de ciertas situaciones dicotómicas en las que sólo estamos interesados en si ha ocurrido o no

un cierto suceso, como pudiera ser que el individuo seleccionado al azar posea o no la enfermedad. Volvamos ahora al ejemplo 2 en el que estábamos interesados en una población de 1000 individuos de los que 50 padecen una cierta enfermedad. Supongamos que extraemos al azar 3 individuos de esa población; esta situación es totalmente equivalente a la que hemos descrito en este ejemplo, reemplazando el suceso “sacar cara” por el suceso “estar enfermo” (la probabilidad es en ambos casos 0.05) y la variable S describiría el número de enfermos entre los tres seleccionados. La distribución del número de enfermos es la que hemos descrito anteriormente. \triangleleft

Ejemplo 2.11. (n lanzamientos independientes de una moneda) Generalicemos algo más el ejemplo anterior. Supongamos ahora que disponemos de una moneda en la que la probabilidad de sacar cara es un cierto número $p \in [0, 1]$, y consideremos el experimento aleatorio que consiste en el lanzamiento de esa moneda un cierto número n de veces. El espacio muestral sería el producto cartesiano $\Omega = \{0, 1\}^n$. Denotemos por X_i el resultado del lanzamiento i -ésimo (es decir, X_i vale 0 o 1 según que en el i -ésimo lanzamiento se haya obtenido cruz o cara, respectivamente) Denotemos $q = 1 - p$ y tomemos $k_1, \dots, k_n \in \{0, 1\}$. Bajo la hipótesis de independencia de los lanzamientos, la probabilidad de que el resultado del experimento sea la n -upla (k_1, \dots, k_n) es

$$P(k_1, \dots, k_n) = p^{k_1} q^{1-k_1} \dots p^{k_n} q^{1-k_n} = p^{\sum_{i=1}^n k_i} q^{n - \sum_{i=1}^n k_i}$$

Supongamos ahora que sólo estamos interesados en el número de caras obtenidas en los n lanzamientos (despreciando el orden en el que aparecieron esas caras); estamos interesados solamente en la v.a. $S = \sum_{i=1}^n X_i$ definida en Ω y a valores en $\{0, 1, 2, \dots, n\}$. Argumentos análogos a los del ejemplo anterior (basados en la hipótesis de independencia de los lanzamientos) prueban ahora que la distribución de la v.a. S asigna al suceso elemental $k \in \{0, 1, \dots, n\}$ probabilidad

$$P^S(k) = P(S = k) = \binom{n}{k} p^k q^{n-k}$$

La distribución de la v.a. S es una de las más importantes del cálculo de probabilidades: se conoce como distribución binomial de parámetros n y p , y la denotaremos por $B(n, p)$. Es, como hemos visto, una probabilidad en el conjunto $\{0, 1, \dots, n\}$ y las probabilidades de los sucesos elementales se obtienen a partir de la fórmula precedente. Podemos decir, entonces, que la distribución de la suma S de n variables aleatorias independientes X_1, \dots, X_n cada una de ellas con distribución de Bernoulli de parámetro p es una distribución binomial de parámetros n y p . La distribución del ejemplo anterior es la distribución binomial $B(3, 0.05)$.

\triangleleft

2.2. Variables Aleatorias Discretas:

La Distribución Binomial

2.2.1. Variable aleatoria discreta. Función de probabilidad

Aunque también se llama discreta a una v.a. que toma una cantidad infinita numerable de valores numéricos, la definición siguiente será más que suficiente para nuestros objetivos.

DEFINICIÓN 2.12. (V.a. discreta) Una v.a. discreta es una v.a. que toma una cantidad finita de valores numéricos.

DEFINICIÓN 2.13. (Función de probabilidad) La aplicación $f : x_i \in \{x_1, \dots, x_n\} \mapsto f(x_i) = p_i$, que asocia a cada suceso elemental x_i su probabilidad p_i , se llama función de probabilidad de la v.a. discreta X .

Observación 2.14. Si $X : (\Omega, P) \longrightarrow \Omega'$ es una v.a. discreta entonces Ω' es un subconjunto finito de \mathbb{R} ; supongamos, por ejemplo, que $\Omega' = \{x_1, \dots, x_n\}$. La distribución de X es una probabilidad P^X en $\{x_1, \dots, x_n\}$ que quedará perfectamente determinada por las probabilidades de los sucesos elementales x_i , $1 \leq i \leq n$, es decir, por los números

$$p_i = P^X(x_i) = P(X = x_i)$$

p_1, \dots, p_n son números ≥ 0 que suman 1 pues $\sum_{i=1}^n p_i = \sum_{i=1}^n P(X = x_i) = P(X \in \{x_1, \dots, x_n\}) = P(\Omega) = 1$. Conocidos esos valores ya podemos calcular la probabilidad de que X tome valores en cualquier subconjunto $A \subset \Omega'$: basta sumar los p_i correspondientes a los sucesos elementales x_i que están en A , es decir,

$$P^X(A) = P(X \in A) = \sum_{i=1}^n p_i I_A(x_i)$$

donde $I_A(x_i) = 1$ si $x_i \in A$ y $I_A(x_i) = 0$ si $x_i \notin A$. Por tanto, la función de probabilidad de una v.a. discreta determina completamente la distribución de esa v.a. \triangleleft

2.2.2. Distribuciones de Bernoulli, binomial y uniforme discreta

Todos los ejemplos de v.a. considerados en el tema anterior son v.a. discretas, pues son las más sencillas. Entre las distribuciones de probabilidad discretas más importantes vamos a destacar tres, todas ellas conocidas ya de uno u otro modo. En los tres casos consideramos una v.a. $X : (\Omega, P) \rightarrow \Omega'$.

DEFINICIÓN 2.15. Distribución de Bernoulli) Diremos que la v.a. X tiene distribución de Bernoulli de parámetro $p \in [0, 1]$ si $\Omega' = \{0, 1\}$ y $P(X = 1) = p$ y $P(X = 0) = 1 - p$.

DEFINICIÓN 2.16. (Distribución binomial) Diremos que X tiene distribución binomial de parámetros $n \in \mathbb{N}$ y $p \in [0, 1]$ si $\Omega' = \{0, 1, \dots, n\}$ y

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n$$

DEFINICIÓN 2.17. (Distribución uniforme en un conjunto finito) Sea $n \in \mathbb{N}$. Diremos que X tiene distribución uniforme en el conjunto finito $\{1, \dots, n\}$ si $\Omega' = \{1, \dots, n\}$ y

$$P(X = k) = \frac{1}{n}, \quad k = 1, \dots, n$$

Observación 2.18. Notar que las probabilidades binomiales $\binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, 1, \dots, n$, son números positivos y que suman 1 pues, de acuerdo con la fórmula del binomio de Newton,⁽¹⁾ la suma de esas probabilidades coincide con $[p + (1-p)]^n = 1^n = 1$. Esta distribución se denotará por $B(n, p)$. La distribución binomial ya fue introducida como ejemplo en el tema anterior. En realidad, la distribución de Bernoulli de parámetro p es un caso particular de la binomial: concretamente es la distribución $B(1, p)$. La distribución binomial aparece siempre que se considera una v.a. que describe el número de individuos de entre n elegidos al azar que verifican una cierta característica dicotómica (como puede ser “estar o no enfermo”); se entiende que, una vez seleccionado un individuo y estudiada de él la característica que nos interesa, el individuo es devuelto a la población de tal suerte que podría volver a ser elegido posteriormente (se habla entonces de muestreo con reemplazamiento).⁽²⁾ Dicho más técnicamente, la distribución de la suma de n v.a. independientes con distribución de Bernoulli de parámetro p es la distribución binomial $B(n, p)$ (recordar el último ejemplo del tema anterior); de hecho, se puede probar que la distribución binomial es aditiva en el

¹ ¡De ahí el nombre de distribución binomial!

² En un tema posterior volveremos sobre esta cuestión, pues en la práctica suele utilizarse en poblaciones finitas el muestreo sin reemplazamiento que, en lugar de a la binomial, conduce a la llamada distribución hipergeométrica que no estudiaremos aquí.

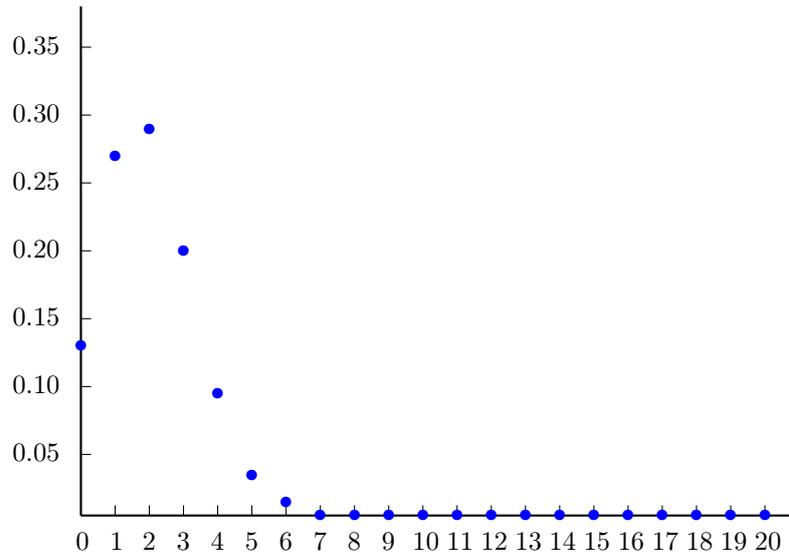


Figura 2.1: Distribución binomial $B(20, 0,1)$

sentido de que si X_1, \dots, X_m son v.a. independientes y si X_i tiene distribución binomial $B(n_i, p)$, $1 \leq i \leq m$, entonces la suma $\sum_{i=1}^m X_i$ tiene distribución binomial $B(\sum_{i=1}^m n_i, p)$. Su función de probabilidad en el caso $n = 20$ y $p = 0.1$ es la de la Figura 2.1:

La tabla II contiene las probabilidades para la distribución binomial $B(n, p)$ para valores de n menor o igual que 20 y ciertos valores de p inferiores a 0.5; no se incluyen valores de p mayores que 0.5 pues, si X es una v.a. con distribución $B(n, p)$ e Y es una v.a. con distribución $B(n, 1 - p)$, entonces, para cada $1 \leq k \leq n$, $P(X = k) = P(Y = n - k)$. \triangleleft

Observación 2.19. La distribución uniforme en un conjunto finito de puntos corresponde al caso clásico en que se considera un número finito n de sucesos elementales equiprobables (la probabilidad común no puede ser otra que $1/n$). Su función de probabilidad es la de la Figura 2.2 \triangleleft

2.2.3. Momentos de una distribución discreta

Asociados a la distribución de probabilidad de una v.a.r. (discreta o no) se consideran una serie de valores, llamados momentos de esa distribución, que nos proporcionan información parcial sobre la misma. Entre ellos destacaremos en principio dos: la media y la varianza, que, en el caso discreto, se definen a continuación.

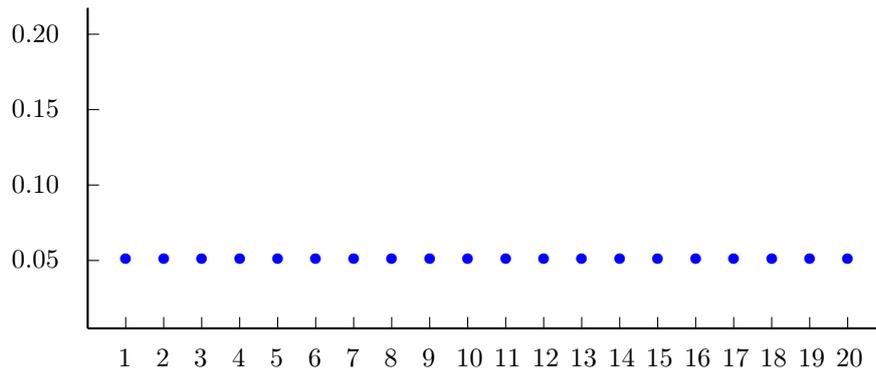


Figura 2.2: Distribución uniforme sobre $\{1, \dots, 20\}$

DEFINICIÓN 2.20. Consideremos para ello una v.a. discreta $X : (\Omega, P) \rightarrow \mathbb{R}$ que toma una cantidad finita de valores x_1, \dots, x_n con probabilidades respectivas $P(X = x_1) = p_1, \dots, P(X = x_n) = p_n$.

i) La media $\mu = E(X)$ de esa v.a. (o de la distribución de esa v.a.) se define por

$$\mu = \sum_{i=1}^n x_i p_i$$

ii) La varianza $\sigma^2 = \text{Var}(X)$ de esa v.a. (o de la distribución de esa v.a.) se define por

$$\sigma^2 = \sum_{i=1}^n (x_i - \mu)^2 p_i$$

La raíz cuadrada σ de la varianza σ^2 se llama desviación típica.

Observación 2.21. La media de la distribución de una v.a. es simplemente una media ponderada de los valores x_i que toma esa v.a., donde los “pesos” son las probabilidades con que la variable toma esos valores; se considera como una medida de posición de esa distribución. Para una v.a. que toma los valores -1 y 1 con probabilidad 1/2 cada uno (ver la Figura 2.3), podemos pensar en un cuerpo formado por dos bolas, cada una de ellas de masa 1/2, situadas en los puntos -1 y 1 de la recta real. En esta imagen la media de esa v.a. se corresponde exactamente con el centro de gravedad de ese cuerpo; si el valor 1 fuese tomado con probabilidad mayor que el -1 la media se desplazaría hacia la derecha de la misma manera que lo haría el centro de gravedad si la bola de la derecha fuese más pesada que la de la izquierda. Esta imagen gráfica justifica expresiones como la de pesos para referirnos a probabilidades. La varianza de la distribución de una v.a. es simplemente una media ponderada del cuadrado de las desviaciones a la media de los valores que toma la variable; se considera como una medida de dispersión de esa distribución. Para interpretar la desviación típica, podemos pensar que, aproximadamente, el 95 % de la distribución de una

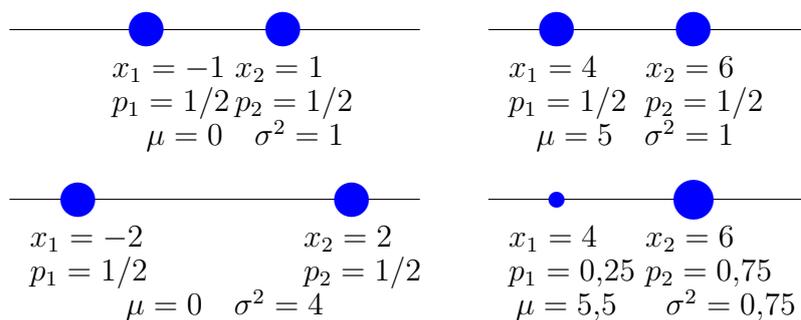


Figura 2.3: Imagen gráfica de la localización y dispersión de una distribución

v.a. se encuentra a un máximo de dos desviaciones típicas de la media; eso es exactamente así para la distribución normal que estudiaremos más adelante, y aproximadamente válido para otras muchas distribuciones (véanse la propiedad (P4) en la página 83 y la observación 2.42). En la Observación 0.37 habíamos establecido la analogía existente entre la cuasi-varianza de un conjunto de datos y el concepto físico de momento de inercia; esa analogía es igualmente válida para la varianza de una variable discreta. El gráfico precedente representa las distribuciones de cuatro v.a. discretas cada una de las cuales toma solamente dos valores posibles (-1 y 1 para la primera, -2 y 2 para la tercera y 4 y 6 para la segunda y la cuarta) cada uno de ellos con probabilidad 1/2 en los tres primeros casos y con probabilidades 0.25 y 0.75 en el cuarto caso. Una imagen física de esas distribuciones consiste en un cuerpo de masa 1 dividido en dos trozos (masas puntuales) colocados en los valores de la v.a. correspondiente con masas proporcionales a las probabilidades con que la v.a. toma esos valores; la media de la distribución nos da la posición del centro de gravedad del cuerpo. Las dos distribuciones de arriba tienen la misma varianza (sus masas están igualmente dispersas) pero diferente media (al sumar 5 unidades a cada uno de los valores de la primera variable para obtener la segunda, la media ha quedado aumentada en esas mismas 5 unidades). Las dos distribuciones de la izquierda en el gráfico poseen la misma media (ambas se encuentran distribuidas alrededor de la misma media $\mu = 0$) pero tienen distintas varianzas (la masa de la distribución de abajo está más dispersa que la de la de arriba). Las dos distribuciones de la derecha en el gráfico poseen distintas medias (nótese que, como indicábamos anteriormente, al desplazarse parte de la masa de una bola hacia la otra, el centro de gravedad del cuerpo se desplaza en esa misma dirección, siendo la interpretación probabilística análoga); nótese que, como cabría esperar, la dispersión de la de abajo es menor que la de arriba. \triangleleft

Observación 2.22. En el tema 1 habíamos definido la media y la varianza de un conjunto de datos numéricos x_1, \dots, x_n como $\bar{x} = \frac{1}{n} \sum_i x_i$ y $s^2 = \frac{1}{n-1} \sum_i (x_i - \bar{x})^2$. Nótese que estas no son más que la media y la varianza de una v.a.r. que toma los valores x_1, \dots, x_n con probabilidad $1/n$ cada uno (salvo que, por razones que aclararemos en las páginas 94 y 137, el denominador de la varianza es $n - 1$ y no n). Eso queda justificado por el hecho de que, aunque en estadística descriptiva los datos x_1, \dots, x_n nos interesan por sí mismos, en inferencia estadística acostumbra a ser valores de una cierta variable medida en

individuos que han sido seleccionados al azar de una población (constituyendo una muestra representativa de esta) y parece razonable que todos esos individuos sean ponderados del mismo modo (con el factor $1/n$) a la hora de calcular el valor medio de esa muestra. \triangleleft

Las siguientes son propiedades de la media y la varianza de una v.a. real que no probaremos:

Proposición 2.23. i) *La media de la suma de v.a. reales (v.a.r.) es la suma de las medias, es decir, si X_1, \dots, X_n son v.a.r., entonces $E(\sum_{i=1}^n X_i) = \sum_{i=1}^n E(X_i)$.*

ii) *La suma de las varianzas de v.a.r. independientes es la suma de las varianzas, es decir, si X_1, \dots, X_n son v.a. reales independientes, entonces $\text{Var}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \text{Var}(X_i)$.*

iii) *Si $a \in \mathbb{R}$, $E(a + X) = a + E(X)$ y $E(a \cdot X) = a \cdot E(X)$.*

iv) *Si $a \in \mathbb{R}$, $\text{Var}(a + X) = \text{Var}(X)$ y $\text{Var}(aX) = a^2 \text{Var}(X)$.*

v) *En general, $\text{Var}(X) = E[(X - E(X))^2] = E(X^2) - E(X)^2$.*

Ejemplo 2.24. Al experimento aleatorio que consiste en el lanzamiento al aire de un dado perfecto hemos asignado el espacio de probabilidad $(\{1, 2, 3, 4, 5, 6\}, P)$ donde $P(i) = 1/6$, $1 \leq i \leq 6$. Se nos propone el siguiente juego: lanzamos el dado y recibimos dos euros si sacamos un 6, un euro si sacamos un 4 o un 5 y pagamos dos euros si sacamos un 1, un 2 o un 3. La v.a. $X : \{1, 2, 3, 4, 5, 6\} \rightarrow \{-2, 1, 2\}$, definida por $X(1) = X(2) = X(3) = -2$, $X(4) = X(5) = 1$ y $X(6) = 2$, recoge las ganancias que obtenemos en un lanzamiento del dado. En este caso $\Omega' = \{-2, 1, 2\}$ y la distribución P^X de X es una probabilidad en Ω' que a los sucesos elementales $\{-2\}$, $\{1\}$ y $\{2\}$ asigna probabilidades:

$$P^X(\{-2\}) = P(X = -2) = P(\{1, 2, 3\}) = P(\{1\}) + P(\{2\}) + P(\{3\}) = \frac{1}{2}$$

$$P^X(\{1\}) = P(X = 1) = P(\{4, 5\}) = P(\{4\}) + P(\{5\}) = \frac{1}{3}$$

$$P^X(\{2\}) = P(X = 2) = P(\{6\}) = \frac{1}{6}$$

pues $X^{-1}(\{-2\}) = \{\omega \in \{1, 2, 3, 4, 5, 6\} : X(\omega) = -2\} = \{1, 2, 3\}$ y, análogamente, $X^{-1}(\{1\}) = \{4, 5\}$ y $X^{-1}(\{2\}) = \{6\}$. Esa distribución P^X nos indica cómo se distribuyen nuestras ganancias en un lanzamiento.

La media de X (ganancia media, o esperada, en un lanzamiento del dado) es $E(X) = (-2) \cdot \frac{1}{2} + 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{6} = -1 + \frac{2}{3} = -\frac{1}{3}$, que nos indica que el juego es desfavorable para nosotros, pues en media en cada partida perdemos un tercio de euro. En una partida concreta puede ocurrir cualquier cosa, pero, jugando partidas sucesivas de ese juego, a la larga terminaremos perdiendo un tercio de euro por partida jugada; este hecho queda justificado por la llamada ley fuerte de los grandes números, resultado debido a Kolmogorov que enunciaremos informalmente en una observación de la página 139. \triangleleft

A continuación recogemos la media y la varianza de las distribuciones discretas que hemos considerado anteriormente.

Proposición 2.25. i) La media y la varianza de la distribución de Bernoulli de parámetro p son

$$\mu = p \quad \text{y} \quad \sigma^2 = p(1 - p)$$

ii) La media y la varianza de la distribución binomial $B(n, p)$ son

$$\mu = np \quad \text{y} \quad \sigma^2 = np(1 - p)$$

iii) La media y la varianza de la distribución uniforme sobre un número finito n de puntos son

$$\mu = \frac{n + 1}{2} \quad \text{y} \quad \sigma^2 = \frac{(n + 1)(2n + 1)}{6}$$

Observación 2.26. Para la afirmación referente a la distribución de Bernoulli, nótese que $\mu = 0 \cdot (1 - p) + 1 \cdot p = p$ y $\sigma^2 = (0 - p)^2 \cdot (1 - p) + (1 - p)^2 \cdot p = [p + (1 - p)]p(1 - p) = p(1 - p)$. Para la demostración de las afirmaciones relativas a la distribución binomial, téngase en cuenta que $B(n, p)$ es la distribución de la suma de n v.a. independientes con distribución de Bernoulli de parámetro p y hágase uso de las propiedades precedentes de la media y la varianza. \triangleleft

Concluimos el tema con un ejemplo.

Ejemplo 2.27. Supongamos conocido que un traumatólogo tiene un 80% de éxito en sus intervenciones quirúrgicas para implantar prótesis de rodilla a pacientes con artrosis. Un individuo con problemas de artrosis en sus dos rodillas desea operarse de ambas. Consideremos la v.a. X_1 (respectivamente, X_2) que toma los valores 1 o 0 según que la intervención de la primera rodilla (respectivamente, la segunda rodilla) resulte exitosa o no. Entonces X_1 y X_2 tienen distribuciones de Bernoulli $B(1, 0.8)$. Si esas dos variables se suponen independientes (que podemos interpretar como que el hecho de que la primera intervención resulte o no exitosa no influye en modo alguno en el resultado de la segunda intervención), la v.a. $X = X_1 + X_2$, que describe el número de rodillas exitosamente operadas, tiene distribución binomial $B(2, 0.8)$. Su función de probabilidad es:

$$\begin{aligned} f(0) &= P(X = 0) = P(X_1 = 0, X_2 = 0) = P(X_1 = 0)P(X_2 = 0) \\ &= 0.2 \cdot 0.2 = 0.04 \end{aligned}$$

$$\begin{aligned} f(1) &= P(X = 1) = P(X_1 = 1, X_2 = 0) + P(X_1 = 0, X_2 = 1) \\ &= P(X_1 = 1)P(X_2 = 0) + P(X_1 = 0)P(X_2 = 1) \\ &= 0.8 \cdot 0.2 + 0.2 \cdot 0.8 = 0.32 \end{aligned}$$

$$\begin{aligned} f(2) &= P(X = 2) = P(X_1 = 1, X_2 = 1) = P(X_1 = 1)P(X_2 = 1) \\ &= 0.8 \cdot 0.8 = 0.64 \end{aligned}$$

en particular, la probabilidad de que ambas intervenciones resulten exitosas es $P(X = 2) = f(2) = 0.64$, y la probabilidad de que al menos una de las dos intervenciones resulte exitosa es $P(X \geq 1) = f(1) + f(2) = 0.96$. La media de X es $E(X) = np = 2 \cdot 0.8 = 1.6$ (haciendo uso de la definición de media ese valor es $E(X) = 0 \cdot 0.04 + 1 \cdot 0.32 + 2 \cdot 0.64 = 1.6$, que podemos interpretar diciendo que, en media, 1.6 de cada 2 intervenciones de artrosis de rodilla le resultan exitosas a ese traumatólogo –o, si se prefiere, una media de 16 rodillas de cada 20 resultarán exitosamente operadas– y la varianza es $\text{Var}(X) = npq = 2 \cdot 0.8 \cdot 0.2 = 0.32$ (como ejercicio, inténtese calcular la varianza de X haciendo uso de la definición de varianza). \triangleleft

2.3. Variables Aleatorias Continuas:

La Distribución Normal

2.3.1. Variable aleatoria continua:

Función de densidad. Momentos

Entre las v.a. que toman valores numéricos hemos distinguido dos tipos principales: aquellas que toman una cantidad finita de valores (que hemos llamado v.a. discretas) y las v.a. continuas que estudiamos en este tema. Las v.a. continuas toman una cantidad no numerable de valores; en principio, podemos pensar que pueden tomar cualquier valor de un cierto intervalo de la recta real. En este caso, la complejidad matemática es mayor que en el caso discreto. Suelen considerarse como continuas variables como pueden ser la estatura, el peso, la presión sanguínea sistólica, el nivel de colesterol en suero, etc, en grandes poblaciones. Una situación típica es que para una v.a. continua X , la probabilidad de que X tome un cierto valor x es nula (pensar, por ejemplo, en la probabilidad de que la estatura de un individuo elegido al azar de una población sea exactamente $x = 1.7500\dots$). Puede argumentarse, por ejemplo, que siendo finito el número de españoles, la v.a. que mide la estatura de un español elegido al azar es discreta. Debe tenerse en cuenta que esa variable, así considerada, es efectivamente discreta; pero debe quedar claro también que en la práctica uno de los principales usos de las v.a. continuas es la de aproximar a variables discretas que resultan poco manejables por tomar un número de valores que, aunque finito, es excesivamente grande.

Como hemos dicho anteriormente, la probabilidad de que una v.a. continua tome un valor concreto es nula (ésta es de hecho la definición de v.a. continua que podemos encontrar en libros avanzados de probabilidad). Resultan en este caso más interesantes el cálculo de probabilidades como las siguientes: dados dos números $a < b$, ¿cuál es la probabilidad de que

X tome valores entre a y b ?, es decir, ¿cómo podemos calcular $P^X([a, b]) = P(a \leq X \leq b)$? Para calcular esas probabilidades se usa normalmente el concepto de función de densidad que definimos a continuación.

DEFINICIÓN 2.28. (Función de densidad) Diremos que una v.a. continua $X : (\Omega, P) \rightarrow \mathbb{R}$ tiene función de densidad (o que su distribución P^X tiene función de densidad) si existe una aplicación no negativa $f : \mathbb{R} \rightarrow [0, +\infty[$ tal que, cualesquiera que sean $a < b$,

$$P(a \leq X \leq b) = \int_a^b f(x) dx$$

Observación 2.29. Lo dicho en la definición anterior es también válido en los casos $a = -\infty$ y $b = +\infty$, de tal suerte que $P(X \leq b) = P(-\infty \leq X \leq b) = \int_{-\infty}^b f(x) dx$ y $P(a \leq X) = P(a \leq X \leq +\infty) = \int_a^{+\infty} f(x) dx$. Además $\int_{-\infty}^{+\infty} f(x) dx = 1$, pues esa integral coincide con la probabilidad de que X tome valores en $\mathbb{R} =]-\infty, +\infty[$, lo cual ocurre siempre. Recíprocamente, cualquier función real no negativa con integral 1 es la función de densidad de una distribución de probabilidad en \mathbb{R} . \triangleleft

Observación 2.30. La función de densidad desempeña para v.a. continuas el mismo papel que la función de probabilidad en el caso discreto. La función de densidad determina completamente la distribución de esa v.a. continua. \triangleleft

Observación 2.31. La probabilidad de que $a \leq X \leq b$ no es, entonces, otra cosa que el área que queda bajo la curva $y = f(x)$ sobre el intervalo $[a, b]$. \triangleleft

Observación 2.32. (★) No toda v.a. continua posee función de densidad; las que sí la admiten son las llamadas v.a. absolutamente continuas, un subconjunto de las continuas. Pero todas las v.a. continuas que consideraremos en este libro son, de hecho, absolutamente continuas. \triangleleft

Como hemos dicho anteriormente, asociados a la distribución de probabilidad de una v.a.r. (discreta o no) se consideran una serie de números, llamados momentos de esa distribución, que nos proporcionan información parcial sobre la misma. Hemos destacado dos momentos principales: la media y la varianza; en el caso continuo se definen a continuación.

DEFINICIÓN 2.33. Sea $X : (\Omega, P) \rightarrow \mathbb{R}$ una v.a. continua con función de densidad $f : \mathbb{R} \rightarrow [0, +\infty[$. Entonces la media $\mu = E(X)$ y la varianza $\sigma^2 = \text{Var}(X)$ de esa v.a. X (o de la distribución de esa v.a.) se definen por

$$\mu = \int_{-\infty}^{+\infty} x \cdot f(x) dx$$

$$\sigma^2 = \int_{-\infty}^{+\infty} (x - \mu)^2 \cdot f(x) dx$$

La raíz cuadrada σ de la varianza σ^2 se llama desviación típica.

Observación 2.34. En el caso continuo, la media y la varianza reciben una interpretación análoga a la que ya habíamos dado en el caso discreto como medidas de posición y de dispersión, respectivamente, de la distribución de la v.a. X . De hecho, la definición de media en el caso continuo recuerda a la que dimos en el caso discreto, pues no es más que una suma continua (integral) de los valores x que toma la variable ponderados por el valor $f(x)$ de la función de densidad, que desempeña ahora el papel que la función de probabilidad jugaba en el caso discreto. Las propiedades de la media y la varianza que recogíamos en el tema anterior para el caso discreto son también ciertas en el caso continuo: la media de la suma es la suma de las medias, la varianza de la suma de v.a. independientes es la suma de las varianzas, etc... \triangleleft

2.3.2. La distribución normal. Propiedades

La mayor parte de los métodos estadísticos básicos que se usan en la práctica requieren solamente el conocimiento de cuatro distribuciones de probabilidad continuas: la normal, la t de Student, la chi-cuadrado de Pearson y la F de Fisher. Las tres últimas aparecen de forma natural en problemas de inferencia estadística sobre muestras de una distribución normal. De momento, presentaremos la distribución normal o de Gauss.

DEFINICIÓN 2.35. (Distribución normal) Sean $\mu \in \mathbb{R}$ y $\sigma^2 > 0$. Diremos que una v.a. $X : (\Omega, P) \rightarrow \mathbb{R}$ tiene distribución normal $N(\mu, \sigma^2)$ de parámetros μ y σ^2 si tiene como función de densidad la aplicación

$$x \in \mathbb{R} \longrightarrow f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

Observación 2.36. Se comprueba que esa función f así definida es no negativa y tiene integral 1 y, por tanto, es la densidad de una distribución de probabilidad que hemos llamado normal de parámetros μ y σ^2 . La representación gráfica de esa densidad tiene la forma de campana que podemos apreciar en la figura siguiente, razón por la cual se llama campana de Gauss a la representación gráfica de la densidad de la distribución normal $N(0, 1)$. \triangleleft

Algunas de las principales propiedades de la representación gráfica de la densidad de la distribución normal son las siguientes:

- (P1) (Media y varianza de la distribución normal) Se prueba que la media de una v.a. X con distribución $N(\mu, \sigma^2)$ es $E(X) = \mu$ y la varianza es $\text{Var}(X) = \sigma^2$. Se suele decir

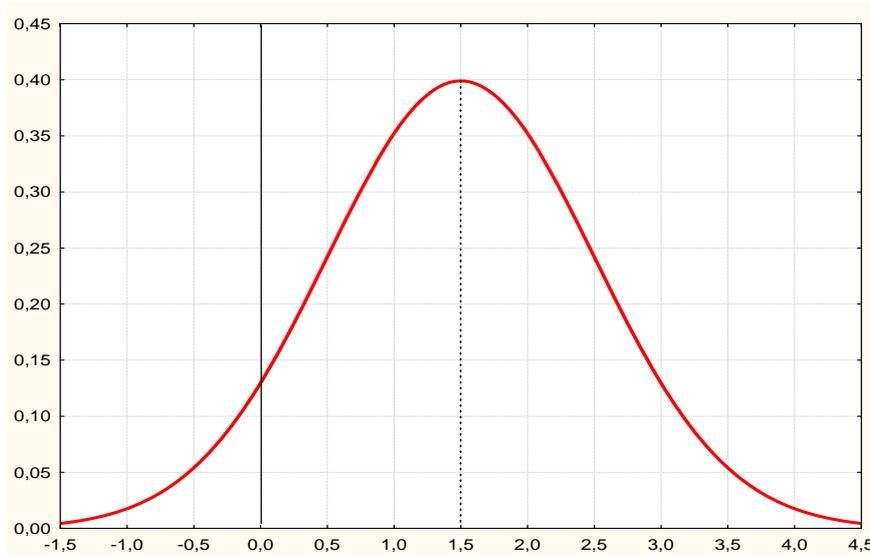


Figura 2.4: Función de densidad de la distribución $N(1.5, 1)$

por ello que $N(\mu, \sigma^2)$ es la distribución normal de media μ y varianza σ^2 . La posición relativa de la curva en el eje de abscisas lo determina el valor de la media μ (más a la derecha cuanto mayor es μ); por ello hemos interpretado μ como un parámetro de posición. Alcanza un máximo en el punto μ . El mayor o menor aplastamiento de la curva lo determina el valor de la desviación típica σ (la curva es más aplastada cuanto mayor es σ); por ello hemos interpretado σ^2 como un parámetro de dispersión. La curva tiene dos puntos de inflexión: las abscisas correspondientes son $\mu - \sigma$ y $\mu + \sigma$.

- (P2) (Mediana y simetría de la normal) Es simétrica respecto a la vertical trazada sobre la media μ . En particular, a la derecha y a la izquierda de esa vertical queda bajo la curva el mismo área, que es, por tanto, igual a $1/2$. Se dice, por ello, que μ es también mediana de la distribución $N(\mu, \sigma^2)$). Análogamente, si $x \in \mathbb{R}$, a la izquierda del punto $\mu - x$ queda bajo la curva el mismo área que a la derecha de $\mu + x$.
- (P3) (Tipificación) Una propiedad interesante de la distribución normal es que si la v.a. $X : (\Omega, P) \rightarrow \mathbb{R}$ tiene distribución $N(\mu, \sigma^2)$ y si $a, b \in \mathbb{R}$ entonces la v.a. $aX + b$ tiene distribución $N(a\mu + b, a^2\sigma^2)$. Se sigue de ello que si X tiene distribución $N(\mu, \sigma^2)$,

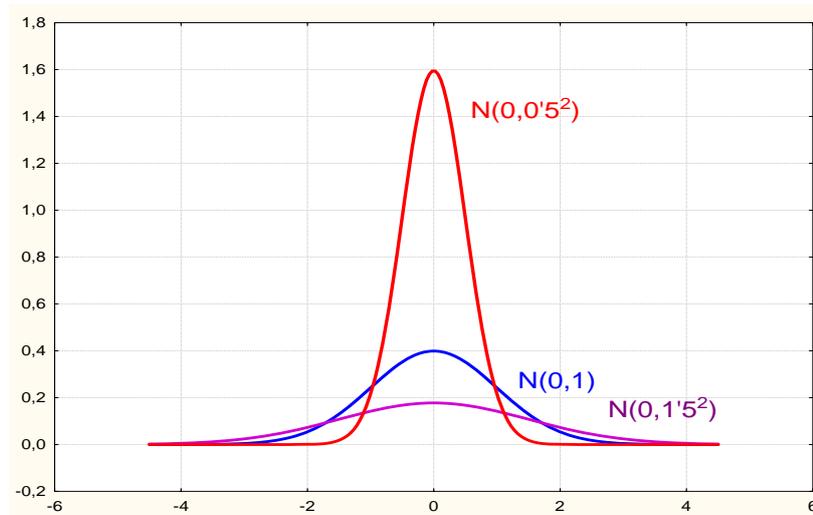


Figura 2.5: Distribuciones normales con media cero y desviaciones típicas 0.5, 1 y 1.5

entonces la v.a.

$$\frac{X - \mu}{\sigma}$$

tiene distribución $N(0, 1)$. El cálculo de probabilidades del tipo $P(X \leq a)$ para una v.a. X con distribución $N(0, 1)$ puede hacerse con la ayuda de la tabla III.2 en la que aparecen esas probabilidades para ciertos valores de a o con la ayuda de algún programa estadístico para ordenador; si X , en lugar de tener distribución $N(0, 1)$, tiene distribución $N(\mu, \sigma^2)$ entonces, de acuerdo con lo dicho, se tiene

$$P(X \leq a) = P\left(\frac{X - \mu}{\sigma} \leq \frac{a - \mu}{\sigma}\right)$$

y esta última probabilidad se puede calcular usando las tablas de la distribución $N(0, 1)$. El proceso descrito en el que hemos convertido una distribución $N(\mu, \sigma^2)$ en una $N(0, 1)$ sin más que restar μ y dividir por σ recibe el nombre de tipificación de la v.a. X .

- (P4) El eje de las equis es asíntota horizontal a la curva tanto en $-\infty$ como en $+\infty$. Así pues, una v.a. con distribución $N(\mu, \sigma^2)$ puede tomar teóricamente cualquier valor real, es decir, asigna probabilidad positiva a cualquier intervalo de la recta, aunque

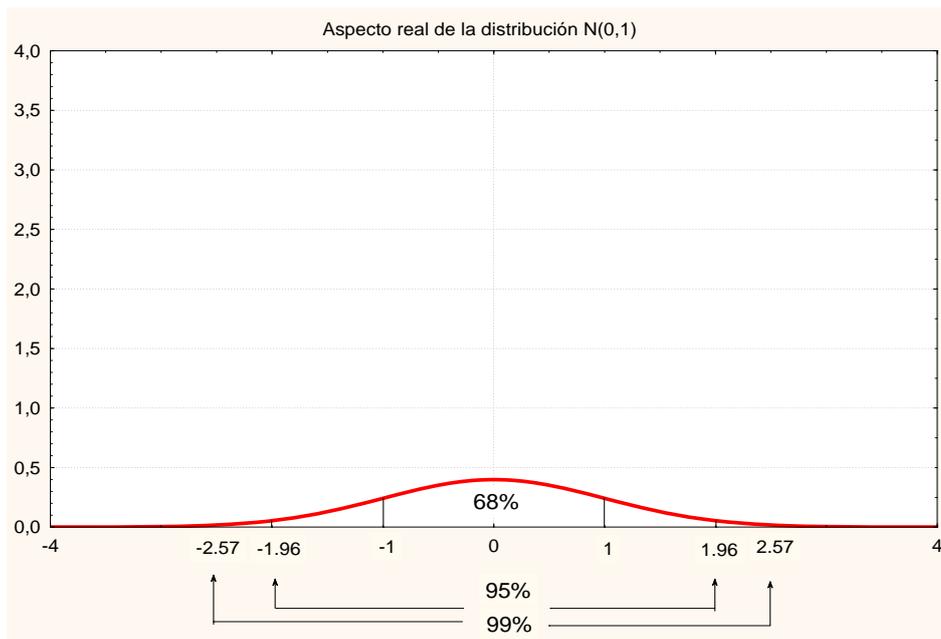


Figura 2.6: Aspecto real de la distribución $N(0, 1)$

esa probabilidad sea muy pequeña si el intervalo está lejos de la media. De hecho se verifica que si X tiene distribución $N(\mu, \sigma^2)$ entonces

$$\begin{aligned}
 P(\mu - 0.67\sigma \leq X \leq \mu + 0.67\sigma) &= 0.5 & P(\mu - \sigma \leq X \leq \mu + \sigma) &= 0.6827 \\
 P(\mu - 1.96\sigma \leq X \leq \mu + 1.96\sigma) &= 0.95 & P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) &= 0.9544 \\
 P(\mu - 2.57\sigma \leq X \leq \mu + 2.57\sigma) &= 0.99 & P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) &= 0.9973
 \end{aligned}$$

Observación 2.37. La expresión “aspecto real” de la distribución $N(0, 1)$ en la Figura 2.6 alude al hecho de que en los ejes de abscisas y ordenadas se ha utilizado la misma longitud para el intervalo unidad, aunque en la mayoría de los libros de estadística (y en este también) la función de densidad de la distribución normal estándar aparece mucho más pronunciada. He aquí un ejemplo más (véase la observación sobre gráficos engañosos en la página 34) de cómo cambios de escala en algunos de los ejes influyen decisivamente en la impresión que el gráfico nos proporciona. ◀

- (P5) (Cuantiles de la distribución normal) Para ciertos valores de x , la tabla III.2 nos permite calcular probabilidades del tipo $F(x) = P(X \leq x)$ cuando X tiene distribución $N(0, 1)$ y, mediante la tipificación, también podemos calcular esas probabilidades cuando X tiene una distribución normal cualquiera. Precisemos algo más esta afir-

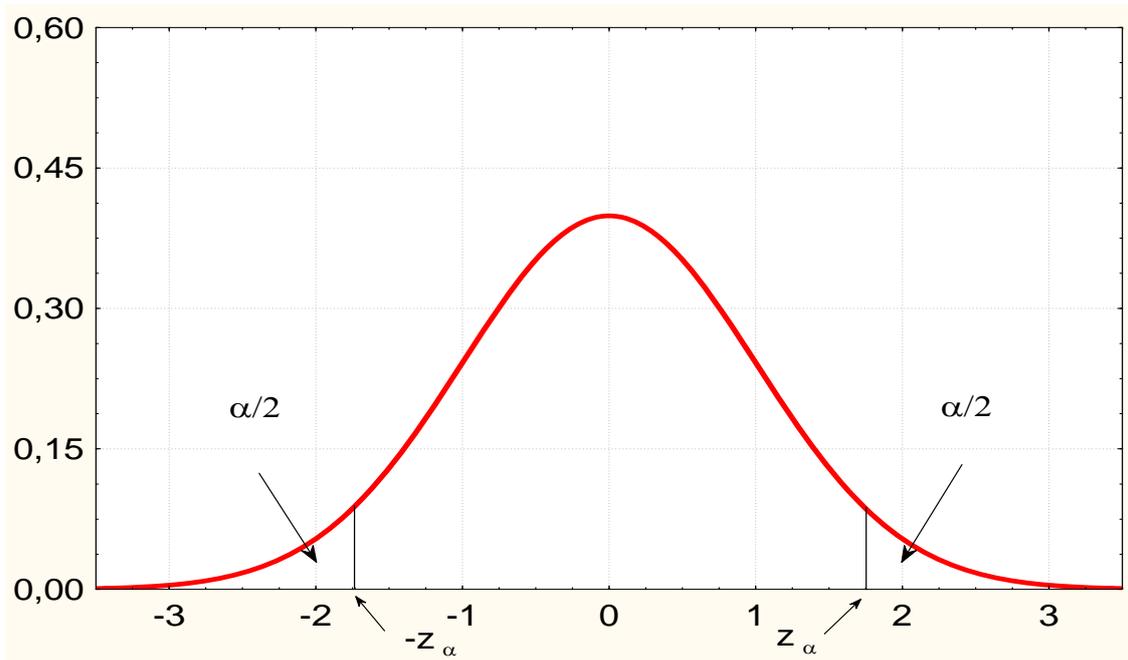


Figura 2.7: Cuantiles de la distribución $N(0, 1)$

mación. Dado un número $\alpha \in]0, 1[$ denotaremos por z_α el número real para el que el intervalo $[-z_\alpha, z_\alpha]$ contiene un $(1 - \alpha) \cdot 100\%$ de la distribución $N(0, 1)$.

Dicho de otro modo, z_α queda definido por $N(0, 1)([-z_\alpha, z_\alpha]) = 1 - \alpha$. Se define el cuantil de orden $\alpha \in]0, 1[$ (llamado también percentil $100 \cdot \alpha$) de una distribución continua como el número real que deja a su izquierda un $\alpha \cdot 100\%$ de la distribución. De acuerdo con esa definición, z_α es el cuantil de orden $1 - \alpha/2$ de la distribución $N(0, 1)$, pues fuera del intervalo $[-z_\alpha, z_\alpha]$ queda un área igual a α y, por simetría, sobre cada una de las colas $]-\infty, -z_\alpha]$ y $[z_\alpha, +\infty[$ quedará un área exactamente igual a $\alpha/2$. Para valores usuales de α , esos cuantiles aparecen tabulados en la tabla III.1.

Observación 2.38. Históricamente, la distribución normal apareció en el estudio de modelos probabilísticos para errores aleatorios: supongamos que deseamos determinar un cierto valor μ usando un instrumento de medida para el que errores positivos y negativos ocurren con la misma frecuencia y para el que valores próximos al valor μ ocurren con mayor frecuencia que valores alejados de μ . Los valores observados en diferentes realizaciones del experimento pueden interpretarse como valores de una cierta v.a. cuya distribución es la normal. \triangleleft

Observación 2.39. La distribución normal suele usarse como modelo probabilístico para el estudio de ciertas variables biológicas (por ejemplo, varios estudios epidemiológicos a gran escala sugieren el uso de la distribución normal en el estudio del nivel de colesterol en suero para grandes poblaciones); en otros casos puede ocurrir que una sencilla función, como el logaritmo, de una cierta variable biológica siga una distribución normal (por ejemplo, esos mismos estudios sugieren que el logaritmo de los niveles de triglicéridos en suero se distribuye normalmente). ◁

Observación 2.40. Pero la principal razón del amplio uso de la distribución normal en inferencia estadística es el llamado teorema del límite central que comentaremos más adelante (pág. 98). ◁

Observación 2.41. También en el caso continuo podemos proponer una imagen física que nos ayude a comprender el concepto de distribución de probabilidad continua. Con ese objetivo, nos apoyaremos en la densidad de la distribución normal. Pensemos en una barra metálica imaginaria de masa 1, longitud infinita y grosor nulo (en el caso discreto habíamos considerado masas puntuales). Suponemos que la masa está continua y desigualmente repartida a lo largo de la barra, de tal suerte que ésta es más densa en unos tramos que en otros. La función de densidad nos permite visualizar cómo está repartida la masa: concretamente, la masa de un cierto intervalo o tramo de la barra es igual al área que queda sobre ese intervalo bajo la gráfica de la función de densidad; es, por tanto, esta función la que nos permite averiguar cómo de densa –de ahí el nombre de función de densidad– es la barra en cada tramo que nos interese. Análogamente, la distribución de probabilidad de una v.a. continua se encuentra, en general, desigualmente repartida a lo largo de la recta real, de tal suerte que la probabilidad de que la variable tome valores en un cierto intervalo varía, en general, cuando el mismo se desplaza sobre la recta; por ejemplo, la probabilidad de que la estatura de una persona adulta de una cierta población se encuentre entre 160 cm y 190 cm, intervalo de longitud 30, es muy superior a la de que se encuentre entre 220 cm y 250 cm, intervalo éste también de longitud 30. También ahora, la media de esa distribución de probabilidad continua se corresponde con el centro de gravedad de la barra, y la varianza de la distribución se corresponde con el momento de inercia de la barra alrededor de su centro de gravedad. ◁

Observación 2.42. (★) Para una v.a.r. cualquiera X con media μ y desviación típica σ se verifica la siguiente desigualdad (llamada de Chebyshev): Si $k > 0$, $P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - k^{-2}$. En particular, para $k = 2$ (respectivamente, $k = 3$) se verifica que a una distancia máxima de dos desviaciones típicas (respectivamente, tres desviaciones típicas) de la media podemos encontrar al menos el 75 % (respectivamente, el 88.88 %) de la distribución de X . La afirmación que hacíamos en la observación 0.36 es un caso particular de esta. Como acabamos de ver, en el caso de la distribución normal podemos mejorar con mucho esa afirmación, pues esos porcentajes son, respectivamente, el 95.44 % y el 99.73 %.

◁

2.3.3. Aplicación al diagnóstico clínico: intervalos de normalidad

Describimos a continuación una aplicación de la distribución normal al diagnóstico clínico. En la práctica médica se usan con frecuencia intervalos que nos permiten clasificar a un individuo como normal o patológico según que una cierta característica numérica de ese individuo caiga dentro o fuera de ese intervalo. Son los llamados intervalos o límites de normalidad o de tolerancia. Esos intervalos pueden ser de tres tipos: 1) si se considera el nivel de glucosa en sangre, interesa buscar un límite superior de forma que a un individuo que lo supere se le declare diabético; 2) si se considera el número de espermatozoides por mm^3 interesa conocer un límite inferior por debajo del cual el individuo será declarado estéril; 3) si se trata de la concentración de potasio en plasma interesa conocer tanto un límite inferior como un límite superior, fuera de los cuales la vida del individuo corre peligro. Los individuos son considerados normales cuando la característica de interés cae dentro del intervalo de normalidad.

Asumiendo que la característica numérica que nos interesa sigue una distribución normal en la población, podemos determinar los límites de tolerancia como se explica en el ejemplo siguiente.

Ejemplo 2.43. Consideremos en primer lugar un ejemplo correspondiente al primer caso. Supongamos conocido que el nivel de colesterol en los individuos normales es una v.a. X con distribución normal $N(175, 13^2)$, y que estamos interesados en determinar un valor máximo para la v.a. X . Puesto que, bajo la distribución normal, es positiva la probabilidad de que la variable tome valores mayores que cualquier número fijado de antemano (aunque muy pequeña si el número es mucho mayor que la media), la pregunta así formulada no tiene sentido. Eso nos obliga a estar dispuestos a aceptar un pequeño error, y podemos estar interesados en determinar un valor M de forma que, por ejemplo, el 95 % de los individuos normales de la población tengan su nivel de colesterol inferior a M . De la tabla III.1 se deduce que $M = 175 + 1.645 \cdot 13 = 196.385$. Si a continuación adoptamos el criterio de declarar enfermo (hipercolesteronémico, en este caso) a todos los individuos que tengan nivel de colesterol superior a M , debemos tener en cuenta que un 5 % de los individuos normales son declarados enfermos erróneamente. Ese 5 % es el error del intervalo de normalidad. En general, una expresión para el límite superior de normalidad M es

$$M = \mu + \sigma z_{2\alpha}$$

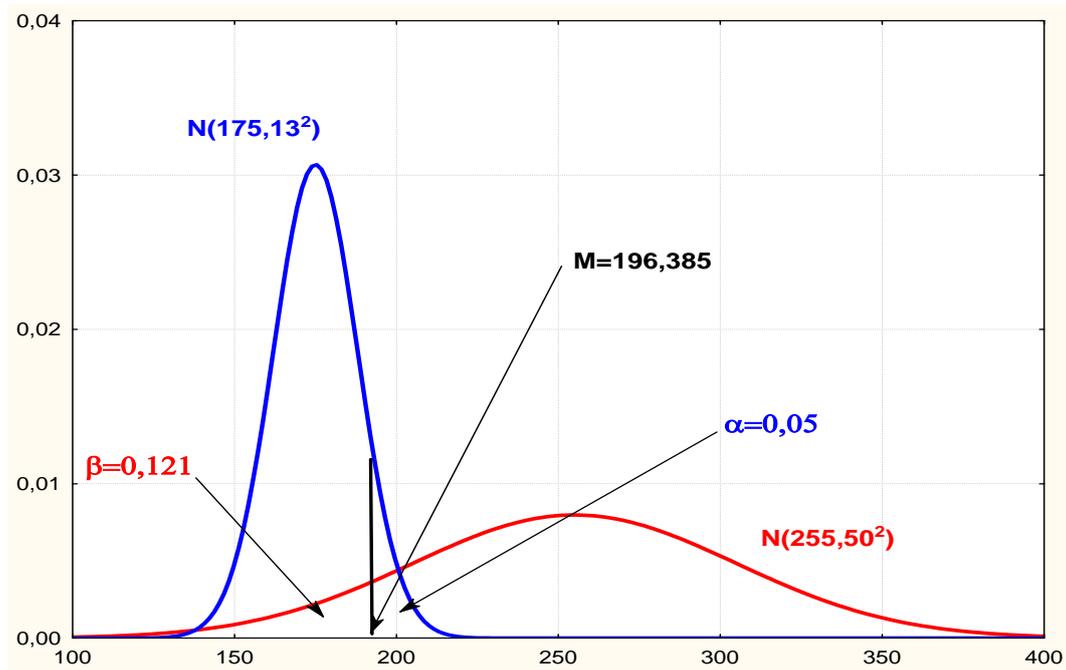


Figura 2.8: Criterio de diagnóstico y errores del mismo

si se desea un error $\alpha \in]0, 1[$ para el intervalo de normalidad, siendo μ y σ la media y la desviación típica de la característica de interés para los individuos normales. Análogamente, un límite inferior de tolerancia es

$$m = \mu - \sigma z_{2\alpha}$$

y un intervalo de tolerancia con límites inferior y superior y con error α es

$$[\mu - \sigma z_{\alpha}, \mu + \sigma z_{\alpha}]$$

Hasta ahora sólo hemos considerado uno de los dos errores posibles: el error α de declarar enfermo a un individuo normal. Obtenemos así un test de diagnóstico para esa enfermedad con un porcentaje del $\alpha \cdot 100\%$ de falsos positivos. Podemos considerar también el error de que un individuo enfermo sea declarado indebidamente normal; ese error se denota por β , y $\beta \cdot 100\%$ es el porcentaje de falsos negativos. En este ejemplo, supongamos que el nivel de colesterol de los individuos hipercolesteronémicos es una v.a. Y con distribución normal $N(255, 50^2)$. Entonces, el error β coincide con la probabilidad

$$P(Y \leq 196.385) = P\left(\frac{Y - 255}{50} \leq \frac{196.385 - 255}{50}\right) = P(Z \leq -1.17) = 0.121$$

de que un individuo hipercolesteronémico tenga nivel de colesterol inferior a 196.385 y sea, por tanto, declarado normal. El test de diagnóstico de la hipercolesteronemia así diseñado tiene un 5% de falsos positivos y un 12.1% de falsos negativos; por tanto, la especificidad es

del 95 % y la sensibilidad, del 87.9 %. Debe notarse que el 5 % de falsos positivos lo hemos fijado nosotros, mientras que el porcentaje de falsos negativos viene condicionado por ese hecho. <

Observación 2.44. Mejor que fijar el porcentaje de falsos positivos en un 5 % y obtener a partir de ahí el porcentaje de falsos negativos (12.1 % en el ejemplo), en la práctica, antes de fijar definitivamente el valor del límite superior de normalidad, se acostumbra a desplazar el valor de M sobre la recta a derecha e izquierda hasta encontrar un test con sensibilidad y especificidad que parezcan simultáneamente razonables, teniendo en cuenta que, como se observa en la figura, una ganancia en una de ellas conlleva una pérdida en la otra. <

Observación 2.45. La tabla siguiente recoge algunas variables cuantitativas de interés en análisis clínicos junto con intervalos de normalidad para las mismas facilitados por los Servicios de Bioquímica y Hematología de un hospital universitario español.

Bioquímica	Int. normal.	Hemograma	Int. normal.
Glucosa (mg/dl)	[70, 110]	Leucocitos (por l.)	$[4.4 \cdot 10^9, 11.3 \cdot 10^9]$
Urea (mg/dl)	[10, 40]	Hematíes (por l.)	$[4.1 \cdot 10^{12}, 5.1 \cdot 10^{12}]$
Acido úrico (mg/dl)	[3.5, 7]	Hemoglobina (g/dl)	[12, 15.3]
Creatinina (mg/dl)	[0.5, 1.1]	Hematocrito (%)	[36, 46]
Colesterol (mg/dl)	[150, 200]	Plaquetas (por l.)	$[150 \cdot 10^9, 410 \cdot 10^9]$
Triglicéridos (mg/dl)	[70, 170]	Neutrófilos (%)	[40.5, 71.8]
Proteínas totales (g/dl)	[6.5, 8]	Linfocitos (%)	[20, 45.5]
Calcio en suero (mg/dl)	[8.5, 10.5]	Monocitos (%)	[2.8, 9.2]
Fósforo en suero (mg/dl)	[3, 4.5]	Eosinófilos (%)	< 4
Colinesterasa (mg/dl)	[4600, 14400]	Basófilos (%)	< 1.5
Transaminasa GPT (UI/l)	[10, 40]	Fibrinógeno (mg/dl)	[150, 475]

Tabla 6.1. Intervalos de normalidad para ciertas variables de interés. <

2.4. Aproximación Probabilística al Concepto de Muestra

2.4.1. Muestreo aleatorio simple

En el tema de introducción a esta obra nos habíamos referido al concepto de muestra como un subconjunto de una cierta población Ω . Al proceso de extracción de una muestra de una población se le suele llamar muestreo. Más adelante comentaremos distintos tipos de muestreo. De momento pretendemos hacer una aproximación probabilística al concepto de muestra. Supongamos que estamos interesados en el estudio de una cierta característica de los individuos de una población Ω y que esa característica viene definida por la v.a. $X : \Omega \rightarrow \Omega'$. Siendo a veces muy costoso (o, incluso, imposible) estudiar todos los individuos de la población, cabe pensar en seleccionar adecuadamente un cierto número n de ellos, estudiar para ellos la característica que nos interesa y extrapolar las conclusiones extraídas a toda la población. Por ejemplo, si pretendemos obtener información sobre la estatura media de los españoles de 18 años, cabe pensar en seleccionar una muestra del conjunto de españoles que tienen esa edad y aproximar la estatura media buscada por la estatura media de los individuos de la muestra (es decir, la media muestral). El adverbio adecuadamente hace alusión al hecho de que, para que la extrapolación de conclusiones sea válida, en el proceso de selección de individuos no debe aparecer ningún tipo de preferencias por parte del seleccionador; en lugar de hablar de seleccionar adecuadamente hablaremos de seleccionar al azar, en un sentido que precisaremos posteriormente.

DEFINICIÓN 2.46. (Muestreo aleatorio simple) Para el caso de poblaciones finitas, el muestreo aleatorio simple queda caracterizado por el hecho de que cada muestra posible de n individuos tiene la misma probabilidad de ser extraída de la población.

Podemos distinguir dos tipos de muestreo aleatorio simple: el muestreo aleatorio simple con reemplazamiento (en el que cada individuo seleccionado es devuelto a la población tras haber sido estudiado y puede volver a ser seleccionado) y el muestreo aleatorio simple sin reemplazamiento (en el que un individuo que haya sido seleccionado en la muestra es retirado de la población y no podrá ser seleccionado en lo sucesivo). Consideraremos ahora solamente el muestreo con reemplazamiento.

Elegir con reemplazamiento n individuos de la población finita Ω equivale a elegir un elemento $\omega = (\omega_1, \dots, \omega_n)$ de Ω^n ; si a la hora de elegir el individuo i -ésimo de la población no se tiene en cuenta cuáles fueron las elecciones precedentes (es decir, si los n individuos se extraen independientemente los unos de los otros) y si la probabilidad de elegir un individuo cualquiera de la población no se modifica de una extracción a otra, la probabilidad P en Ω induce en el conjunto producto Ω^n una probabilidad P^n (que llamaremos probabilidad producto) definida por

$$P^n(\omega_1, \dots, \omega_n) = P(\omega_1) \cdots P(\omega_n)$$

Suponiendo además que todos los individuos de la población finita Ω tienen la misma probabilidad de ser elegidos (esa probabilidad será entonces igual a $1/N$ si N es el número de individuos de la población Ω), entonces

$$P^n(\omega_1, \dots, \omega_n) = P(\omega_1) \cdots P(\omega_n) = \frac{1}{N^n}$$

y todas las n -uplas $(\omega_1, \dots, \omega_n) \in \Omega^n$ tienen la misma probabilidad ($1/N^n$) de ser seleccionadas; por tanto, el espacio de probabilidad (Ω^n, P^n) corresponde al muestreo aleatorio simple con reemplazamiento de n individuos de la población Ω . Nótese que, como no podía ser de otro modo, el espacio de probabilidad correspondiente a muestras de tamaño n es diferente al espacio de probabilidad (Ω, P) correspondiente a la elección al azar de un único individuo de la población: en el caso del muestreo aleatorio simple con reemplazamiento, el nuevo espacio de probabilidad es el espacio producto (Ω^n, P^n) .

Estudiar las características $X(\omega_i)$ de los individuos ω_i , $1 \leq i \leq n$, equivale a considerar las v.a. X_i , $1 \leq i \leq n$, definidas en Ω^n por $X_i(\omega_1, \dots, \omega_n) = X(\omega_i)$. Se prueba que (considerando en Ω^n la probabilidad producto P^n) las v.a. X_i , $1 \leq i \leq n$, son independientes e

idénticamente distribuidas, y que cada X_i tiene la misma distribución que X (se dice por ello que las X_i son copias independientes de la v.a. X ; en el tema 4 habíamos visto cómo los tres lanzamientos X , Y y Z de una moneda, que como aplicaciones son diferentes, tienen la misma distribución). Así pues, las v.a. X_1, \dots, X_n son independientes e idénticamente distribuidas.

Observación 2.47. (★) La definición de la probabilidad producto P^n dada anteriormente tiene perfecto sentido incluso en el caso de que no todos los sucesos elementales sean equiprobables y aunque Ω sea finito o infinito numerable. La probabilidad producto se puede definir también en el caso continuo (en que Ω es \mathbb{R} y la probabilidad P viene definida por un densidad) pero eso queda lejos del alcance de esta obra; mencionaremos, sin embargo, que incluso en ese caso el espacio de probabilidad producto (Ω^n, P^n) es el que corresponde al muestreo aleatorio simple con reemplazamiento en una población infinita no numerable como es Ω en el que los individuos son elegidos independientemente unos de otros. ◁

2.4.2. Concepto de muestra

La introducción precedente nos conduce a la siguiente definición probabilística de muestra.

DEFINICIÓN 2.48. (Muestra de tamaño n) Llamaremos muestra de tamaño n a una colección de n v.a. X_1, \dots, X_n independientes e idénticamente distribuidas definidas en un mismo espacio de probabilidad (Ω, P) y a valores en un conjunto Ω' . Puesto que todas las v.a. X_i tienen la misma distribución, diremos también que se trata de una muestra de tamaño n de esa distribución.

Observación 2.49. Tras un muestreo aleatorio simple con reemplazamiento en una población Ω , se obtiene una n -upla $(\omega_1, \dots, \omega_n) \in \Omega^n$, y si sólo estamos interesados en una característica numérica descrita por la v.a. $X : \Omega \rightarrow \mathbb{R}$ de los individuos de la población, la muestra es, en realidad, una n -upla $(x_1, \dots, x_n) \in \mathbb{R}^n$ de números reales, donde $x_i = X(\omega_i) = X_i(\omega_1, \dots, \omega_n)$; esa n -upla (x_1, \dots, x_n) se considera como una realización concreta de la muestra X_1, \dots, X_n . Tras otra extracción de n individuos de la misma población Ω , se obtendrá una n -upla (x'_1, \dots, x'_n) en principio diferente de la anterior, pero que se considera asimismo como otra realización concreta de la misma muestra X_1, \dots, X_n . ◁

Observación 2.50. (Tablas de números aleatorios) La obtención de una muestra requiere, en primer lugar, la identificación de toda la población a estudiar; una vez enumerados todos

los individuos de la población, mediante un procedimiento análogo a un sorteo se pueden extraer al azar una serie de números, y los individuos correspondientes constituyen una muestra aleatoria de la población. La realización de un sorteo de ese tipo puede resultar tedioso; con el objeto de simplificar ese proceso se han construido las llamadas tablas de números aleatorios. Una tabla de números aleatorios es una lista de dígitos del 0 al 9 que se suponen dispuestos al azar de forma que todos los dígitos del 0 al 9 tienen la misma probabilidad ($1/10$) de aparecer, todos los pares de dígitos desde el 00 hasta el 99 tienen la misma probabilidad ($1/100$) de aparecer, todas las ternas de dígitos desde el 000 hasta el 999 tienen la misma probabilidad ($1/1000$) de aparecer, . . . Puesto que la tabla es necesariamente finita (la mayor que se conoce tiene un millón de dígitos aleatorios y fue construida por la Rand Corporation), nos conformaremos con que eso sea aproximadamente así para grupos de n dígitos con n pequeño. Esas tablas se manejan del siguiente modo: ante una población finita con N individuos, lo primero que debemos hacer es asignar a cada individuo de la población un número entre 0 y $N - 1$; si queremos extraer una muestra de tamaño n de esa población y n es una cifra con 3 dígitos, se eligen de la tabla n bloques consecutivos de tres cifras teniendo en cuenta que 000 es 0, que 087 es 87, etc; si sólo queremos números al azar entre el 99 y el 450, rechazaremos los bloques de tres cifras que no estén entre ambos y seguiremos tomando bloques de 3 dígitos hasta completar los n deseados; si se desea un muestreo sin reemplazamiento, se desechan los números repetidos. Si se desean números al azar del intervalo $[0, 1]$ con una precisión de, por ejemplo, cuatro cifras decimales, se eligen de la tabla bloques consecutivos de cuatro dígitos que serán los decimales de los números buscados, siendo 0 la parte entera. Actualmente, existen programas estadísticos que generan números aleatorios (o, más precisamente, números pseudoaleatorios). La tabla I contiene 3000 dígitos pseudoaleatorios generados con la ayuda del programa estadístico STATISTICA. \triangleleft

Observación 2.51. Desgraciadamente, en las ciencias de la salud el muestreo aleatorio simple, aunque es un objetivo ideal a cumplir, rara vez se verifica, pues el investigador se ve condicionado a elegir sus datos entre los individuos que le llegan al hospital, o entre los animales de experimentación que le proporcionan, etc. Conviene pues distinguir entre la población objetivo (sobre la que queremos información) y la población de muestreo (aquella subpoblación de la población objetivo de la que realmente se extrajo la muestra) y estudiar la muestra obtenida para ver si todas aquellas características relevantes para el objeto de nuestro estudio son conformes a lo que ocurre en la población objetivo; por ejemplo, conviene comprobar si en la muestra las distribuciones por sexo, por edades y por otras variables que consideremos relevantes para nuestro estudio son análogas a las de la población objetivo. \triangleleft

Observación 2.52. Se habla, a veces, de muestreo probabilístico cuando se seleccionan al azar y con reemplazamiento individuos de una población Ω que no tiene por qué ser finita y que, aunque lo fuese, no considere a todos sus individuos equiprobables; sólo es imprescindible la hipótesis de independencia en la selección de los individuos –lo que queda garantizado si a la hora de elegir al individuo i -ésimo no tenemos en cuenta lo ocurrido en elecciones precedentes– y la hipótesis de que la probabilidad de elegir a un individuo se mantenga inalterable de una elección a otra. \triangleleft

Ejemplo 2.53. No sería un muestreo aleatorio simple un caso como el siguiente: para seleccionar 10 individuos de una clase de cien alumnos, se colocan estos por orden alfabético, se enumeran del uno al cien, se elige un número del uno al cien al azar y se eligen el alumno correspondiente y los nueve siguientes. \triangleleft

2.4.3. Momentos muestrales: sus distribuciones

Anteriormente hemos considerado la media $E(X)$ y la varianza $\text{Var}(X)$ de una v.a. $X : \Omega \rightarrow \mathbb{R}$ como los principales momentos de esa v.a. (o de su distribución P^X). En este tema consideramos una muestra X_1, \dots, X_n de tamaño n de una distribución de probabilidad en \mathbb{R} y definiremos a partir de ella dos nuevas v.a. que llamaremos media muestral y varianza muestral.

DEFINICIÓN 2.54. Sea $X_1, \dots, X_n : (\Omega, P) \rightarrow \mathbb{R}$ una muestra de tamaño n .

1) (Media muestral) Se define la media muestral de esa muestra como la v.a. \bar{X} definida por

$$\bar{X} = \frac{X_1 + \dots + X_n}{n}$$

2) (Varianza muestral) Se define la varianza muestral para esa muestra como la v.a. S^2 definida por

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n-1} \left(\sum_{i=1}^n X_i^2 - n\bar{X}^2 \right)$$

Observación 2.55. Dividir la suma de cuadrados por $n-1$ y no por n a la hora de definir la varianza muestral S^2 queda justificado por el apartado (iii) del siguiente teorema, pues, en la terminología que introduciremos en el tema de estimación en la Segunda Parte, eso significa que S^2 es un estimador insesgado de la varianza σ^2 . Otros autores prefieren reservar el nombre de varianza muestral para $\frac{n-1}{n}S^2$. \triangleleft

De esas dos nuevas v.a. (o de sus distribuciones) definidas a partir de las n v.a. que constituyen la muestra nos interesan algunos resultados como el siguiente.

TEOREMA 2.56. Con las notaciones de la definición anterior, se verifican las proposiciones siguientes:

(i) Puesto que las X_i tienen la misma distribución, tienen también la misma media μ . La media de la media muestral es también igual a μ , es decir,

$$E(\bar{X}) = \mu$$

(ii) Puesto que las X_i tienen la misma distribución, tienen también la misma varianza σ^2 . La varianza de la media muestral \bar{X} es

$$\text{Var}(\bar{X}) = \frac{\sigma^2}{n}$$

(iii) La varianza muestral S^2 tiene media σ^2 , es decir,

$$E(S^2) = \sigma^2$$

Para muestras de una distribución normal, el resultado precedente puede ser mejorado.

TEOREMA 2.57. Sea X_1, \dots, X_n una muestra de una distribución normal $N(\mu, \sigma^2)$ (es decir, las X_i son v.a. independientes e idénticamente distribuidas con distribución común $N(\mu, \sigma^2)$). Entonces:

(i) \bar{X} tiene distribución $N(\mu, \sigma^2/n)$.

(ii) (Distribución χ^2 de Pearson) La distribución de la v.a.

$$\frac{(n-1)S^2}{\sigma^2}$$

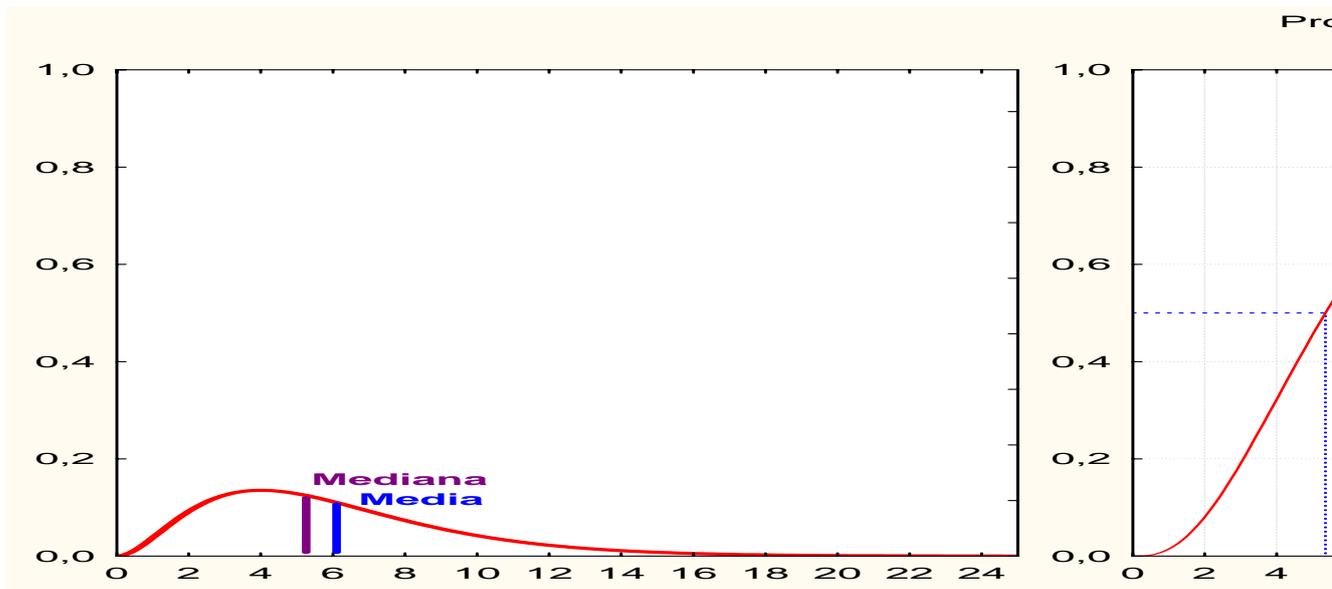
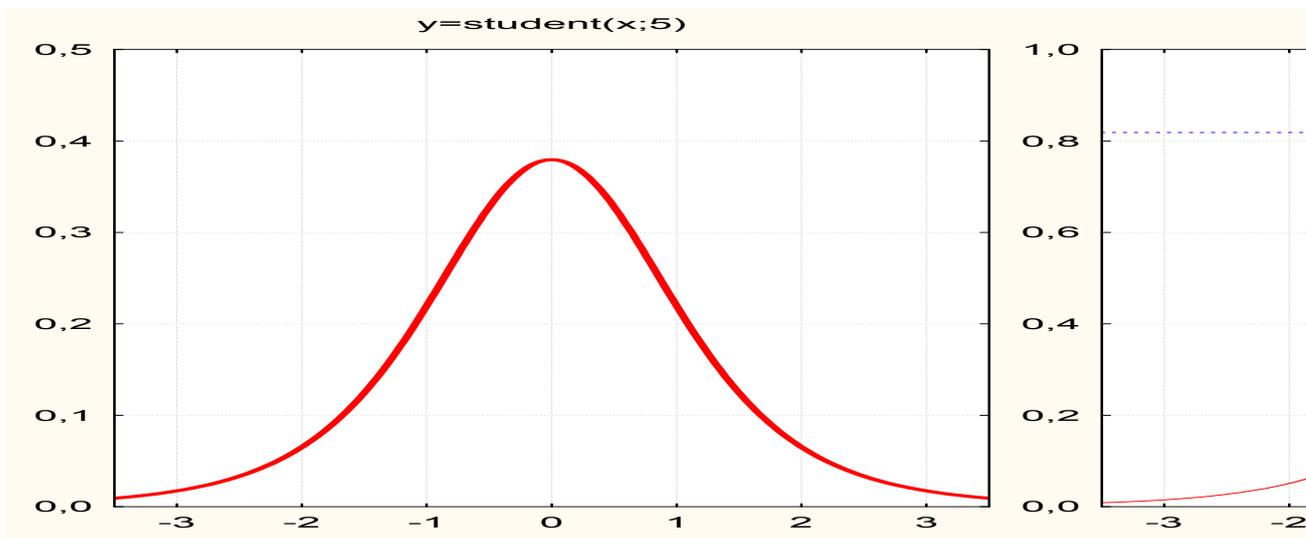
es la llamada distribución χ^2 de Pearson con $n-1$ grados de libertad, que denotaremos por $\chi^2(n-1)$.

(iii) (Distribución t de Student) La distribución de la v.a.

$$\sqrt{n} \frac{\bar{X} - \mu}{S}$$

es la llamada distribución t de Student con $n-1$ grados de libertad, que denotaremos por $t(n-1)$.

(iv) (Distribución F de Fisher) Supongamos ahora que tenemos dos muestras de tamaños n_1 y n_2 independientes de dos distribuciones normales $N(\mu_1, \sigma_1^2)$ y $N(\mu_2, \sigma_2^2)$, respectivamente, y denotemos por S_1^2 y S_2^2 las varianzas muestrales correspondientes a la primera y a

Figura 2.9: Densidad de la distribución $\chi^2(6)$ Figura 2.10: Densidad de la distribución $t(5)$

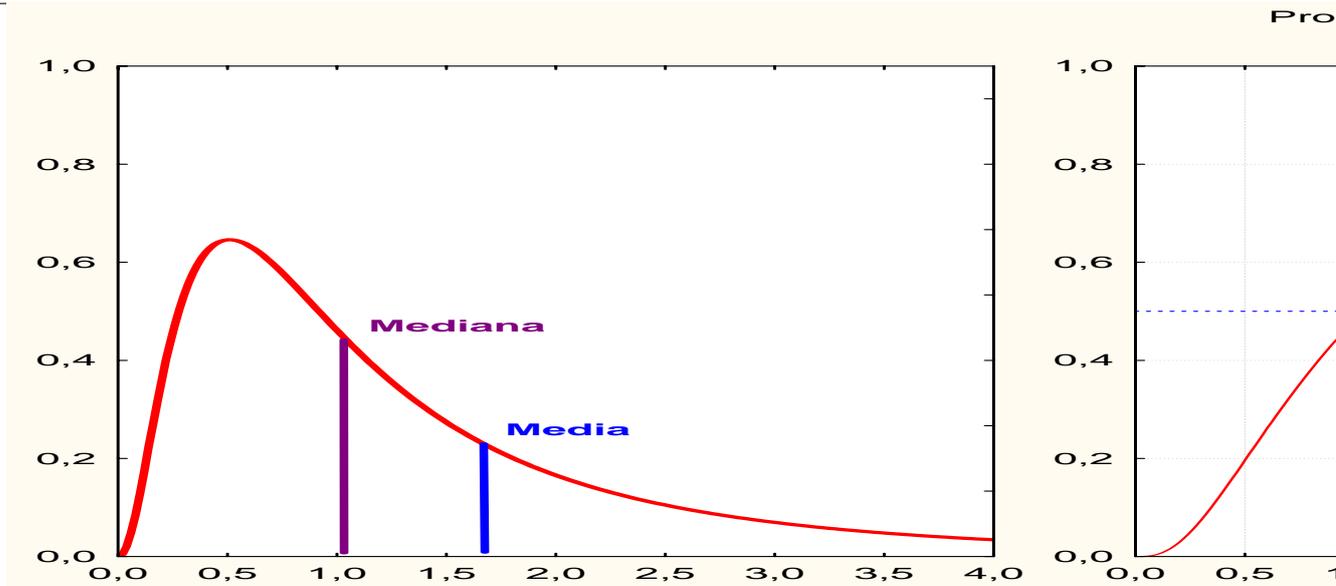


Figura 2.11: Densidad de la distribución $F(7, 5)$

la segunda muestras, respectivamente. La distribución de la v.a.

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$$

es la llamada distribución F de Fisher con $(n_1 - 1, n_2 - 1)$ grados de libertad, que denotaremos por $F(n_1 - 1, n_2 - 1)$.

Observación 2.58. Nótese que la densidad de la distribución $t(n)$, aunque no es una campana de Gauss, recuerda la densidad de la distribución $N(0, 1)$; de hecho, a medida que aumenta el número n de grados de libertad esa densidad se aproxima más y más a la campana de Gauss. La tabla IV contiene ciertos cuantiles de la distribución t para valores de n inferiores a 100; para tamaños muestrales superiores a 100 podemos utilizar los cuantiles z_α de la distribución $N(0, 1)$ en lugar de $t_\alpha(n)$. Para valores de n no tabulados, pero comprendidos entre dos valores que sí han sido tabulados, en ésta y en otras tablas podemos aproximar los cuantiles por interpolación lineal, como se explica a continuación con un ejemplo: el método de interpolación lineal actúa como si $t_{0,05}(n)$ fuese una función lineal a trozos de n ; puesto que en la tabla IV de la distribución t no aparece el valor $t_{0,05}(42)$, pero sí los valores $t_{0,05}(40) = 2.021$ y $t_{0,05}(45) = 2.014$, aproximaremos $t_{0,05}(42)$ por el valor en el punto 42 de la recta que une los puntos $(40, 2.021)$ y $(45, 2.014)$, cuya ecuación es $y - 2.021 = \frac{2.014 - 2.021}{45 - 40}(x - 40)$; la aproximación deseada es, por tanto, $y = 2.021 + \frac{-0.007}{5}(42 - 40) = 2.0182$; el valor real, calculado con la ayuda de un programa estadístico, es 2.018082. ◀

Observación 2.59. Las densidades de las distribuciones χ^2 y F son densidades de v.a. positivas y, por tanto, dan probabilidad nula al intervalo $]-\infty, 0[$. ◀

Observación 2.60. La distribución F también se suele conocer como distribución de Snedecor o, incluso, de Fisher-Snedecor. \triangleleft

2.4.4. Teorema del límite central

La importancia de la distribución normal en estadística queda plenamente justificada por el siguiente resultado.

TEOREMA 2.61. (Teorema del límite central) *Sea X_1, X_2, \dots una muestra de una distribución real con media μ y varianza σ^2 . Si n es suficientemente grande, la media muestral \bar{X} de X_1, \dots, X_n verifica que*

$$\sqrt{n} \frac{\bar{X} - \mu}{\sigma}$$

tiene aproximadamente distribución normal $N(0, 1)$; o bien, para n grande, \bar{X} tiene aproximadamente distribución $N(\mu, \sigma^2/n)$; o bien, para n grande, $\sum_{i=1}^n X_i$ tiene aproximadamente distribución $N(n\mu, n\sigma^2)$.

Observación 2.62. El teorema del límite central es uno de los principales del cálculo de probabilidades y convierte a la distribución normal en la distribución más importante de la estadística pues, además del uso que anteriormente le hemos atribuido en teoría de errores y como modelo avalado por estudios epidemiológicos para ciertas variables biológicas, nos permite utilizar la distribución normal para aproximar la distribución de otras variables de las que sabemos positivamente que no son normales (porque sean, por ejemplo, discretas) cuando el tamaño de la muestra es grande. Tengamos en cuenta que de la distribución común de las X_i sólo suponemos que tiene media y varianza finita, y por lo demás puede ser cualquier distribución (podría ser discreta, como la binomial, la uniforme, etc..., o cualquier distribución continua), y nos asegura que la media muestral obtenida a partir de una muestra de tamaño suficientemente grande de esa distribución tiene distribución aproximadamente normal. \triangleleft

Observación 2.63. Buena parte de los métodos estadísticos que presentaremos a lo largo de esta obra se basan en la suposición de normalidad de las observaciones. Lo dicho en el apartado anterior nos permite utilizar con éxito esos mismos métodos, incluso en el caso de que la hipótesis de normalidad no esté garantizada, si disponemos de tamaños de muestra suficientemente grandes. La distribución normal se convierte así en una distribución especialmente sencilla (sólo depende de dos parámetros: la media μ y la varianza σ^2), y los métodos estadísticos diseñados para ella pueden usarse para otras distribuciones en el caso de grandes muestras. \triangleleft

Observación 2.64. El teorema del límite central viene también a justificar el uso de la distribución normal como modelo apropiado para muchas variables biológicas de interés o en teoría de errores del siguiente modo: si el resultado numérico de un experimento aleatorio depende de una gran cantidad de factores (digamos, es la media de una gran cantidad de factores) que actúan independientemente unos de otros (eso justificaría la hipótesis de independencia) y contribuyendo cada uno de ellos en la misma forma al resultado final (eso justificaría la hipótesis de que las X_i tienen la misma distribución), entonces el valor obtenido puede considerarse aproximadamente como una observación concreta de una v.a.r. que sigue una distribución normal. \triangleleft

2.4.5. Aproximación de la distribución binomial por la normal

Veamos ahora cómo podemos aproximar la distribución binomial por la normal haciendo uso del teorema del límite central.

La necesidad de esas aproximaciones viene justificada por el hecho de que trabajar con la distribución binomial $B(n, p)$ cuando n es grande resulta excesivamente engorroso. Puesto que la distribución binomial $B(n, p)$ se puede describir como la suma de v.a. independientes de Bernoulli de parámetro p , el teorema del límite central prueba que, si n es grande, es posible aproximar la distribución binomial por una distribución normal de la misma media y la misma varianza que la binomial. Concretamente, si X es una v.a. con distribución binomial $B(n, p)$ e Y es una v.a. con distribución $N(np, np(1-p))$, entonces la distribución de X se aproxima por la de Y del siguiente modo: Si $k = 1, \dots, n-1$, la probabilidad $P(X = k)$ no se aproxima por $P(Y = k)$, que es nula, sino por $P(k - 1/2 \leq Y \leq k + 1/2)$; $P(X = 0)$ se aproxima por $P(Y \leq 1/2)$ y $P(X = n)$ se aproxima por $P(Y \geq n - 1/2)$; dados dos números reales $a \leq b$, la probabilidad $P(a \leq X \leq b)$ puede aproximarse por el área bajo la densidad de la distribución $N(np, np(1-p))$ entre los puntos $a - 1/2$ y $b + 1/2$, es decir, por $P(a - 1/2 \leq Y \leq b + 1/2)$. En el gráfico siguiente se observa la aproximación de la distribución binomial $B(25, 0.4)$ (cuya función de probabilidad está representada por pequeños círculos negros) por la distribución normal $N(10, 6)$. El rectángulo del gráfico sobre el punto $k \in \{0, 1, \dots, 25\}$, de base 1, tiene área igual a su altura, que coincide con la probabilidad de que $X = k$; esa probabilidad se aproxima por el área que queda bajo la

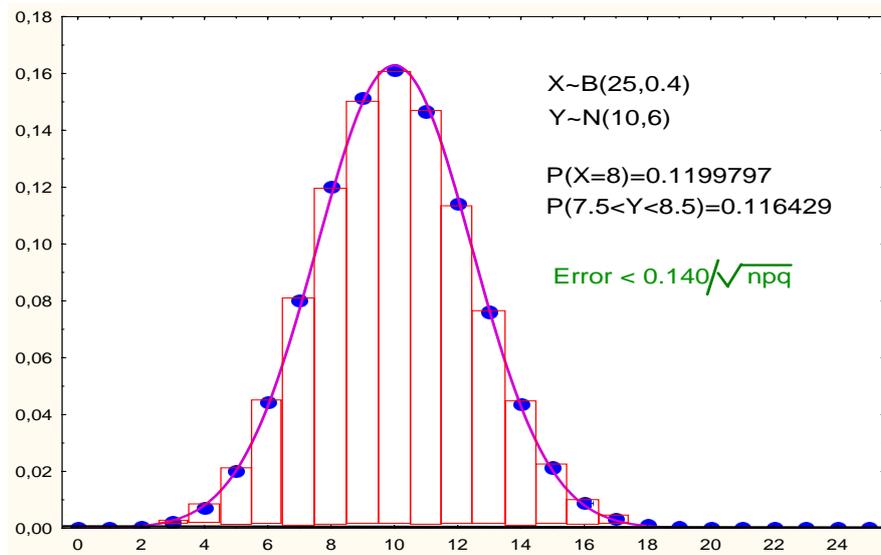


Figura 2.12: Aproximación de la binomial por la normal

densidad de la normal en el intervalo $[k - 1/2, k + 1/2]$.

Observación 2.65. (Corrección por continuidad) El hecho de sumar y restar $1/2$ en la aproximación suele llamarse corrección por continuidad. Se ha calculado que una cota superior para el error cometido en esa aproximación con corrección por continuidad es $0.140 / \sqrt{np(1-p)}$. \triangleleft

Observación 2.66. (Condiciones de validez de la aproximación) A la hora de aplicar la aproximación de la binomial por la normal suele exigirse en la práctica la siguiente condición: $np(1-p) \geq 5$. \triangleleft

Observación 2.67. (Distribución binomial y distribución hipergeométrica) Una última aproximación que nos interesa es la aproximación del muestreo sin reemplazamiento por el muestreo con reemplazamiento, es decir, la aproximación de la distribución hipergeométrica por la binomial según se explica a continuación. Hemos comentado en temas precedentes que la distribución binomial se usa, por ejemplo, como modelo para el número de individuos que poseen una cierta enfermedad entre n individuos seleccionados al azar de una población. Hay que tener en cuenta que eso es exactamente así en el caso de que cada individuo, una vez decidido si posee o no la enfermedad en cuestión, sea devuelto a la población de manera que pueda ser seleccionado posteriormente (se habla en ese caso de muestreo con reemplazamiento); si la población total tiene un número finito N de individuos y los individuos tras ser examinados no se devuelven a la población (se habla entonces de muestreo sin reemplazamiento), la distribución del número de individuos enfermos entre los n seleccionados de ese modo no sería ya la distribución binomial, sino que es una distribución discreta llamada hipergeométrica que no estudiaremos en detalle; no obstante, se prueba que la distribución

hipergeométrica puede aproximarse por la binomial cuando el cociente n/N es pequeño y N es grande (en la práctica suelen exigirse las siguientes condiciones para que la aproximación se considere aceptable: $n/N \leq 0.1$ y $N > 40$). <

Observación 2.68. (Máquina de Galton) La imagen anterior representa una máquina (llamada de Galton) en la que se introducen bolas por el embudo superior; tras sucesivos choques con los pernos dispuestos en filas de la máquina, cada bola cae en uno de los 8 depósitos en la parte inferior. Si, en cada perno, cada bola tiene probabilidad $1/2$ de ir a la derecha o a la izquierda, cada bola cae en el depósito k -ésimo con probabilidad $f(k) = \binom{7}{k}2^{-7}$, donde f es la función de probabilidad de la distribución $B(7, 1/2)$. Si se lanza un número grande de bolas, éstas terminarán dibujando en la parte inferior algo parecido a una campana de Gauss, poniendo de manifiesto el llamado efecto límite central, es decir, la aproximación de la binomial $B(7, 1/2)$ por la distribución normal $N(3.5, 1.75)$ ³ (pues el número de bolas en el depósito k -ésimo es proporcional a $f(k)$).

De hecho, si son 64 las bolas que introducimos por el embudo superior y, en cada perno, se fuesen a derecha e izquierda exactamente la mitad de las bola que llegan a él, obtendríamos la siguiente distribución:

			64												
			32		32										
			16		32		16								
			8		24		24		8						
			4		16		24		16		4				
			2		10		20		20		10		2		
			1		6		15		20		15		6		1

La tabla siguiente, conocida como triángulo de Tartaglia (también llamado de Pascal), recoge en cada perno el número de caminos diferentes que, partiendo del embudo, conducen al mismo. Concretamente, al perno k -ésimo de la fila n -ésima ($n = 0, 1, \dots, k = 0, 1, \dots, n$) llegan exactamente $\binom{n}{k}$ caminos:

³Se trabaja con 7 filas por no complicar el ejemplo, aunque para garantizar las condiciones de validez de aproximación de la binomial por la normal que hemos fijado necesitaríamos un número de filas superior a 20.

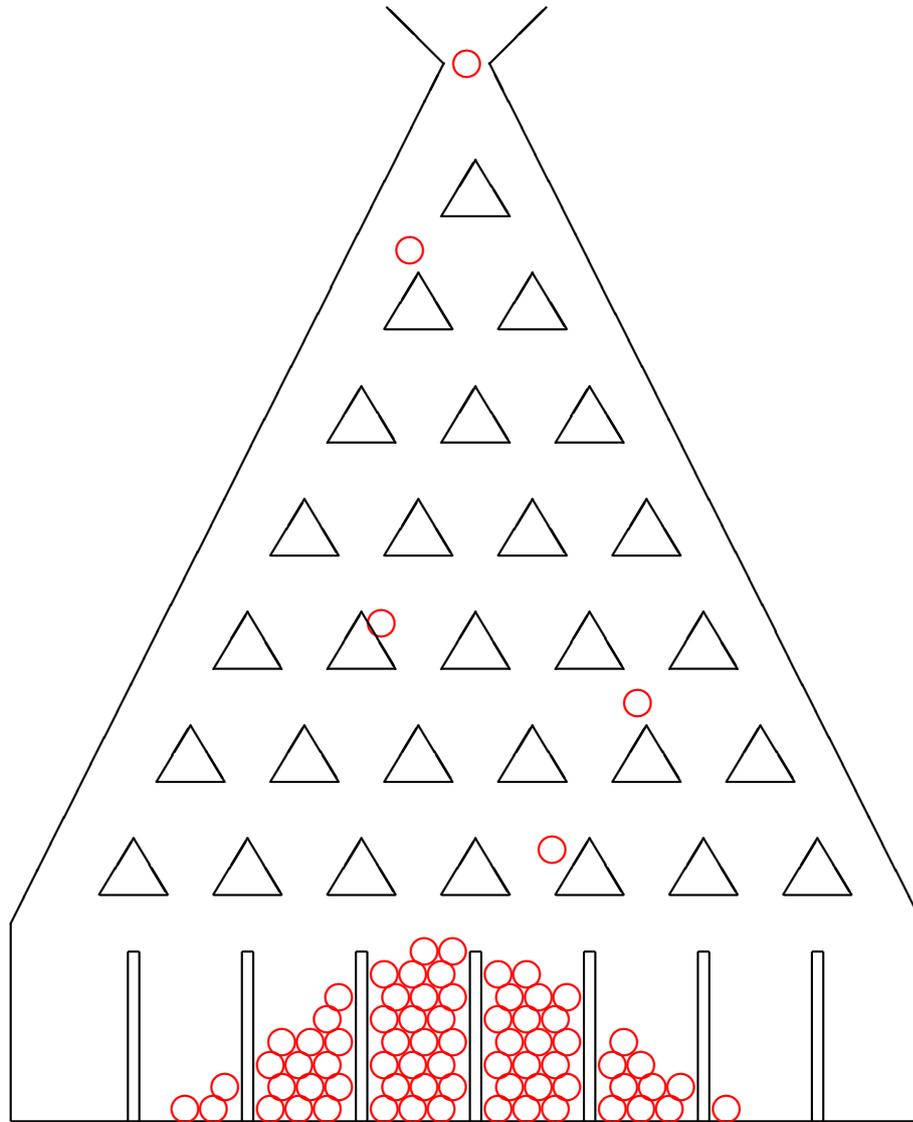


Figura 2.13: Máquina de Galton

			1							
			1		1					
			1	2		1				
			1	3	3		1			
			1	4	6	4		1		
			1	5	10	10	5		1	
			1	6	15	20	15	6		1

△

2.4.6. Muestreo estratificado

Ya hemos comentado que el problema de selección de muestras de una población es, en general, un problema importante y difícil. El método de muestreo utilizado debe garantizar la muestra representatividad de la muestra, para que las conclusiones obtenidas de ésta sean extrapolables a la población objeto de estudio. Los modelos estadísticos que proponemos en la segunda parte de esta obra para llevar a cabo la extrapolación pueden ser más o menos sensibles a las suposiciones (como las de normalidad, igualdad de varianzas, etc) que hacemos sobre las variables que generan los datos obtenidos en la experimentación, pero son dramáticamente sensibles a la presencia inadvertida (o, lo que es más grave, ignorada) en la población objetivo de subpoblaciones heterogéneas en lo que a la distribución de las variables de interés se refiere. Parece obvio, por ejemplo, que suponer normalidad para la distribución de la estatura de los jóvenes de una población es cuestionable, y no tanto por la naturaleza de la variable estatura como por la influencia de sobre ella de otras variables como pudieran ser la edad y el sexo; más razonable sería la suposición de normalidad para la estatura de los varones de 20 años de esa población.

En ocasiones la población aparece dividida de manera natural en subpoblaciones o estratos. La población de jóvenes a que nos referíamos puede ser dividida en varones y hembras si entendemos que el sexo es una variable relevante en el estudio que deseamos realizar, o en varios estratos atendiendo además a la variable edad, de forma que los estratos sean homogéneos en cuanto a sexo y edad; un país aparece naturalmente dividido en regiones (o en provincias); un área de salud está dividida en varias zonas de salud con características económicas y socioculturales peculiares. En ese caso, podemos aprovechar la información disponible sobre los diferentes estratos a la hora de extraer la muestra y, más que realizar un muestreo aleatorio simple en toda la población, podemos repartir el tamaño muestral deseado entre los diferentes estratos, de tal suerte que los datos obtenidos nos permitan hacer inferencia estadística tanto sobre la población en general como sobre cada uno de los estratos considerados. Un muestreo de este tipo, en el que se realiza un muestreo aleatorio simple sobre cada uno de los estratos, recibe el nombre de muestreo estratificado. En la medida en que los estratos sean homogéneos (en relación a las variables incluidas en el estudio), las inferencias que podamos llevar a cabo en cada uno de ellos serán fiables, y tanto más cuanto mayor sea el tamaño muestral en el estrato en cuestión. Pero, para que la inferencia en la población total sea fiable, conviene ser cuidadoso a la hora de repartir el tamaño muestral global entre los diferentes estratos, con la intención de que un estrato tenga en la muestra un peso similar al que de hecho tiene en la población general. En cualquier caso, estratificando conseguimos eliminar posibles fuentes de variabilidad en la muestra. Es conveniente, por ejemplo, en estudios multicéntricos (en los que aparecen implicados varios hospitales), estratificar por centro. Distinguiremos aquí tres formas posibles de reparto (se suele hablar técnicamente de afijación) del tamaño muestral: la afijación uniforme, en la que la muestra se reparte por igual entre los distintos estratos, la afijación proporcional, en la que la muestra se reparte entre los estratos proporcionalmente al tamaño que éstos tienen en la población total, y la afijación óptima, en la que, además del tamaño de los estratos, se tiene en cuenta la variabilidad de los estratos.

¿Verdadero o Falso? Capítulo 2

Decidir si cada una de las proposiciones siguientes es verdadera (V) o falsa (F), justificando la respuesta.

- 2.1** Una pregunta de un examen de Bioestadística contiene 5 proposiciones sobre las que el alumno debe decir si son verdaderas o falsas (supongamos que no se puede dejar sin respuesta ninguna pregunta); cada respuesta acertada suma dos puntos y cada respuesta errónea resta 1 punto. Si a cada una de las afirmaciones responde V o F según que salga cara o cruz, resp., en lanzamientos consecutivos de una moneda equilibrada, la puntuación media obtenida es igual a 5 puntos.
- 2.2** Dos v.a.r. normales que tienen la misma media y la misma varianza tienen también la misma distribución.
- 2.3** Si estamos dispuestos a admitir el error de declarar no normales a un 5% de los individuos normales, un límite inferior de normalidad para una v.a. X con distribución $N(2.5, 2)$ es 0.174. **SOLUCIÓN: PÁG. 341**
- 2.4** Si X tiene distribución $N(5, 1.1^2)$ entonces $P(X \geq 3.9) = 0.8413$.
- 2.5** Si X es una v.a.r. con distribución $t(n)$, entonces $P(X \leq -a) + P(X \leq a) = 1$.
- 2.6** Sean X una v.a.r. con distribución $N(4, 4\sigma^2)$ e Y una v.a.r. con distribución $N(0, 1)$. Entonces $P(X - 4 < 2\sigma) = P(Y < 1)$. **SOLUCIÓN: PÁG. 341**
- 2.7** Sea X una v.a.r. con distribución normal $N(2, 1)$. Entonces $P(2X - 1 \leq 2) > 0.275$.
- 2.8** Sea $a > 0$. Si X es una v.a.r. con distribución $N(-2, a)$, entonces $P(X \geq -5) + P(X > 1) = 1$.
- 2.9** Si X_1, \dots, X_n es una muestra de tamaño n de una distribución real, la v.a. S^2 tiene distribución $\chi^2(n - 1)$. **SOLUCIÓN: PÁG. 341**

- 2.10** El intervalo $[\mu - \sigma, \mu + \sigma]$ contiene aproximadamente el 95% de la distribución $N(\mu, \sigma^2)$.
- 2.11** La función de probabilidad de una v.a. numérica discreta X asigna a cada valor x de la variable la probabilidad de que X tome un valor menor o igual que x .
- 2.12** Si X e Y son dos v.a.r. independientes cada una de ellas con distribución $N(0, 1)$, entonces $P(\{X \leq 1\} \cap \{Y \leq 1\}) = 0.7078$. **SOLUCIÓN: PÁG. 341**
- 2.13** Sea X una v.a.r. con distribución normal $N(2, 1)$. Entonces $P(X = 2) = 1/2$.
- 2.14** Sean $a > 0$ y X una v.a. con distribución normal de media 8; entonces $P(X > 8 + a) < \frac{1}{2}$.
- 2.15** Si X es una v.a. con distribución $N(\mu, \sigma^2)$, entonces $P(X \geq \mu - \sigma^2) = P(X \leq \sigma^2 - \mu)$. **SOLUCIÓN: PÁG. 341**
- 2.16** Si X es una v.a.r. cualquiera, entonces $P(X \leq 1) + P(X > 1) < 1$.
- 2.17** Es razonable suponer independientes dos v.a. que describen dos características numéricas diferentes en un mismo individuo elegido al azar en una cierta población.
- 2.18** Se eligen 15 alumnos al azar y con reemplazamiento de un colegio en el que el 70% de los alumnos han pasado la varicela. El número medio de alumnos que han pasado la varicela entre los 15 elegidos es igual a 7. **SOLUCIÓN: PÁG. 341**
- 2.19** Si X es una v.a.r. con distribución normal $N(1, 4)$, entonces $P(X \geq -1.7) = 0.0885$.
- 2.20** En una población Ω , denotaremos por $X(\omega)$ la presión sanguínea de un individuo ω . Consideremos los sucesos $A = \{X \geq 90\}$ y $B = \{75 \leq X \leq 100\}$. Entonces $A \cap B = \{X \geq 90\}$.
- 2.21** Sea X una v.a.r. con distribución normal $N(2, 2)$. Entonces la v.a. $3X - 2$ tiene distribución $N(4, 18)$. **SOLUCIÓN: PÁG. 341**

- 2.22** Si X es una v.a.r. con distribución normal de media 5, entonces $P(X < 10) > P(X > 1)$.
- 2.23** Sea X es la v.a. que describe el nivel en suero de una cierta sustancia en los individuos normales de una población y supongamos que tiene distribución $N(5, 1.1^2)$. Entonces la v.a. $5X + 2$ tiene distribución $N(27, 8.05)$.
- 2.24** Sea X_1, \dots, X_{10} una muestra de tamaño 10 de una distribución $N(10, 100)$. La probabilidad de que $\sqrt{10} \cdot |\bar{X} - 10|/S \leq 2,228$ es igual a 0.95. **SOLUCIÓN: PÁG. 341**
- 2.25** Supongamos que el nivel de glucosa en suero tiene distribución normal $N(90, 100)$ en una cierta población. Un límite inferior de normalidad con error del 10% para el nivel de glucosa en suero es igual a 73.55.
- 2.26** La distribución de la presión intraocular en la población general es aproximadamente normal $N(16, 9)$. La probabilidad de que la presión intraocular sea superior a 17.5 es igual a 0.6915.
- 2.27** La distribución de la presión intraocular en la población general es aproximadamente normal $N(16, 9)$. Un intervalo de normalidad para la presión intraocular con un error de intervalo del 5% es $[10, 12, 21, 88]$. **SOLUCIÓN: PÁG. 341**
- 2.28** La varianza de la media muestral de una muestra de tamaño n de una distribución real es siempre mayor o igual que la varianza de esa distribución.

Problemas del Capítulo 2

- 2.1** Se ha estimado que, entre los hombres de una cierta población con edad entre 30 y 40 años que fumaron alguna vez, el número medio de años que estuvieron fumando es 12.8 con una desviación típica de 5 años.
- a) Asumiendo normalidad para los tiempos observados, ¿qué proporción de hombres en ese grupo de edad han fumado durante más de 15 años?
- b) En esa población se eligen al azar 80 individuos con edad entre 30 y 40 años que fumaron alguna vez. ¿Cuál es la probabilidad de que al menos 20 de ellos hayan fumado durante más de 15 años?
- 2.2** Sea X una v.a. que describe el número de casos de otitis en el oído medio en los dos primeros años de vida de los niños de una cierta población. Se supone conocido que la función de probabilidad de X es la aplicación $f : \{0, 1, 2, 3, 4, 5, 6\} \rightarrow [0, 1]$ definida por: $f(0) = 0.129$, $f(1) = 0.264$, $f(2) = 0.271$, $f(3) = 0.185$, $f(4) = 0.095$, $f(5) = 0.039$, $f(6) = 0.017$.
- a) Calcula $E(X)$ e interpreta el valor obtenido.
- b) ¿Cuál es la probabilidad de que, entre 40 niños de más de 2 años elegidos al azar de esa población, al menos 30 hayan sufrido 2 o menos episodios de otitis en el oído medio en sus dos primeros años de vida?
- 2.3** Se sabe que entre los individuos que poseen una cierta enfermedad en una población, el 30% poseen el síntoma S. Entre los individuos de la población que poseen esa enfermedad se eligen 20 al azar. ¿Cuál es la probabilidad de que haya 5 o menos individuos con el síntoma S entre esos 20? **SOLUCIÓN: PÁG. 342**
- 2.4** Consideremos el espacio de probabilidad correspondiente al experimento aleatorio que consiste en un sólo lanzamiento al aire de un dado perfecto definido por el conjunto $\Omega = \{1, 2, 3, 4, 5, 6\}$ y la probabilidad P que asigna el valor $1/6$ a cada uno de los sucesos elementales de Ω . Consideremos también dos v.a. X e Y definidas en Ω y a valores en $\{0, 1\}$ del siguiente modo: si $1 \leq i \leq 6$, $X(i) = 0$ si i es un número primo

y $X(i) = 1$ si i no es primo; $Y(i) = 0$ si i es un número par e $Y(i) = 1$ si i es impar. (Nota: se recuerda que un número primo es aquel que no tiene más divisores que él mismo y la unidad; p.ej.: el 13 es primo y el 27 no).

a) Describe explícitamente los sucesos $\{X = 1\}$ e $\{Y = 0\}$ y las distribuciones P^X y P^Y de X e Y .

b) ¿Son independientes las v.a. X e Y ? (La independencia de esas v.a. equivaldría a que lo sean los sucesos “ser primo” y “ser par”).

c) Imagina que, en lugar de un dado, tenemos un tetraedro regular perfecto (cuatro caras numeradas del 1 al 4 equiprobables). ¿Son independientes los sucesos “ser primo” y “ser par”?

2.5 Supongamos que el peso X en Kg de los niños de 6 años de una cierta población tiene distribución normal de media 20 y desviación típica 2.

a) Calcula $P(18 < X < 22)$ y $P(X > 16)$.

b) Calcula los percentiles 10 y 40 de esa distribución y sus cuantiles 0.25 y 0.75.

c) Determina un intervalo de normalidad para el peso de esas niñas con un error del 10 %.

d) Se extrae una muestra aleatoria de 10 niñas de 6 años de edad. Calcula las probabilidades $P(18 < \bar{X} < 22)$ y $P(S^2 > 1.2)$.

e) ¿Qué distribución tiene la variable $\sqrt{10}(\bar{X} - 20)/S$? Calcula $P(\bar{X} \in [20 - 1.833 \cdot S/\sqrt{10}, 20 + 1.833 \cdot S/\sqrt{10}])$

f) Determina el intervalo que contiene el 95 % central de los valores de \bar{X} y decide si 17 Kg es un valor insólito para el peso medio de una muestra aleatoria de 10 niñas de 6 años.

2.6 Para elegir 10 personas al azar de una población de 780 personas, se enumeran éstas desde el 0 hasta el 779 y se preparan tres bombos de lotería con bolas del mismo peso y tamaño numeradas del 0 al 7 en el primer bombo, del 0 al 7 en el segundo y del 0 al 9 en el tercero. Para seleccionar un individuo, se saca una bola de cada bombo

y, tras anotar el resultado, las bolas son devueltas al bombo; esta operación se repite nueve veces más. ¿Se trata de un muestreo aleatorio simple? **SOLUCIÓN: PÁG. 342**

2.7 Supongamos conocido que el coeficiente de inteligencia (CI) se distribuye en una cierta población según una distribución $N(100, 225)$ y que de una persona con CI superior a 115 se dice que tiene un CI alto. Trabaja a lo largo de este problema con una precisión de dos cifras decimales.

a) ¿Cuál es la probabilidad de que un individuo de la población tenga CI alto?

b) Se eligen al azar 80 individuos de la población. ¿Cuál es el número medio de individuos con CI alto? Calcula la probabilidad de que a lo más un 15% de esos 80 individuos tenga un CI alto.

2.8 Para elegir 3 números entre el 0 y el 99999 mediante un muestreo aleatorio simple sin reemplazamiento, ¿podemos proceder del siguiente modo?: se utilizan cinco bombos de lotería y en cada uno de ellos se introducen 10 bolas exactamente iguales y numeradas del 0 al 9; se extrae una bola de cada bombo y se anota el número de cinco cifras obtenido; sin devolver las bolas extraídas a los bombos respectivos, se extraen dos bolas más de cada uno de ellos, obteniendo así los otros dos números buscados. Justifica la respuesta.

2.9 Se ha estimado que, para un cierto tipo de cáncer, la probabilidad de que un individuo sobreviva seis meses tras un tratamiento de quimioterapia es del 60%, mientras que la de que sobreviva un año es del 45%.

a) Calcula la probabilidad de que un paciente que ha sobrevivido seis meses sobreviva también un año.

b) Calcula la probabilidad de que de cinco pacientes elegidos al azar al menos 3 de ellos sobrevivan seis meses. **SOLUCIÓN: PÁG. 342**

2.10 Imagina que queremos elegir una muestra de n personas mayores de edad en la ciudad de Badajoz para que respondan a una encuesta sobre la atención sanitaria que reciben y que para ello salimos a la calle a las diez de la mañana y pasamos la encuesta a

las n primeras personas mayores de edad que nos cruzamos en un día determinado. ¿Crees que se trata de un muestreo aleatorio simple? Justifica la respuesta exponiendo diversas razones que te hagan creer eso.

2.11 Supongamos conocido que el 20% de los varones mayores de 14 años de la ciudad de Badajoz han sido declarados bebedores de riesgo mediante un test diseñado por Altisent y otros relativo al consumo de alcohol. Supondremos también que en la ciudad de Badajoz hay 45000 individuos varones mayores de 14 años y que, entre ellos, 10000 están adscritos al centro de salud de La Paz, de los cuales 2500 han sido declarados bebedores de riesgo. En el centro de salud de la Zona Centro se encuentran adscritos 4500 varones mayores de 14 años.

a) ¿Cuál es la probabilidad de que a lo sumo 20 entre 100 individuos elegidos al azar de esa población sean bebedores de riesgo?

b) ¿Cuál es la probabilidad de que como mínimo 2 de 10 varones mayores de 14 años elegidos al azar en el centro de salud de La Paz sean declarados bebedores de riesgo?

2.12 Sobre el conjunto $\Omega = \{0, 1, 2, 3, 4, 5\}$, que podría hacer las veces de espacio de las observaciones para el experimento aleatorio que consiste en lanzar al aire un dado (no necesariamente perfecto) con caras numeradas del 0 al 5, construir una probabilidad P (nótese que, para ello, es suficiente determinar las probabilidades de los sucesos elementales) en cada uno de los casos siguientes, y calcula de acuerdo con ella la probabilidad de sacar un número impar:

a) El dado es perfecto.

b) El resultado del lanzamiento sigue una distribución binomial $B(5, 0.5)$.

c) El resultado del lanzamiento sigue una distribución binomial $B(5, 0.1)$. **SOLUCIÓN:**

PÁG. 342

2.13 Supongamos conocido que el contenido medio de glucosa en sangre en un grupo de ratas diabéticas es 1.8 mg/ml, con una desviación típica $\sigma = 0.2$, y que la glucosa se distribuye normalmente en esa población.

a) Se extrae una muestra de tamaño 40 de esa población. Calcula la probabilidad de que al menos 5 de esas ratas tengan un nivel de glucosa inferior a 1.6.

b) Se extrae una muestra de tamaño 4 de esa población. Calcula $P(\bar{X} < 1.6)$.

2.14 Queremos elegir una muestra de n personas mayores de edad en la ciudad de Badajoz para que respondan a una encuesta sobre la atención sanitaria que reciben. Para ello elegimos de la guía telefónica n números al azar y pasamos la encuesta a la persona que coge el teléfono (si es que ésta es mayor de edad). ¿Crees que se trata de un muestreo aleatorio simple? Justifica la respuesta exponiendo diversas razones que te hagan creer eso.

2.15 Un test de diagnóstico que ha sido diseñado para la detección del cáncer de colon en una cierta población posee una sensibilidad del 55 % y una especificidad del 80 %, mientras que la prevalencia del cáncer de colon en esa población es del 15 %.

a) Entre 100 individuos con cáncer de colon, ¿cuál es la probabilidad de que el test resulte positivo para al menos 45 de ellos?

b) Si se eligen al azar 4 individuos para los que el test resultó positivo, ¿cuál es la probabilidad de que a lo sumo dos de ellos posean verdaderamente la enfermedad?

SOLUCIÓN: PÁG. 342

2.16 Supongamos que el 10 % de individuos hipertensos que toman el fármaco antihipertensivo A padecen molestias gastrointestinales (GI), mientras que para los que toman el fármaco B las molestias se producen en un 20 % de los casos. (a) Si las molestias GI de ambos fármacos se producen de forma independiente, ¿cuál es la probabilidad de que un paciente que toman simultáneamente ambos fármacos tenga molestias GI? (b) ¿Cuál es la probabilidad de que, de 100 individuos hipertensos tratados con el fármaco B, menos de la mitad tenga molestias GI?

2.17 Supongamos que los varones españoles sanos entre 65 y 79 años de edad el nivel de ácido úrico en la sangre es aproximadamente normal de media $340\mu\text{mol/l}$ y desviación

típica $\sigma = 80\mu\text{mol/l}$. a) ¿Qué proporción de esos individuos tiene nivel de ácido úrico en sangre entre 300 y 400? b) ¿Qué proporción de muestras de tamaño 4 en esa población tiene media muestral entre 300 y 400?

¿Verdadero o Falso? PRIMERA PARTE

Decidir si cada una de las proposiciones siguientes es verdadera (V) o falsa (F), justificando la respuesta.

- I.1** Si $x_{ij} = i + j$, $1 \leq i \leq 4$, $1 \leq j \leq 3$, entonces $\sum_{j=1}^3 x_{4j} - 1 = 15$.
- I.2** Si $x_{ij} = i + j$, $1 \leq i \leq 4$, $1 \leq j \leq 3$, entonces $\sum_{j=1}^3 (x_{4j} - 1) = 15$.
- I.3** Dos combinaciones de orden 2 en el conjunto $\{a, b, c, d\}$ son distintas si una de ellas contiene un elemento que no contiene la otra o si sus elementos aparecen en distinto orden.
- I.4** En una población Ω , sean H el subconjunto formado por los individuos hipertensos y F el formado por los individuos fumadores. Entonces $(H \cap F)^c = H^c \cup F^c$.
- I.5** La desviación típica de un conjunto de datos numéricos es una medida de dispersión del mismo expresada en las mismas unidades de medida que los datos.
- I.6** Si en un conjunto de datos numéricos, al menor de los datos se le resta una unidad, la media disminuye.
- I.7** En un histograma sobre intervalos de clase de diferente longitud, la altura de cada rectángulo es proporcional a la frecuencia de la clase correspondiente.
- I.8** Los conjuntos de datos numéricos x_1, \dots, x_n y $x_1 + c, \dots, x_n + c$ tienen la misma desviación típica.
- I.9** Si en un conjunto de datos numéricos x_1, \dots, x_n , a todos los datos se le suma la misma cantidad, la varianza no cambia.
- I.10** La media de un conjunto de datos numéricos es el número que mejor aproxima esos datos en el sentido de los mínimos cuadrados.
- I.11** La varianza de un conjunto de datos numéricos se mide en las mismas unidades que ellos.

- I.12** Dos conjuntos de datos con idéntico diagrama de caja coinciden.
- I.13** Sean $x_1 = 2$, $x_2 = -1$ y $x_3 = 3$. Entonces $(\sum_{i=2}^3 x_i)^2 = \sum_{i=1}^3 x_i$.
- I.14** Sea X la variable que registra el “nivel de colesterol” de los individuos de la población española. Los sucesos $\{X \geq 140\}$ y $\{X < 180\}$ son incompatibles.
- I.15** De un colectivo formado por 1000 individuos esquizofrénicos (de los que el 40 % estaban casados), 60 de ellos terminaron suicidándose durante los 10 años siguientes al comienzo del estudio; entre ellos, 24 estaban casados. Si C y S denotan los sucesos “estar casado” y “haberse suicidado”, entonces $P(C|S) \neq P(C)$.
- I.16** Cualquier suceso A en el espacio de probabilidad (Ω, P) es independiente del suceso seguro Ω .
- I.17** Si A es un suceso en el espacio de probabilidad (Ω, P) entonces $P(A|\Omega) = P(A)$.
- I.18** En el experimento aleatorio que consiste en dos lanzamientos de un dado con caras numeradas del 1 al 6, los sucesos $A =$ “la suma de los resultados es ≤ 5 ” y $B =$ “el primer lanzamiento es ≥ 5 ” son incompatibles.
- I.19** Dados dos sucesos cualesquiera A y B , $P(A \cup B) + P(A \cap B) = P(A) + P(B)$.
- I.20** Sean A , B y C sucesos y supongamos que $A \cap B = \emptyset$ y $A \cup B = \Omega$. Entonces $P(C) = P(A) \cdot P(C|A) + P(B) \cdot P(C|B)$.
- I.21** De los mil alumnos de un colegio, 400 son niños y 600 niñas. La mitad de las niñas tienen algún problema de visión. Si también la mitad de los alumnos del colegio tienen algún problema de visión, podemos afirmar que en ese colegio los sucesos “ser niña” y “presentar algún problema de visión” son dependientes.
- I.22** En una población de 1200 habitantes, 450 personas sufrieron la gripe el pasado invierno, 240 de las cuales eran varones. Sabiendo que hay 640 varones en la población podemos afirmar que los sucesos “ser mujer” y “haber sufrido la gripe” en esa población son independientes.

- I.23** Los tests de diagnóstico con pocos falsos negativos resultan útiles a la hora de confirmar la enfermedad.
- I.24** $P(A|B) = P(B|A)$ si $P(A), P(B) > 0$.
- I.25** Sean A y B sucesos tales que $A \subset B$. Entonces $P(B|A) = 1$.
- I.26** En el experimento aleatorio que consiste en dos lanzamientos de una moneda perfecta, los sucesos “sacar dos caras” y “sacar una cara y una cruz” son equiprobables.
- I.27** El valor predictivo positivo de un test de diagnóstico es la probabilidad de que el test resulte positivo para un paciente del que se sabe que tiene la enfermedad.
- I.28** En una población de 1600 habitantes, 650 son del grupo sanguíneo A, 340 de los cuales son mujeres. Sabiendo que hay 840 mujeres en la población podemos afirmar que los sucesos “ser mujer” y “no ser del grupo A” en esa población no son independientes.
- I.29** Consideremos una urna que contiene bolas del mismo peso, forma y tamaño de las que 4 son blancas, 1 negra, 4 rojas y una azul. Los sucesos A = “sacar bola blanca o roja” y B = “sacar bola negra o roja” en una extracción son independientes.
- I.30** Si A y B son sucesos tales que $P(B|A) = P(B)$, entonces A y B son independientes.
- I.31** Utilizamos “consumo elevado de grasas” como test de diagnóstico para predecir el desarrollo de una enfermedad cardiovascular (ECV) en una determinado sector de una población. Si sabemos que tienen un elevado consumo de grasas el 80% de las personas con ECV y el 35% de las personas sin ECV, la probabilidad de falso negativo del test es 0.2.
- I.32** En un lanzamiento de un dado equilibrado, los sucesos $\{3, 4\}$ y $\{3, 5, 6\}$ son independientes.
- I.33** En el conjunto de las familias de un país con padre, madre y 2 hijos, la unión de los sucesos “el padre tiene gripe” y “la madre tiene gripe” es el suceso “ambos padres tienen gripe”.

- I.34** Se seleccionan 10 alumnos de una clase de 100 ordenados alfabéticamente del siguiente modo: se elige un número k al azar del 1 al 10 y se seleccionan los alumnos $k, k + 10, k + 20, \dots, k + 90$. Se trata de un muestreo aleatorio simple sin reemplazamiento.
- I.35** Cuanto mayor es la varianza de una distribución normal más “aplastada” es su función de densidad.
- I.36** Si X es una v.a. con distribución normal de media 5 entonces $P(X \leq 0.5) < P(X \geq 10)$.
- I.37** Si X es una v.a.r. con distribución $N(2.5, 4)$ entonces $P(X > 0) = 0.8944$.
- I.38** Sean $a, \sigma^2 > 0$. Si X es una v.a.r. con distribución $N(0, \sigma^2)$, entonces $P(X \geq -a) + P(X \leq a) > 1$.
- I.39** Si X_1, \dots, X_n es una muestra de tamaño n de una distribución $N(\mu, \sigma^2)$ y $a \in \mathbb{R}$, entonces $\text{Var}(a\bar{X}) = \frac{a^2\sigma^2}{n}$.
- I.40** Se lanzan dos dados perfectos al aire n veces. La distribución del número de ocasiones en que la suma es igual a 7 es la distribución binomial $B(n, 1/6)$.
- I.41** Si X_1 y X_2 son v.a.r. independientes e idénticamente distribuidas con distribución $N(0, 1)$, entonces $X_1 + X_2$ tiene distribución $N(0, 2)$.
- I.42** Dos v.a.r. con distinta distribución tienen también medias distintas.
- I.43** Si X_1, \dots, X_{20} es una muestra de tamaño 20 de una distribución $N(\mu, 2)$, entonces $\text{Var}(\bar{X}) = 0.1$
- I.44** Una v.a. discreta X que sólo toma dos valores tiene distribución de Bernoulli.
- I.45** Tipificar una v.a. X con distribución normal $N(\mu, \sigma^2)$ consiste en restar a X su media μ y dividir la diferencia por la varianza.

- I.46** En el experimento que consiste en el lanzamiento de dos dados perfectos, sea X la v.a. suma de los resultados obtenidos. La contraimagen del suceso $\{X \text{ es par}\}$ es el conjunto de pares $(i, j) \in \{1, \dots, 6\}^2$ tales que $i = j$.
- I.47** Sean X e Y v.a. discretas independientes con distribuciones respectivas $B(10, 0.4)$ y uniforme sobre el conjunto $\{1, 2, 3, 4\}$. Entonces $P(\{X \leq 2\} \cap \{Y > 2\}) = 0.167$.
- I.48** Dos v.a.r. con la misma distribución tienen también la misma media y la misma varianza.
- I.49** Sean X una v.a. con distribución $N(\mu, \sigma^2)$ e Y una v.a.r. con distribución $N(0, 1)$. Entonces $P(|X - \mu| \leq \sigma) = P(-1 \leq Y \leq 1)$.
- I.50** Sean $\mu \in \mathbb{R}$ y X e Y v.a.r. con distribuciones respectivas $N(\mu, 1)$ y $N(\mu, 4)$. Entonces $P(\mu - 2 \leq Y \leq \mu + 2) = P(\mu - 1 \leq X \leq \mu + 1)$.
- I.51** Si en una cierta población el cinco por ciento de los niños desarrollan una bronquitis crónica en el primer año de vida, la distribución del número de niños entre veinte elegidos al azar de esa población que desarrollaron esa enfermedad en su primer año de vida es una distribución uniforme discreta en el conjunto $\{1, 2, \dots, 20\}$.
- I.52** Sea X una v.a.r. con distribución normal $N(2, 4)$. Entonces la v.a. $\frac{X-2}{2}$ tiene distribución $N(0, 1)$.
- I.53** Se propone el siguiente juego: pagamos 3 euros por lanzar un dado equilibrado, y ganamos 1 euro si obtenemos un 1, 2 euros si obtenemos un dos, etc. El juego es desfavorable para el jugador.
- I.54** Dos v.a.r. que tienen la misma media y la misma varianza tienen también la misma distribución.
- I.55** La media de una v.a.r. discreta con distribución uniforme es la media aritmética de los valores de la variable.
- I.56** Es razonable suponer independientes dos v.a. que describen una misma característica numérica en diferentes individuos elegidos al azar en una cierta población.

- I.57** Supongamos que la temperatura (tomada bajo el brazo y en grados centígrados) de los individuos no enfermos de una población sigue una distribución normal $N(36.8, 0.03)$. Sabiendo que una temperatura de C grados centígrados equivale a $F := 32 + \frac{9}{5}C$ grados fahrenheit, se verifica que la temperatura en grados fahrenheit sigue una distribución normal $N(98.24, 0.0972)$.
- I.58** Si el nivel de colesterol de los individuos de una población sigue una distribución normal de media 175 y desviación típica 25, aprox. un 68% de los individuos de la población tendrán nivel de colesterol comprendido entre 150 y 200.
- I.59** La distribución binomial $B(100, 0.2)$ se puede aproximar por la distribución normal $N(20, 16)$.
- I.60** El colegio oficial de odontólogos y estomatólogos de una región española ha realizado un estudio aleatorio entre los pacientes que acudieron a sus consultas la semana pasada. De 1000 fichas analizadas, 500 habían realizado al menos una visita el año anterior. Se puede concluir que un 50% de la población de esa región acude al dentista al menos una vez al año.

PROBLEMAS DE LA PRIMERA PARTE

- I.1** En un juego con una baraja española de 40 cartas, cada jugador recibe 4 cartas. ¿Cuántos grupos distintos de 4 cartas se pueden formar?
- I.2** Cuatro individuos forman parte de un ensayo clínico y cada uno de ellos recibirá uno y sólo uno de cuatro fármacos diferentes. ¿Cuántas posibilidades tenemos de asignar un fármaco a cada individuo?
- I.3** a) Sean $x_1 = 2$, $x_2 = -1$, $x_3 = 1$, $x_4 = 2$ y $x_5 = 3$. Calcula a.1) $\sum_{i=1}^5 x_i$, a.2) \bar{x} , a.3) $\sum_{i=2}^4 x_i^2$, a.4) $\sum_{i=1}^5 x_i - 3$, a.5) $\sum_{i=1}^5 (x_i - 3)$, a.6) $\sum_{i=1}^3 \sum_{j=2}^4 x_i x_j$.
b) Sea $x_i = i + 1$, $i = 0, 1, \dots, 6$. Calcula b.1) $\sum_{i=0}^6 x_i^2$, b.2) $\sum_{i=1}^4 \frac{x_i}{x_{i+1}}$ y b.3) $\sum_{i=1}^4 \frac{x_i}{x_{i+1}}$.
- I.4** (★) Realiza una estadística descriptiva para los siguientes conjuntos de datos que incluya media, mediana, desviación típica, rango intercuartílico, valores mínimo y máximo, histograma (con 5 intervalos de clase) o diagrama de barras, y diagrama de caja.
- a) 1, 2, 3, 4, 5, 6, 7, 8, 9 y 10.
b) 1, 1, 3, 3, 5, 5, 7, 7, 10 y 10.
c) 1, 1, 1, 1, 1, 10, 10, 10, 10 y 10.
d) 1, 1, 1, 1, 1, 1, 1, 1, 1 y 10.
e) 1, 5, 5, 5, 5, 5, 5, 5, 5 y 10.
f) 51.6, 48.7, 42.6, 38, 50.1, 60.2, 41.3, 47.6, 40.6, 49.5, 42, 55.1, 46.2, 52.1.
- I.5** Los 720 alumnos de un colegio han sido clasificados según su color de pelo (claro u oscuro) y según las notas de la primera evaluación (aprobado o suspenso), habiéndose observado los siguientes hechos: 1) esas dos variables dicotómicas son independientes en esa población; 2) 600 alumnos aprobaron la primera evaluación; 3) 540 alumnos tienen el pelo oscuro.
- ¿Cuántos alumnos con el pelo claro han suspendido?

- I.6** En un juego con una baraja española de 40 cartas, cada jugador recibe 4 cartas. ¿Cuál es la probabilidad de que un jugador reciba los cuatro ases?
- I.7** Se ha estimado que, para un cierto tipo de cáncer, la probabilidad de que un individuo sobreviva seis meses tras un tratamiento de quimioterapia es del 60 %, mientras que la de que sobreviva un año es del 45 %. Calcula la probabilidad de que un paciente que ha sobrevivido seis meses sobreviva también un año.
- I.8** (Continuación del Ejemplo 1.40) En un estudio sobre afecciones cardiacas en individuos entre 40 y 50 años de edad en una cierta población, se han considerado como factores de riesgo el hecho de ser fumador (suceso F) y un consumo excesivo de grasas saturadas en la dieta habitual (suceso G), habiéndose detectado que las probabilidades de los sucesos $F^c \cap G^c$, $F \cap G^c$, $F^c \cap G$ y $F \cap G$ son 0.05, 0.1, 0.35 y 0.5 entre los individuos con algún tipo de afección coronaria (suceso E), y 0.5, 0.3, 0.15 y 0.05 entre los individuos que no poseen problemas cardiacos.
- a) Calcula las probabilidades de los siguientes sucesos: “ser fumador”, “consumo excesivo de grasas saturadas”, “ser fumador y consumir demasiadas grasas saturadas”.
- b) ¿Son independientes los sucesos F y G ? Compara $P(F)$ y $P(F|G)$ e interpreta el resultado.
- I.9** Supongamos conocido que el 20 % de los varones mayores de 14 años de la ciudad de Badajoz han sido declarados bebedores de riesgo mediante un test diseñado por Altisent y otros relativo al consumo de alcohol. Supondremos también que en la ciudad de Badajoz hay 45000 individuos varones mayores de 14 años y que, entre ellos, 10000 están adscritos al centro de salud de La Paz, de los cuales 2500 han sido declarados bebedores de riesgo. En el centro de salud de la Zona Centro se encuentran adscritos 4500 varones mayores de 14 años.
- a) ¿Son independientes los sucesos “ser bebedor de riesgo” y “estar adscrito al centro de salud de La Paz”? Justifica la respuesta e interprétala en términos del consumo de alcohol entre varones de más de 14 años en el área de salud de La Paz y en la ciudad de Badajoz.

b) ¿Cuántos varones mayores de 14 años de la Zona Centro serían declarados bebedores de riesgo si los sucesos “ser bebedor de riesgo” y “estar adscrito al centro de salud de la Zona Centro” fuesen independientes?

c) ¿Son independientes los sucesos “estar adscrito al centro de salud de la Zona Centro” y “estar adscrito al centro de salud de la La Paz”?

d) Se han elegido al azar 5 individuos varones mayores de 14 años de un mismo centro de salud y se ha observado que tres de ellos son bebedores de riesgo; si se sabe que ese centro es el de La Paz o el de la Zona Centro, ¿a qué centro de salud crees que pertenecen esos 5 individuos? (*Indicación:* utiliza el principio de máxima verosimilitud esbozado en el problema 5).

I.10 Cuatro individuos forman parte de un ensayo clínico y cada uno de ellos recibirá uno y sólo uno de cuatro fármacos diferentes. Si eres uno de esos cuatro individuos y prefieres recibir el fármaco A, ¿cuál es la probabilidad de que ello ocurra?

I.11 Supongamos que el diámetro en micras (denotémoslo por X) de los hematíes de los individuos normales de una población sigue una distribución $N(7.5, 0.04)$ y que nos interesa conocer un valor máximo para el diámetro de los hematíes de los individuos normales.

a) Determina un límite superior M de normalidad para X con un error del 5%.

b) Supongamos ahora que adoptamos el criterio de declarar enfermo (cirrótico, en este caso) a todo individuo con diámetro de hematíes superior a M ; ¿qué porcentaje de individuos normales serán declarados cirróticos incorrectamente con este nuevo test de diagnóstico?

c) Supongamos conocido que el diámetro de los hematíes de los individuos cirróticos de la población sigue una distribución $N(8.5, 0.36)$; ¿qué porcentaje de individuos cirróticos serán declarados incorrectamente normales?

d) ¿Qué sensibilidad y especificidad tiene el test de diagnóstico que acabamos de construir?

e) Si la prevalencia de la cirrosis en esa población es del 10 %, determina los valores predictivos positivo y negativo del test.

f) ¿Qué ocurre con la sensibilidad y la especificidad del test de diagnóstico que acabamos de diseñar si el error del límite superior de normalidad es del 10 %?

I.12 El contenido de una cierta sustancia por cada cápsula de un medicamento sigue una distribución normal de media 15 mg y desviación típica 3 mg.

a) Si la cápsula es efectiva cuando el contenido de esa sustancia es superior a 12 mg, ¿cuál es la probabilidad de que la cápsula sea efectiva?

b) En un frasco de 20 unidades, ¿cuál es la probabilidad de que haya como máximo una cápsula no efectiva?

c) ¿Cuál es la probabilidad de que el contenido de sustancia sea inferior a 15 mg entre las cápsulas efectivas?

I.13 Se desea extraer una muestra con reemplazamiento de tamaño 12 de un grupo de 60 pacientes. Para ello se numeran los pacientes desde el 00 hasta el 59 y se eligen 12 bloques consecutivos de dos dígitos de una tabla de números aleatorios y se procede del siguiente modo: si el número obtenido está entre 00 y 59, el paciente correspondiente es elegido; si no, al número extraído se le resta 59 y se elige el paciente correspondiente a la diferencia. ¿Es éste un muestreo aleatorio simple?

I.14 Los 720 alumnos de un colegio han sido clasificados según su color de pelo (claro u oscuro) y según las notas de la primera evaluación (aprobado o suspenso), habiéndose observado los siguientes hechos: 1) esas dos variables dicotómicas son independientes en esa población; 2) 600 alumnos aprobaron la primera evaluación; 3) 540 alumnos tienen el pelo oscuro.

Se eligen al azar y con reemplazamiento 64 alumnos del colegio. ¿Cuál es la probabilidad de que 25 de ellos tengan el pelo claro o hayan suspendido?

I.15 Para extraer una muestra de tamaño 10 de una población de 1500 personas, se numeran éstas desde el 0 hasta el 1499 y se preparan tres bombos de lotería con bolas

exactamente iguales numeradas del 0 al 14 en el primer bombo y del 0 al 9 en los otros dos. Para seleccionar un individuo, se saca una bola de cada bombo y, tras anotar el resultado, las bolas son devueltas al bombo; esta operación se repite nueve veces más. ¿Se trata de un muestreo aleatorio simple?

I.16 De una baraja española de 40 cartas se eligen 4 al azar con reemplazamiento. Denotemos por X la v.a. que describe el número de oros que aparecen entre esas cuatro cartas.

a) ¿Cuál es la distribución de la v.a. X ?

b) Si f denota la función de probabilidad de la v.a. X , determina los valores $f(0)$ y $f(1)$.

I.17 Se considera el experimento aleatorio que consiste en extraer una carta de una baraja española de 40 cartas. Denotemos por X (respectivamente, por Y) la v.a. que asigna a cada carta el palo (respectivamente, el número) de la misma (supongámoslas numeradas del 1 al 10, en lugar de 1, 2, ..., 7, 10, 11 y 12).

a) Describir el espacio de probabilidad que corresponde a ese experimento aleatorio (antes de extraer una carta se baraja el mazo, de tal suerte que todas las cartas se pueden suponer equiprobables). ¿Cuál es la distribución de probabilidad de X ?

b) ¿Son independientes las v.a. X e Y ? Nota: para probar que X e Y son independientes habrá que probar que, cualesquiera que sean $i = \text{oro, copa, espada, bastos}$ y $j = 1, 2, \dots, 10$, $P(\{X = i\} \cap \{Y = j\}) = P(X = i)P(Y = j)$; para probar que no son independientes, basta encontrar $i \in \{\text{oro, copa, espada, bastos}\}$ y $j \in \{1, 2, \dots, 10\}$ tal que $P(\{X = i\} \cap \{Y = j\}) \neq P(X = i)P(Y = j)$.

I.18 La tabla siguiente contiene las distribuciones categorizadas por sexo de los grupos sanguíneos para los individuos de una población (p. ej., la probabilidad de ser varón y del grupo A es 0.352):

	A	B	AB	O	Total
Varón	0.352	0.064	0.024	0.360	0.80
Mujer	0.054	0.048	0.008	0.090	0.20
Total	0.406	0.112	0.032	0.450	1.00

¿Son independientes las v.a. que asignan a cada individuo su sexo y su grupo sanguíneo?

- I.19** Se dispone de un tetraedro irregular cuyas caras están numeradas del 1 al 4. Se ha comprobado que el 1 se obtiene con probabilidad $1/2$, el 2 con probabilidad $1/4$, el 3 con probabilidad $1/6$ y el 4 con probabilidad $1/12$. Se efectúan dos lanzamientos independientes del tetraedro y se denota por X el resultado del primer lanzamiento y por Y el resultado del segundo lanzamiento.
- Describe el espacio de probabilidad correspondiente a ese experimento.
 - Determina todos y cada uno de los sucesos elementales de los sucesos $A = \{X - Y = 2\}$, $B = \{X + Y > 5\}$ y $C = \{|X - Y| = 2\}$.
 - Decide si los sucesos A y B del apartado (b) son independientes.
- I.20** Imagina que queremos elegir una muestra de n personas mayores de edad en la ciudad de Badajoz para que respondan a una encuesta sobre la atención sanitaria que reciben y que para ello salimos a la calle a las diez de la mañana y pasamos la encuesta a las n primeras personas mayores de edad que nos cruzamos en un día determinado. ¿Crees que se trata de un muestreo aleatorio simple? Justifica la respuesta exponiendo diversas razones que te hagan creer eso.
- I.21** En una cierta población, $[150, 200]$ es un intervalo de normalidad con un error del 5 % para el nivel de colesterol en suero (en mg/dl). Supongamos que el nivel de colesterol sigue una distribución normal.
- Determina la media y la desviación típica de esa distribución.
 - ¿Qué porcentaje de individuos de la población tienen nivel de colesterol superior a 225?

c) Para una muestra de tamaño 50 de esa población, determina las distribuciones de la media y varianza muestrales.

I.22 Se dispone de una moneda de la que se sabe que la probabilidad de sacar cara ($=1$) es $1/3$ y la de sacar cruz ($=0$) es $2/3$.

a) Describe el espacio de probabilidad que corresponde al experimento aleatorio que consiste en dos lanzamientos independientes de esa moneda.

b) Si X (respectivamente, Y) es la v.a. que describe el resultado del primer (respectivamente, del segundo) lanzamiento, determina las distribuciones de las v.a. $X + Y$ y $X - Y$.

c) Decide si las v.a. $X + Y$ y $X - Y$ son independientes.

I.23 Se dispone de una moneda con canto gordo de la que se sabe que la probabilidad de sacar cruz ($=0$) es 0.4, la de sacar cara ($=1$) es 0.4 y la probabilidad de que caiga de canto ($=2$) es 0.2.

a) Describir el espacio de probabilidad que corresponde al experimento aleatorio que consiste en dos lanzamientos independientes de esa moneda.

b) Si X (respectivamente, Y) es la v.a. que describe el resultado del primer (respectivamente, del segundo) lanzamiento, determina la distribución de la v.a. $X + Y$.

c) Decide si los sucesos $\{X + Y \leq 2\}$ y $\{X - Y = 1\}$ son independientes.

I.24 En 30 lanzamientos independientes de un dado perfecto, ¿cuál es la probabilidad de obtener 15 o menos números primos? Nota: recuerda que un número se dice primo si no posee otros divisores que él mismo y la unidad.

I.25 Un tetraedro regular tiene sus caras numeradas del 1 al 4; supongamos que el tetraedro es perfecto en el sentido de que los sucesos elementales 1, 2, 3 y 4 son equiprobables. Se lanza al aire dos veces el tetraedro y se denota por X la suma de los dos resultados obtenidos.

a) ¿Cuál es la función de probabilidad de la v.a. X ?

b) ¿Cuál es la media de X ?

c) ¿Son independientes los sucesos $A = \{X \text{ es impar}\}$ y $B = \{X \leq 3\}$?

I.26 En unas elecciones a delegado de alumnos en primer curso del Grado en Medicina, se encuentran presentes 100 alumnos: 70 chicas y 30 chicos. a) Se han de seleccionar, al azar y sin reemplazamiento, dos vocales entre los alumnos. ¿Cuál es la probabilidad de que los dos vocales sean chicos? b) Decide si, en esa clase, son independientes los sucesos “ser fumador” y “ser chica” si se sabe que fuman 14 chicas y 6 chicos. Justifica la respuesta. c) Se eligen al azar y con reemplazamiento 20 alumnos de la clase. Calcula la probabilidad de que 3 o más de ellos sean fumadores. d) De los cien alumnos, 50 han pasado alguna vez por tutorías de Bioestadística: 30 chicas, de las que 6 son fumadoras, y 20 chicos, de los que 2 son fumadores. El hecho de haber pasado por tutorías, ¿afecta a la relación entre los sucesos “ser fumador” y “ser chica” establecida en el apartado b)?; ¿y a la relación entre los sucesos “ser fumador” y “ser chico”? Justificar ambas respuestas.

I.27 Un ginecólogo con pocos escrúpulos afirma poseer un método, basado exclusivamente en la exploración visual del rostro de la embarazada, para detectar, por el módico precio de 50 euros, el sexo del bebé en la primera visita a su consulta; tanto es así que se declara dispuesto a devolver el dinero en caso de equivocación. En realidad, se juega a cara o cruz si decidirá niño o niña. Justifica matemáticamente por qué se le acusa de tener pocos escrúpulos.

SEGUNDA PARTE

Inferencia Estadística

Capítulo **3**

Introducción a la Inferencia
Estadística

3.1. Cálculo de Probabilidades e Inferencia Estadística

3.1.1. Distinción entre Probabilidad y Estadística: Parámetros de una distribución

Hemos comentado anteriormente que el cálculo de probabilidades, además del interés que tiene en sí mismo (también desde el punto de vista de la bioestadística), es una herramienta fundamental para la inferencia estadística y que la inferencia estadística es el objetivo último de este libro de bioestadística. Recíprocamente, de algún modo, la inferencia estadística, además de justificar la extrapolación de las conclusiones obtenidas a partir de una muestra al resto de la población de la que se extrajo, puede considerarse también como un paso previo para el cálculo de probabilidades. Intentaremos a continuación justificar esta afirmación, al mismo tiempo que pretendemos dejar clara la diferencia entre ambas disciplinas.

Formalmente, el cálculo de probabilidades comienza con un espacio de probabilidad (Ω, P) , de tal suerte que, dado un suceso $A \subset \Omega$, conocemos perfectamente su probabilidad $P(A)$; en la práctica, la dificultad en el cálculo de $P(A)$ puede estribar en la correcta identificación del suceso A con un determinado subconjunto de Ω , o en la habilidad de uso de las propiedades de la probabilidad. En ocasiones, el suceso A queda identificado como el conjunto de sucesos elementales para los que una cierta v.a. $X : \Omega \rightarrow \Omega'$ toma valores en un cierto subconjunto $A' \subset \Omega'$, en cuyo caso $P(A) = P(X \in A') = P^X(A')$ queda perfectamente determinada por la distribución P^X de la v.a. X ; si esta distribución es conocida (por ejemplo, binomial $B(n, p)$ de parámetro p , o normal $N(\mu, \sigma^2)$ de parámetros μ y σ^2), el cálculo de la probabilidad de A es sencillo, sin más que usar la función de probabilidad en el caso discreto o la función de densidad en el caso continuo.

Por ejemplo, la distribución binomial $B(n, p)$ de parámetro p (n suele ser determinado de antemano por el investigador y no lo consideraremos como parámetro) ha sido introducida como la distribución del número de caras en n lanzamientos independientes de una moneda si la probabilidad de sacar cara en un lanzamiento es p (o, más generalmente, la distribución del número de éxitos en n realizaciones independientes de un mismo experimento si la probabilidad de éxito en una realización del experimento es p y la de fracaso $1-p$). Conocida esa probabilidad $p \in]0, 1[$, podemos calcular la probabilidad de que una v.a. con distribución $B(n, p)$ tome valores en algún subconjunto de $\{0, 1, \dots, n\}$ haciendo uso de la función de probabilidad de esa distribución (en un ejemplo continuo como el de la normal, haremos uso de la función de densidad para calcular probabilidades); en la práctica, basta saber usar las tablas de la binomial y de la normal (véanse las tablas II y III al final de este manual).

Pero, en una situación real, el parámetro suele ser desconocido, y esto es lo que diferencia el cálculo de probabilidades de la inferencia estadística; en el caso de la moneda, ante una moneda concreta, nada nos asegura a priori cuál es la probabilidad p de sacar cara en un lanzamiento. El objetivo de la inferencia estadística es precisamente ése, obtener información sobre el parámetro desconocido tras una serie de realizaciones del experimento. Tras lanzar la moneda al aire n veces, anotamos el número k de caras obtenidas y, a partir de ese valor experimental y con el uso adecuado de la inferencia estadística, pretendemos obtener información sobre el parámetro desconocido p (la probabilidad desconocida p de sacar cara).

Si, tras la aplicación del método estadístico correspondiente, admitimos que la probabilidad de sacar cara es efectivamente $1/2$, ya podemos de nuevo calcular probabilidades haciendo uso de la distribución binomial.

En resumen, en inferencia estadística se suele suponer que los resultados del fenómeno aleatorio que nos interesa son valores de una cierta v.a. de cuya distribución podemos admitir que es de un tipo conocido (binomial, normal, . . .), pero se supone desconocida, en el sentido de que se desconocen los parámetros concretos (p, μ, σ^2, \dots). El término “parámetro” se suele utilizar en estadística para designar una cantidad obtenida a partir de la distribución de probabilidad de la variable que nos interesa estudiar en una población, como pueden ser la media o la varianza de la misma, y que, en ocasiones, determinan esa distribución,

como ocurre con el parámetro p de la distribución binomial, o los parámetros μ y σ^2 de la distribución normal. Puesto que habitualmente no tenemos acceso a toda la población, rara vez podemos calcular exactamente un parámetro; pero, a partir de una muestra de la población, y con la ayuda de un “estadístico” adecuado, es decir, de una función apropiada de los datos u observaciones, podemos estimar ese parámetro.

Cuando sólo se suponen desconocidos uno, dos o, en general, un número finito de parámetros de la distribución se suele hablar de inferencia paramétrica. Pero en otras ocasiones se supone desconocido, no sólo un número finito de parámetros de la distribución de la v.a. que se observa, sino que también se supone desconocido el tipo de la distribución de la v.a. (sólo se asume, por ejemplo, que esa distribución tiene densidad); se habla en ese caso de inferencia no paramétrica.

3.1.2. Estructura estadística. Estadísticos

Desde un punto de vista formal, la diferencia entre el cálculo de probabilidades y la inferencia estadística es que, mientras aquel comienza con un espacio de probabilidad (Ω, P) donde P es una probabilidad en Ω , la inferencia estadística comienza con lo que llamaremos estructura estadística (también llamada espacio estadístico, experimento estadístico o, simplemente, modelo estadístico).

DEFINICIÓN 3.1. (Estructura estadística) Una estructura estadística (o modelo estadístico) es un par (Ω, \mathcal{P}) , donde \mathcal{P} es una familia o conjunto de probabilidades en el conjunto Ω . Los elementos de Ω , en este contexto, se llamarán observaciones.

Observación 3.2. A diferencia del espacio de probabilidad, la segunda componente del par que define una estructura estadística, en lugar de una probabilidad en Ω , es una familia \mathcal{P} de probabilidades en Ω que se obtiene al considerar todos los posibles valores que pueda tomar el parámetro desconocido. Hacemos la suposición fundamental de que \mathcal{P} contiene la verdadera distribución de la variable que observamos. \triangleleft

Ejemplo 3.3. Por ejemplo, si la proporción p de individuos que poseen una cierta enfermedad en una población (a p se le llama también “prevalencia” de la enfermedad en esa población) es igual a 0.35, el espacio de probabilidad asociado al número de individuos enfermos entre 150 extraídos al azar (y con reemplazamiento) de esa población es $(\{0, 1, \dots, 149, 150\}, B(150, 0.35))$. Si la prevalencia p de la enfermedad es desconocida en

esa población, en lugar de ese espacio de probabilidad debemos considerar la estructura estadística $(\{0, 1, 2, \dots, 148, 149, 150\}, \{B(150, p) : p \in [0, 1]\})$. \triangleleft

Ejemplo 3.4. Si la distribución de la característica que nos interesa se considera normal $N(\mu, \sigma^2)$ de parámetros $\mu \in \mathbb{R}$ y $\sigma^2 > 0$ desconocidos, la estructura estadística vendría dada por $\Omega = \mathbb{R}$ y $\mathcal{P} = \{N(\mu, \sigma^2) : \mu \in \mathbb{R}, \sigma^2 > 0\}$. Por ejemplo, la estructura estadística apropiada para el nivel de colesterol de un español elegido al azar sería $(\mathbb{R}, \{N(\mu, \sigma^2) : \mu \in \mathbb{R}, \sigma^2 > 0\})$, pues el nivel de colesterol de un español elegido al azar es un número real que se considera como valor concreto tomado por una v.a. de cuya distribución en toda la población estamos suponiendo que es normal con media μ y varianza σ^2 desconocidos. El parámetro μ representa el nivel medio de colesterol en toda la población, mientras que la varianza desconocida σ^2 es una medida de la dispersión de los niveles de colesterol de los españoles alrededor de su media. \triangleleft

Observación 3.5. (★) Del mismo modo que habíamos introducido el espacio de probabilidad producto podemos introducir ahora la estructura estadística producto como la correspondiente a una muestra de tamaño n de esa distribución desconocida; por ejemplo, la estructura estadística producto $(\mathbb{R}^n, \{N(\mu, \sigma^2)^n : \mu \in \mathbb{R}, \sigma^2 > 0\})$ viene a ser el modelo apropiado para el estudio de los niveles de colesterol de n españoles elegidos al azar (y con reemplazamiento) de la población. Aunque estas son cuestiones que quedan lejos de las sencillas pretensiones de este libro. \triangleleft

DEFINICIÓN 3.6. (Estadístico) Un estadístico es una aplicación $X : (\Omega, \mathcal{P}) \rightarrow \Omega'$.

Observación 3.7. A veces se usa en inferencia estadística una terminología diferente a la del cálculo de probabilidades para referirnos a conceptos análogos. Por ejemplo, a los elementos de Ω les hemos llamado sucesos elementales en cálculo de probabilidades y se les suele llamar observaciones en inferencia estadística. En cálculo de probabilidades hemos llamado variable aleatoria a una aplicación $X : (\Omega, \mathcal{P}) \rightarrow \Omega'$, mientras que en inferencia estadística hablamos de estadístico. Así pues, un estadístico no es más que una función de las observaciones. Como ejemplos concretos de estadísticos, mencionemos aquí la media y varianza muestrales \bar{X} y S^2 de una muestra X_1, \dots, X_n de una distribución real desconocida definidas como sabemos por

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{y} \quad S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad \triangle$$

3.1.3. Los dos grandes problemas de la Inferencia Estadística

Así pues, la inferencia estadística comienza con una estructura estadística (Ω, \mathcal{P}) , donde Ω es el conjunto de las observaciones o resultados posibles del fenómeno aleatorio que estudiamos y cuya distribución suponemos desconocida; pero hacemos la hipótesis fundamental

de que es una de las probabilidades de la familia \mathcal{P} ; el objetivo de la inferencia estadística consiste en la obtención de información sobre ese parámetro desconocido haciendo uso del resultado concreto $\omega \in \Omega$ obtenido en el experimento con la intención de tomar una decisión, que puede ser de dos tipos: dar una estimación de ese parámetro, junto con una medida del error cometido en la estimación, o decidir entre una hipótesis que ha sido formulada sobre el parámetro y su contraria (por ejemplo, podemos estar interesados en decidir si la probabilidad desconocida p de sacar cara es o no igual a $1/2$). Eso nos conduce a los dos grandes problemas de la inferencia estadística: la estimación (en sus dos variantes: estimación puntual y estimación conjuntista) y el contraste de hipótesis. Conviene dejar claro que nunca debe tomarse como definitiva la decisión tomada a partir de la observación ω , pues, debido a la naturaleza aleatoria de las observaciones, en una segunda realización del experimento podemos obtener otra observación ω' que puede dar lugar a una decisión diferente. Así pues, en inferencia estadística tomamos decisiones a partir de observaciones aleatorias y nunca podremos tener la certeza absoluta de haber acertado en la decisión tomada. Siendo la verdadera distribución desconocida, ni siquiera tiene sentido preguntarse si en esa decisión hemos cometido o no un error; no obstante, sí podemos preguntarnos por la probabilidad de cometer un error, y la mayor parte de los métodos estadísticos que presentaremos en este texto han sido diseñados con el objetivo de que la probabilidad de cometer un error sea lo más pequeña posible. Notemos, en fin, que no sólo la naturaleza aleatoria de las observaciones nos invita a ser cautos en la interpretación de la decisión que finalmente tomamos sobre la distribución desconocida que observamos; hay que ser conscientes de que esa decisión depende también de las suposiciones que hemos hecho para asignar un modelo estadístico (es decir, una estructura estadística) al fenómeno aleatorio que estudiamos.

3.2. Estimación

3.2.1. Estimación puntual. Estimador

Dentro de la teoría de estimación podemos distinguir entre la estimación puntual, que tiene como objetivo proporcionar una estimación del parámetro desconocido, y la estimación conjuntista (también llamada, en el caso más sencillo, estimación por intervalos de confianza), cuyo objetivo es proporcionar una idea del error cometido en la estimación en un sentido a precisar posteriormente. En este tema haremos una primera introducción a la estimación puntual.

DEFINICIÓN 3.8. (Estimador y estimación) Si θ es un parámetro real desconocido de la distribución de la variable que observamos y Ω es el conjunto de los resultados posibles (u observaciones) del experimento, un estimador del parámetro θ es un estadístico $T : \Omega \rightarrow \mathbb{R}$. Para cada resultado posible $\omega \in \Omega$ del experimento, $T(\omega)$ es una estimación de θ .

Observación 3.9. Si en inferencia estadística se pretende obtener información sobre la distribución desconocida (por ejemplo, sobre un parámetro desconocido de esa distribución) que estamos observando a partir de las observaciones, en estimación puntual en concreto se pretende obtener una estimación sobre un parámetro desconocido de esa distribución, es decir, se pretende obtener a partir de las observaciones un valor “aproximado” de ese parámetro desconocido. Para ello buscamos un estimador del parámetro desconocido, es decir, una función de las observaciones que utilizaremos para obtener la estimación concreta del parámetro. Hemos escrito la palabra aproximado entre comillas pues queremos distinguir la idea de estimación de la de aproximación del siguiente modo: mientras que en una aproximación a un valor que nos interesa es posible encontrar una cota tan pequeña como queramos del error cometido en la aproximación (pensemos, p.ej., en el número π , que habitualmente aproximamos por 3.1416, aunque teóricamente es posible conseguir valores tan aproximados al número π como deseemos), en la estimación de un parámetro desconocido sólo podemos hablar de la probabilidad de haber cometido un error de cierta magnitud, debido a la naturaleza aleatoria de las observaciones en las que se apoya la estimación. \triangleleft

Observación 3.10. Nos limitaremos a presentar los estimadores usuales de ciertos parámetros que son especialmente interesantes, como pueden ser los del parámetro p de una binomial o los de la media y la varianza de una distribución normal. Quede claro de momento que, si

Ω es el conjunto de las observaciones del experimento aleatorio que estamos considerando (Ω puede ser \mathbb{R} si estamos interesados en una característica real de los individuos de una cierta población y sólo seleccionamos un individuo, o $\Omega = \mathbb{R}^n$ si seleccionamos una muestra de tamaño n de esa población, o $\Omega = \{0, 1, \dots, n\}$ si sólo estamos interesados en el número de veces que ocurre un cierto suceso en n repeticiones del experimento), y si $T : \Omega \rightarrow \mathbb{R}$ es el estimador que usaremos para el parámetro θ , entonces, una vez realizado el experimento y observado el valor ω , consideraremos $T(\omega)$ como estimación de θ ; si el experimento aleatorio se realiza en otra ocasión podemos obtener un resultado ω' diferente y la estimación $T(\omega')$ de θ sería distinta. \triangleleft

3.2.2. Estimación de los parámetros de una distribución normal y de una proporción

(a) ESTIMACIÓN DE LA MEDIA μ Y LA VARIANZA σ^2 DE UNA DISTRIBUCIÓN NORMAL:

Para estimar la media μ de una distribución normal $N(\mu, \sigma^2)$ de parámetros μ y σ^2 desconocidos a partir de una muestra X_1, \dots, X_n de tamaño n de esa distribución, utilizaremos como estimador la media muestral

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

Para estimar la varianza desconocida σ^2 , utilizaremos como estimador la varianza muestral

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

Observación 3.11. Por ejemplo, si asumimos que el nivel de colesterol se distribuye normalmente en una cierta población, aunque con media μ y varianza σ^2 desconocidas, seleccionados n individuos al azar de esa población y determinados sus niveles de colesterol x_1, \dots, x_n , la media aritmética $\bar{x} = (x_1 + \dots + x_n)/n$ es una estimación del nivel medio poblacional de colesterol μ y $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$ es una estimación de la varianza poblacional σ^2 . \triangleleft

Observación 3.12. (★) La media muestral es también el estimador que se suele utilizar para estimar la media desconocida de otras distribuciones distintas de la normal; incluso es, desde distintos puntos de vista, el estimador de la media preferido en inferencia no paramétrica donde sólo se asume de la distribución desconocida que tiene densidad y media finita. No obstante, para cierto tipo de distribuciones es preferible el uso de otros estimadores de la media μ diferentes de la media muestral. \triangleleft

(b) ESTIMACIÓN DE UNA PROPORCIÓN p (O DEL PARÁMETRO p DE UNA DISTRIBUCIÓN BINOMIAL): Supongamos que estamos interesados en decidir si los individuos de una población presentan o no una cierta característica (son rubios o no, son altos o no, . . . – suponiendo que, ante un individuo cualquiera de la población, es posible determinar sin ambigüedad si posee o no esa característica). La proporción p de individuos que presentan esa característica se supone desconocida y es el objetivo de nuestro estudio. Para ello seleccionamos una muestra de tamaño n de la población y determinamos el número $x \in \{0, 1, \dots, n\}$ de individuos de la muestra que poseen esa característica. x se considera como una observación particular de una v.a. X con distribución binomial $B(n, p)$ de parámetro desconocido p . x/n es la proporción de individuos de la muestra que poseen esa característica, y se utilizará como estimador de la proporción desconocida p . Así pues, el estimador del parámetro desconocido p de una distribución binomial $B(n, p)$ es la proporción muestral

$$\hat{p} = X/n$$

donde X es la variable que representa el número de individuos con esa característica presentes en la muestra de tamaño n .

Observación 3.13. En inferencia estadística existen diversos criterios, que no estudiaremos aquí en detalle, a la hora de elegir el mejor estimador de un parámetro desconocido. Sí nos interesa destacar que los estimadores que presentaremos en esta obra son buenos estimadores desde diferentes puntos de vista. Por ejemplo, los estimadores de los parámetros μ , σ^2 y p descritos anteriormente son estimadores insesgados de mínima varianza: sin entrar en detalles, que un estimador sea insesgado significa que su media coincide con el parámetro que se quiere estimar, con el fin de que las sobreestimaciones producidas por el estimador sean compensadas con las subestimaciones; y que sea insesgado de mínima varianza significa que es un estimador insesgado que posee menor varianza que la de cualquier otro estimador insesgado, lo que hace que las estimaciones, además de insesgadas, sean lo más precisas posibles. Una imagen física de lo que acabamos de decir puede ser la siguiente: imaginemos que se desea estimar el centro de una diana -que supondremos unidimensional, por simplicidad- mediante disparos con una escopeta de aire comprimido, que hace las veces de estimador; los diferentes disparos -estimaciones- se sitúan entorno a un valor medio, y éste será distinto del centro de la diana cuando la escopeta sea sesgada; entre todas las escopetas insesgadas, preferiremos aquella en la que la dispersión de los disparos ha sido menor. El $n - 1$ en el denominador de la definición de la varianza muestral, en lugar de n que podría parecer más natural, queda explicado por el hecho de que así la varianza muestral es un estimador insesgado de σ^2 . \triangleleft

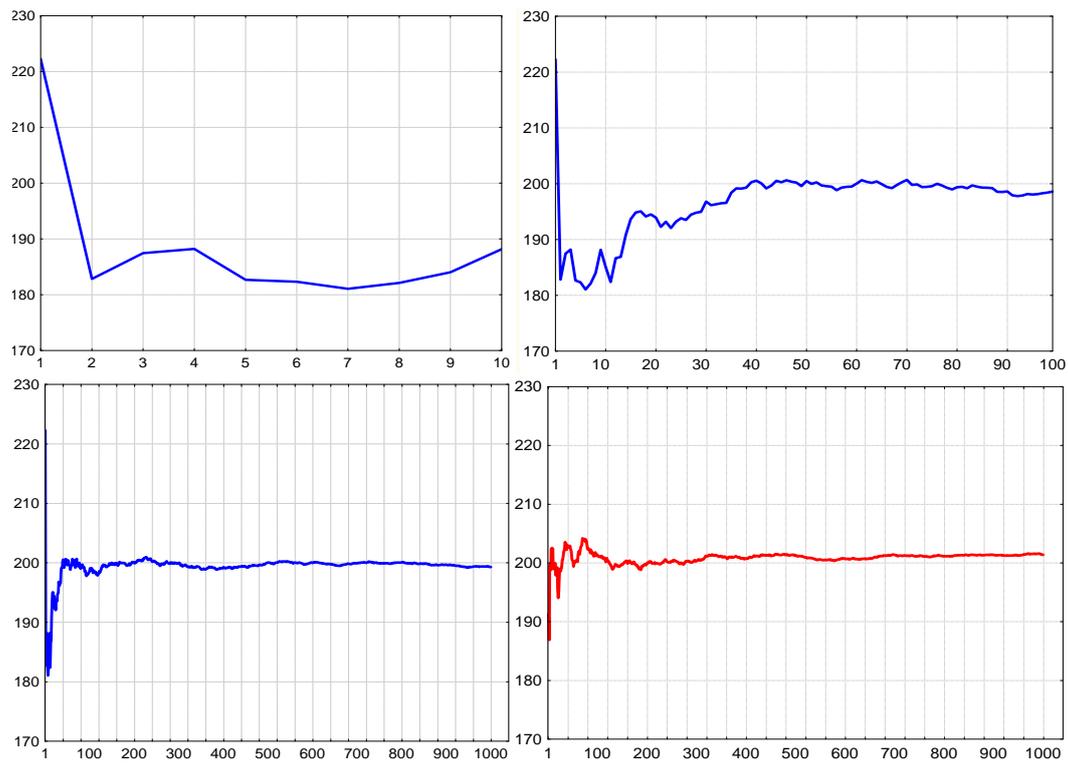


Figura 3.1: Convergencia de la media muestral a la media poblacional

Observación 3.14. (Ley fuerte de los grandes números) Los estimadores anteriores, y otros que veremos posteriormente, presentan además un buen comportamiento asintótico (es decir, estimadores que tienen propiedades deseables cuando el tamaño de la muestra se hace más y más grande); la Figura 3.1 recoge las medias muestrales de tamaños $1, 2, \dots, 1000$ obtenidas a partir de 1000 observaciones independientes de una distribución $N(200, 25^2)$ que, aunque han sido obtenidas por simulación, podrían ser consideradas como los niveles de colesterol de 1000 individuos elegidos al azar de una población grande en la que el nivel de colesterol tenga (aproximadamente) distribución $N(200, 25^2)$. Los tres primeros gráficos corresponden a una misma muestra de tamaño 1000 (presentando la evolución de la media muestral hasta los tamaños 10, 100 y 1000), mientras que el cuarto gráfico corresponde a una segunda muestra de tamaño 1000 de esa misma población. No es casualidad que, cuando el tamaño n de la muestra aumenta, la media muestral se aproxime a la media poblacional (200, en este caso), sino que es consecuencia de un teorema matemático que se conoce como ley fuerte de los grandes números, debida a Kolmogorov, y que, grosso modo, viene a decir que, con probabilidad 1 (que no es lo mismo que siempre, pero casi), cuando el tamaño de la muestra tiende a infinito, la media muestral tiende a la media poblacional. Es esa misma ley la que explica la estabilización de las frecuencias relativas (medias muestrales de una muestra de una distribución de Bernoulli) de un suceso entorno a su probabilidad (media de esa distribución de Bernoulli). \triangleleft

3.2.3. Estimación conjuntista. Intervalos de confianza

Los estimadores puntuales, una vez realizado el experimento aleatorio y observado un valor ω , nos dan una estimación (valor “aproximado”) del verdadero valor del parámetro a estimar. Pero una respuesta de este tipo no es del todo satisfactoria, pues no contiene referencia alguna al error cometido en la estimación. Ya hemos dicho anteriormente que no tiene sentido intentar buscar cotas exactas de ese error. Lo que sí tiene sentido es proporcionar, alrededor de la estimación del parámetro, un conjunto que contenga ese parámetro con probabilidad alta. Éste es el objetivo de la estimación conjuntista o estimación por intervalos.

Por ejemplo, si la característica que nos interesa de los individuos de una cierta población sigue una distribución normal $N(\mu, 625)$ de media desconocida μ y varianza conocida 625 (pensar en el ejemplo del colesterol en la población española considerado anteriormente pero ahora con varianza que se supone conocida e igual a 625), hemos propuesto como estimador de μ la media muestral \bar{X} de tal suerte que tras la realización del experimento (selección de

n españoles al azar y medida de los niveles de colesterol de cada uno de ellos) se obtienen n números reales x_1, \dots, x_n cuya media $\bar{x} = \frac{1}{n}(x_1 + \dots + x_n)$ (media muestral) es la estimación propuesta para la media poblacional μ . Dado un número real $c > 0$, la cuestión de si \bar{x} dista de μ menos que c no tiene sentido, por ser μ desconocida. No obstante, es posible hacer afirmaciones de tipo probabilístico sobre ese error; por ejemplo, podemos afirmar que con probabilidad 0.95 el verdadero y desconocido valor del parámetro μ se encuentra en un cierto intervalo alrededor de la media muestral. Ése es un intervalo cuyos extremos dependen de la observación obtenida en el experimento y que llamaremos intervalo de confianza para la media desconocida μ .

3.2.4. Intervalo de confianza para la media de una distribución normal de varianza conocida

Con la intención de fijar mejor los conceptos que vamos a introducir, nos interesamos en primer lugar por un problema sencillo: el de construir un intervalo de confianza para la media de una distribución normal de varianza conocida a partir de una muestra de tamaño n de esa distribución.

Consideremos el caso de una muestra de tamaño n de una distribución $N(\mu, \sigma_0^2)$ de media desconocida $\mu \in \mathbb{R}$ y varianza conocida $\sigma_0^2 > 0$ (suponer conocida la varianza mientras que la media se supone desconocida no es, en general, una situación realista, pero, desde un punto de vista pedagógico, es una situación lo suficientemente sencilla como para que las definiciones que presentaremos sean más fácilmente comprensibles; suponer también desconocida la varianza complica algo el problema de determinar un intervalo de confianza para μ).⁽¹⁾

Fijemos un valor $\mu \in \mathbb{R}$. Sabemos que la media muestral \bar{X} correspondiente a una muestra X_1, \dots, X_n de una distribución normal $N(\mu, \sigma_0^2)$ tiene distribución $N(\mu, \sigma_0^2/n)$.

¹ La estructura estadística correspondiente a una muestra de tamaño n de la distribución $N(\mu, \sigma_0^2)$ considerada es, entonces, $(\mathbb{R}^n, \{N(\mu, \sigma_0^2)^n : \mu \in \mathbb{R}\})$.

Tipificando se obtiene que

$$\frac{\bar{X} - \mu}{\sigma_0/\sqrt{n}} = \sqrt{n} \frac{\bar{X} - \mu}{\sigma_0}$$

tiene distribución $N(0, 1)$. Puesto que el intervalo $[-1.96, 1.96]$ concentra el 95 % de la distribución $N(0, 1)$, se deduce que la probabilidad de que $\sqrt{n} \frac{\bar{X} - \mu}{\sigma_0}$ esté en el intervalo $[-1.96, 1.96]$ es igual a 0.95. Pero decir que

$$-1.96 \leq \sqrt{n} \frac{\bar{X} - \mu}{\sigma_0} \leq 1.96$$

es equivalente a decir que

$$\bar{X} - 1.96 \frac{\sigma_0}{\sqrt{n}} \leq \mu \leq \bar{X} + 1.96 \frac{\sigma_0}{\sqrt{n}}$$

por tanto, el intervalo

$$\left[\bar{X} - 1.96 \frac{\sigma_0}{\sqrt{n}}, \bar{X} + 1.96 \frac{\sigma_0}{\sqrt{n}} \right]$$

contiene a la media μ con probabilidad 0.95; como eso es cierto cualquiera que sea μ (que lo hemos fijado desde un principio), podemos afirmar que eso también es cierto para el verdadero y desconocido valor μ del parámetro. Se dice por ello que ése es un intervalo de confianza para la media μ al nivel de confianza 0.95 (o al 95 % de confianza).

Algo más generalmente, haciendo uso de los cuantiles z_α de la distribución normal, dado un número $0 < \alpha < 1$ (interesa normalmente tomar valores de α pequeños; el valor $\alpha = 0.05$ es el que habitualmente se utiliza en la literatura científica, pero también suelen considerarse como valores de α los números 0.01 y 0.1 entre otros), el intervalo

$$\left[\bar{X} - z_\alpha \frac{\sigma_0}{\sqrt{n}}, \bar{X} + z_\alpha \frac{\sigma_0}{\sqrt{n}} \right]$$

es un intervalo de confianza para la media desconocida μ al nivel de confianza $1 - \alpha$.

Observación 3.15. Nótese que, más que de un intervalo, se trata de una familia de intervalos: un intervalo para cada muestra x_1, \dots, x_n ; el intervalo puede ser diferente de haberse extraído otra muestra. La afirmación de que ése es un intervalo de confianza para la media (pensar, por ejemplo, en el nivel medio de colesterol) desconocida μ al nivel de confianza 0.95 debe entenderse del siguiente modo: sea cual sea la media desconocida μ y suponiendo que el experimento de extraer una muestra de tamaño n se realiza un número grande de veces, el 95 % de las ocasiones que extraigamos una muestra de tamaño n de la población obtendremos niveles de colesterol x_1, \dots, x_n que dan lugar a un intervalo $[\bar{x} - 1.96\sigma_0/\sqrt{n}, \bar{x} + 1.96\sigma_0/\sqrt{n}]$

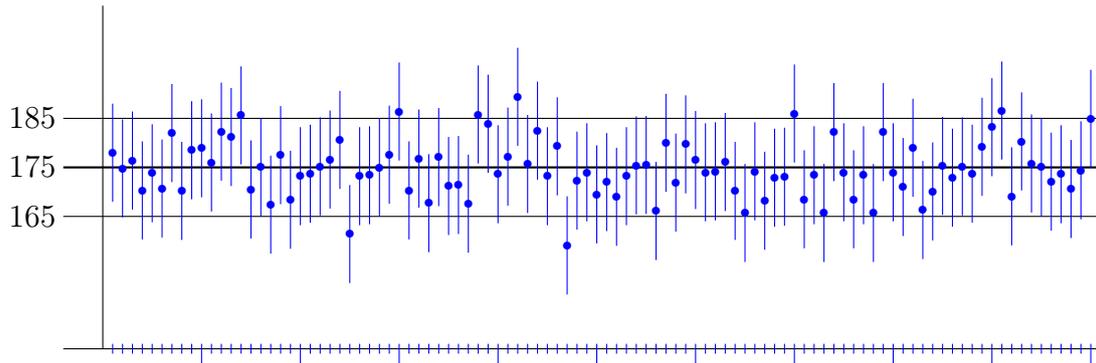


Figura 3.2: Interpretación empírica simulada del concepto de intervalo de confianza

que contiene al parámetro μ ; sólo un 5% de las veces que extraigamos una muestra de tamaño n , la media μ puede quedar fuera del intervalo. Así que, a menos que tengamos la mala suerte de haber obtenido una de ese 5% de muestras conflictivas, obtendremos un intervalo que contiene a la media desconocida, lo cual nos da una buena información sobre esta: concretando algo más, supongamos que sabemos que, en el caso del colesterol, $\sigma_0^2 = 625$ y que $n = 25$ y que, tras la determinación de los niveles de colesterol de 25 españoles elegidos al azar, se ha obtenido el valor medio muestral $\bar{x} = 178.03$; entonces, el intervalo concreto al 95% de confianza obtenido viene a ser $[178.03 - 1.96 \cdot 5, 178.03 + 1.96 \cdot 5]$ y nos queda una confianza del 95% en que la media desconocida (el nivel medio de colesterol en España) está entre los valores 168.23 y 187.83, que no es poca información, y nos da (con probabilidad alta) una medida del error que estamos cometiendo al estimar el nivel medio de colesterol poblacional μ mediante el nivel medio muestral $\bar{x} = 178.03$. No tendría, sin embargo, sentido afirmar que el parámetro desconocido μ está en el intervalo concreto $[178.03 - 9.8, 178.03 + 9.8]$ con probabilidad 0.95, y por eso decimos que tenemos una confianza del 95% de que eso es así. Para confirmar los comentarios precedentes, se ha realizado un ejercicio de simulación (con la ayuda de un programa estadístico) que ha consistido en la extracción de 100 muestras de tamaño 25 de una distribución $N(175, 625)$.

La Figura 3.2 contiene los 100 intervalos de confianza correspondientes, donde se puede comprobar experimentalmente que un alto porcentaje de esos intervalos contiene a la verdadera media 175; si el número de realizaciones del experimento, en lugar de 100 como en este caso, tiende a ∞ , el porcentaje de intervalos que contiene a la verdadera media 175 tiende al 95%, de acuerdo con la interpretación frecuentista de la probabilidad. \triangleleft

Observación 3.16. (★) En general, podemos dar la siguiente definición de intervalo de confianza: Consideremos una estructura estadística (Ω, \mathcal{P}) en la que la familia de probabilidades queda descrita con la ayuda de un parámetro θ , es decir, $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$, y que queremos estimar una función real del parámetro $f : \Theta \rightarrow \mathbb{R}$. Un intervalo de confianza para f al nivel de confianza $1 - \alpha$ es una familia de intervalos $\{[a(\omega), b(\omega)] : \omega \in \Omega\}$ que, para cada $\theta \in \Theta$, verifica lo siguiente

$$P_\theta(\{\omega \in \Omega : a(\omega) \leq f(\theta) \leq b(\omega)\}) = 1 - \alpha$$

En el caso considerado en este tema, $\Omega = \mathbb{R}^n$, $\omega = (x_1, \dots, x_n)$, el parámetro θ es la media desconocida μ , el espacio de parámetros es $\Theta = \mathbb{R}$ y $f(\mu) = \mu$. \triangleleft

3.2.5. Determinación del tamaño de muestra necesario para conseguir una cierta precisión

Continuando con el ejemplo de la sección anterior, es claro que la información que proporciona el intervalo es tanto más precisa cuanto más pequeña es su longitud, que, en este caso, es

$$2z_\alpha \frac{\sigma_0}{\sqrt{n}}$$

A la vista de esta expresión, queda claro que esa longitud depende de tres cantidades: α , σ_0^2 y n . Fijando dos de esas cantidades y variando la tercera se obtienen las siguientes conclusiones: 1) si deseamos una mayor confianza, tendremos que tomar valores de α más pequeños, lo que da lugar a mayores valores de z_α y mayor longitud del intervalo, con lo que una ganancia en la confianza supone una pérdida en la precisión de la estimación, hasta el punto de que, para conseguir un nivel de confianza del 100 %, necesitamos un intervalo de longitud infinita; 2) la varianza σ_0^2 se ha considerado conocida, pero es evidente que una varianza mayor hubiera dado lugar a un intervalo de mayor longitud, con lo que un aumento en la variabilidad de los datos tiene como consecuencia una menor precisión en la estimación; 3) por último, parece lógico exigir en inferencia estadística que, al aumentar el tamaño de la muestra, se consiga una mejor información sobre la distribución desconocida; en el caso del intervalo de confianza eso es evidente, pues un mayor valor de n tiene como consecuencia una disminución de la longitud del intervalo (y, por tanto, un aumento de la precisión en la estimación).

En el caso que estamos considerando (muestra de tamaño n de una distribución normal de varianza conocida) hemos podido observar que la longitud del intervalo de confianza propuesto no depende de la observación concreta (x_1, \dots, x_n) obtenida. Eso es algo que no suele ocurrir en otras situaciones y que ahora nos permite plantear sencillamente un

problema interesante siempre en inferencia estadística: el de la determinación del tamaño de muestra n necesario para que la información proporcionada por la muestra sobre el parámetro desconocido sea suficientemente buena. En nuestro caso, ese problema se plantea del siguiente modo: ¿cuál es el tamaño n de muestra necesario para poder garantizar con una confianza del $100 \cdot (1 - \alpha) \%$ que la estimación \bar{x} no dista de la media desconocida μ más de una cierta cantidad $d > 0$ fijada de antemano? Se trata evidentemente de determinar n de modo que

$$z_\alpha \frac{\sigma_0}{\sqrt{n}} \leq d$$

lo cual se consigue sin más que tomar

$$n \geq \left(\frac{z_\alpha \sigma_0}{d} \right)^2$$

Puesto que el coste de la investigación suele aumentar con el tamaño de la muestra n , conviene tomar el menor valor de n que satisface la desigualdad anterior. En la bibliografía pueden encontrarse soluciones parciales al problema del cálculo del tamaño de muestra para otras situaciones diferentes a la que hemos considerado aquí, aunque nosotros no estudiaremos esos casos.

3.2.6. Cotas superior e inferior de confianza

Hasta el momento sólo hemos considerado intervalos de confianza bilaterales en los que nos interesa tanto una cota inferior como una cota superior para el verdadero valor del parámetro desconocido. Eventualmente podemos estar interesados en determinar solamente una cota superior de confianza al nivel $1 - \alpha$. En el mismo caso de una muestra de tamaño n de una distribución normal de media desconocida μ y varianza conocida σ_0^2 que estamos considerando, ya sabemos que, fijado un valor de μ , la media muestral \bar{X} tiene distribución $N(\mu, \sigma_0^2/n)$ y, por tanto,

$$\sqrt{n} \frac{\bar{X} - \mu}{\sigma_0}$$

tiene distribución $N(0, 1)$; así pues, dado $0 < \alpha < 1$, la probabilidad de que

$$\sqrt{n} \frac{\bar{X} - \mu}{\sigma_0} \geq -z_{2\alpha}$$

es $1 - \alpha$. Entonces

$$\left[-\infty, \bar{X} + z_{2\alpha} \frac{\sigma_0}{\sqrt{n}} \right]$$

es un intervalo de confianza unilateral para μ al nivel $1 - \alpha$.

Del mismo modo, si deseamos una cota inferior de confianza, se tiene que

$$\left[\bar{X} - z_{2\alpha} \frac{\sigma_0}{\sqrt{n}}, +\infty \right]$$

es un intervalo de confianza unilateral para μ al nivel $1 - \alpha$.

Observación 3.17. (Extensión al caso no normal) Los intervalos de confianza que hemos obtenido anteriormente (no sólo en esta sección) se basan en la hipótesis de que estamos observando una muestra de tamaño n de una distribución normal; consecuencia de ello es que la media muestral \bar{X} tiene también distribución normal, lo que nos ha conducido a los intervalos de confianza que hemos obtenido. Nos preguntamos ahora hasta qué punto los intervalos precedentes son aun válidos en el caso no normal. Aún cuando falle la hipótesis de normalidad de las observaciones, el teorema del límite central puede garantizarnos todavía la normalidad aproximada de la media muestral \bar{X} cuando n es grande (a efectos prácticos, la aproximación la venimos considerando aceptable si $n \geq 60$); los intervalos precedentes proporcionan entonces un nivel de confianza aproximado de $1 - \alpha$. \triangleleft

3.3. Contraste de Hipótesis

3.3.1. Introducción

En el tema anterior se ha estudiado el primer gran problema de la inferencia estadística: estimación. Este capítulo está dedicado al otro gran problema de la inferencia estadística: el contraste de hipótesis. Ambos problemas pretenden tomar una decisión sobre una población a partir de una muestra extraída de la misma. En estimación se pretende tomar una decisión sobre el valor de un parámetro poblacional a partir de lo observado en la muestra. En contraste de hipótesis se hace una afirmación (es decir, se formula una hipótesis) sobre algún parámetro poblacional y se pretende tomar una de dos decisiones posibles a partir de lo observado en la muestra: aceptar o rechazar esa hipótesis. Un análisis minucioso del experimento aleatorio puede conducir al investigador a la formulación de una hipótesis cuya validez desea contrastar haciendo uso de los datos. Así, por ejemplo, un médico puede estar interesado en comparar un nuevo tratamiento para una enfermedad diseñado por él con el tratamiento tradicional; o en decidir si el nivel de colesterol medio en una cierta población es igual a 175 o no; o si la proporción de individuos que posee una cierta enfermedad es 0.1 o no. Para tomar la decisión de aceptar o rechazar la hipótesis formulada, el investigador realizará lo que se llama un test de hipótesis.

DEFINICIÓN 3.18. (Hipótesis) Una hipótesis puede definirse como una proposición sobre la distribución, desconocida desde el punto de vista estadístico, de la variable que observamos, normalmente relacionada con uno o más parámetros de la misma. En un problema de contraste de hipótesis hay dos hipótesis claramente definidas: la hipótesis formulada por el investigador, que recibe el nombre de hipótesis nula y se suele denotar por H_0 , y la negación de la hipótesis nula, que llamaremos hipótesis alternativa y se denota por H_1 .

DEFINICIÓN 3.19. (Test de hipótesis) Un test para contrastar la hipótesis nula H_0 contra la hipótesis alternativa H_1 es un estadístico $\phi : \Omega \rightarrow \{0, 1\}$. Si $\omega \in \Omega$ es el resultado

del experimento y $\phi(\omega) = 0$ aceptaremos la hipótesis nula H_0 , y si $\phi(\omega) = 1$ aceptaremos la hipótesis alternativa H_1 , lo que equivale a rechazar la hipótesis nula. Al conjunto $\{\omega \in \Omega/\phi(\omega) = 1\}$ se le llama región crítica del test ϕ , y al conjunto $\{\omega \in \Omega/\phi(\omega) = 0\}$ se le llama región de aceptación del test ϕ .

Observación 3.20. Un test de hipótesis puede describirse como un procedimiento estadístico que nos permite tomar la decisión de aceptar o rechazar la hipótesis formulada a partir de las observaciones. En realidad, un test, más que en forma de estadístico (dependiente, por tanto, únicamente de los datos u observaciones), suele presentarse por medio de su región crítica (o su región de aceptación), de tal suerte que, una vez realizado el experimento y observado el valor ω , sólo tenemos que comprobar si ω cae o no en esa región para saber qué decisión debemos tomar. En el epígrafe siguiente se describe en varias etapas el proceso que seguiremos a la hora de resolver un problema de contraste de hipótesis. Debemos señalar aquí que el contraste de hipótesis no conduce a una demostración de la hipótesis formulada; simplemente, indica cuándo la hipótesis es o no compatible con los datos u observaciones obtenidos en el experimento. \triangleleft

3.3.2. Cuatro etapas en la ejecución de un test de hipótesis

El procedimiento de ejecución de un test de hipótesis se dividirá en cuatro etapas. Para fijar las ideas, consideraremos un problema especialmente sencillo: el contraste de la media de una distribución normal de varianza conocida a partir de una muestra de tamaño n de esa distribución.

Supongamos que estamos interesados en la siguiente cuestión: ¿Puede admitirse que el nivel medio desconocido μ de glucosa en suero en una población (una ciudad, un país,...) es igual a 90? Se ha formulado la hipótesis de que el nivel medio μ de glucosa es igual a 90 y se desea contrastarla. El problema puede quedar motivado por diferentes hechos: supongamos conocido, por ejemplo, que hace 20 años, en un estudio similar en esa población, se llegó a la conclusión de que el nivel medio de glucosa es igual a 90, y se desea saber si en la actualidad se puede sostener esa afirmación; o bien, puede ocurrir que tengamos conocimiento de estudios en otra población afirmando que el nivel medio de glucosa es igual a 90 y deseamos estudiar si eso también es así en esta población.

Aprovechemos este ejemplo concreto para describir los cuatro pasos del procedimiento:

1) ESTRUCTURA ESTADÍSTICA: Antes de realizar un contraste de hipótesis, fijaremos un modelo estadístico (una estructura estadística) para el experimento aleatorio que estudiemos. Para ello debemos comprender la naturaleza de los datos obtenidos en el experimento y hacer un análisis riguroso de las condiciones en que éste se realiza, pues esa información será necesaria a la hora de elegir el test que vamos a utilizar. Pensaremos normalmente que las observaciones obtenidas son valores posibles de una variable aleatoria y debe, por ejemplo, distinguirse si es una variable aleatoria discreta (por ejemplo, si cuenta el número de individuos que poseen una cierta característica dicotómica entre n elegidos al azar) o continua (por ejemplo, si mide una cierta cantidad, como puede ser el nivel de colesterol). En inferencia estadística diferentes suposiciones en el modelo teórico (estructura estadística) conducen a diferentes soluciones (por ejemplo, distintos intervalos de confianza en la estimación por intervalos). En particular, en la teoría de contraste de hipótesis ocurre lo mismo: un cambio en las suposiciones del modelo dan lugar a diferentes tests. Suposiciones habituales son, por ejemplo, las de independencia de las observaciones, normalidad de la distribución de la variable que observamos, etc.

Supongamos que, en el ejemplo del nivel de glucosa, se han realizado 10 mediciones a otros tantos individuos elegidos al azar de la población de interés, siendo la media muestral $\bar{x} = 84.02$. Esos diez datos se considerarán una realización concreta de una muestra de tamaño 10 de una distribución continua (la v.a. que observamos mide el nivel de glucosa), que supondremos normal con varianza conocida 100. La hipótesis de normalidad puede quedar justificada por estudios previos a gran escala, en que el teorema del límite central permita suponer normalidad para la distribución del nivel de glucosa; existen métodos estadísticos (en este libro veremos alguno de ellos) que permiten comprobar la bondad de ajuste a una distribución normal de los datos obtenidos. Sin embargo, no parece muy razonable suponer desconocida la media y conocida la varianza. Posteriormente veremos cómo tratar el caso de varianza desconocida; de momento son razones de tipo pedagógico las que justifican la suposición de varianza conocida (igual a 100, en este caso).

2) HIPÓTESIS: Hay dos hipótesis involucradas en un problema de contraste de hipótesis que deben ser explícitamente establecidas. La primera es la hipótesis que desea contrastarse;

recibe el nombre de hipótesis nula y suele denotarse por H_0 . La otra hipótesis es a menudo, aunque no siempre, la negación de la hipótesis nula; se conoce como hipótesis alternativa y se denotará por H_1 . Al final del procedimiento tendremos que decidir si aceptamos la hipótesis nula o la rechazamos (lo que supondrá la aceptación de H_1). Pero la aceptación de una u otra hipótesis se interpretará de distinta manera. Si la hipótesis nula es aceptada diremos que los datos en los que se ha basado el test no contienen evidencias suficientes para rechazar esa hipótesis; en cambio, si se rechaza la hipótesis nula diremos que los datos no son compatibles con la hipótesis nula. En general, una decisión por H_1 es más fiable que una decisión por H_0 . Las hipótesis para el caso del nivel de glucosa las escribiremos en forma compacta como sigue:

$$H_0 : \mu = 90 \quad , \quad H_1 : \mu \neq 90$$

3) ESTADÍSTICO DE CONTRASTE. SU DISTRIBUCIÓN BAJO LA HIPÓTESIS NULA: El estadístico de contraste es, como cualquier estadístico, una función de las observaciones (datos o muestra) obtenidas; su valor depende de la muestra (otra realización del experimento da lugar a datos diferentes y a un valor distinto para el estadístico de contraste) y la decisión de aceptar o rechazar la hipótesis nula se basará en el valor del estadístico de contraste para los datos obtenidos. Ese valor se llamará valor o cantidad experimental. En la etapa siguiente necesitaremos conocer la distribución del estadístico de contraste cuando la hipótesis nula se supone cierta. En el ejemplo que venimos estudiando, puesto que estamos contrastando una hipótesis sobre la media de una distribución normal con varianza conocida y que la media muestral es un estimador del parámetro desconocido μ , consideraremos el siguiente estadístico de contraste:

$$Z = \frac{\bar{X} - 90}{\sqrt{100}/\sqrt{10}}$$

que mide la discrepancia que existe entre la estimación \bar{X} de μ y el valor que propone para μ la hipótesis nula. El valor experimental es $z_{\text{exp}} = \frac{84.02-90}{\sqrt{100}/\sqrt{10}} = -1.89$. La distribución del estadístico de contraste Z , si la hipótesis nula se supone cierta, es normal de media 0 y varianza 1.

4) REGLA DE DECISIÓN: Para tomar esa decisión, el valor del estadístico de contraste se compara con otro valor o valores, que nos proporcionan ciertas tablas estadísticas (o ciertos programas estadísticos) como las que presentamos al final de este libro, y que dependen de la distribución del estadístico de contraste supuesta cierta la hipótesis nula y de lo que llamaremos el nivel de significación que deseamos para el test. El valor o valores proporcionados por la tabla dividen el conjunto de valores posibles del estadístico de contraste en dos regiones: la región de aceptación y la región de rechazo o región crítica; si el valor del estadístico de contraste cae en la primera se aceptará la hipótesis nula y si cae en la región crítica se rechazará la hipótesis nula. El nivel de significación de un test suele denotarse por la letra griega α y es un número elegido por el investigador antes de realizar el experimento; el valor más frecuentemente utilizado para α es 0.05, aunque también suelen usarse otros valores como 0.01 y 0.1. Al tomar la decisión podemos cometer dos tipos de error, que nos interesa distinguir: rechazar la hipótesis nula cuando es cierta (error que suele llamarse de tipo I) o aceptar la hipótesis nula cuando es falsa (error que suele llamarse de tipo II y se denotará por β ; $1 - \beta$ es la llamada potencia del test).

	H_0 cierta	H_1 cierta
Aceptar H_0	Decisión correcta ($1 - \alpha$)	Error de Tipo II (β)
Aceptar H_1	Error de Tipo I (α)	Decisión correcta ($1 - \beta$)

La región crítica o de rechazo del test se elige de modo que la probabilidad de cometer un error de tipo I sea igual que el nivel de significación α . Como α es pequeño, la probabilidad de rechazar la hipótesis cuando es cierta es pequeña; por tanto, cuando la decisión es rechazar, pueden darse dos casos: o bien la decisión es correcta, o bien la decisión es errónea y hemos cometido un error de tipo I, cosa que ocurre con probabilidad pequeña (en el caso $\alpha = 0.05$, una de cada 20 veces rechazaremos por error una hipótesis nula que es cierta); por eso las decisiones por H_1 se consideran fiables, como hemos indicado anteriormente.

En cambio, la probabilidad β de cometer un error de tipo II (de aceptar H_0 incorrectamente), que no es controlada por el investigador, porque no se pueden controlar simultáneamente las probabilidades de los dos tipos de error, puede ser relativamente grande, aunque suele disminuir aumentando el tamaño de la muestra según veremos más adelante. Así pues,

si la decisión es aceptar H_0 , también se pueden dar dos situaciones: o bien la decisión es correcta, o bien la decisión es errónea y se ha cometido un error de tipo II, cuya probabilidad, a tamaño de muestra fijo, no está controlada y podría ser eventualmente grande, lo que nos lleva a aceptar la hipótesis nula con cierta cautela, afirmando que los datos obtenidos no nos permiten rechazar la hipótesis nula.

Una vez obtenido el valor experimental y obtenida la región crítica del test, se decidirá aceptar la hipótesis nula H_0 (con reservas) si ese valor experimental no está en la región crítica (es decir, si está en la región de aceptación); en ese caso se dice que el test no es significativo al nivel α . Se decidirá rechazar la hipótesis nula H_0 (sin reservas), o aceptar H_1 , si el valor experimental está en la región crítica; en ese caso se dice también que el test es significativo al nivel α . En la sección 10.4 se presenta un ejemplo concreto para distinguir las consecuencias de los dos tipos de error considerados.

En el estudio del nivel de glucosa que usamos como ejemplo, fijaremos $\alpha = 0.05$. Necesitamos también determinar la región crítica. Para ello observemos la definición del estadístico de contraste. Puesto que la media muestral \bar{X} es una estimación de la media desconocida μ , valores demasiado pequeños o demasiado grandes del estadístico de contraste Z deberían hacernos pensar que la hipótesis nula no es cierta, pues si 90 fuese la verdadera media poblacional, la media muestral $\bar{x} = 84.02$ debería estar próxima a ése valor 90. Es decir, valores de Z alejados de 0 nos obligarán a rechazar la hipótesis. ¿Cuán lejos de cero debe estar el valor de Z (valor experimental) para que la decisión sea rechazar la hipótesis?: eso es algo que, como veremos, depende del nivel de significación elegido. De momento, lo dicho nos lleva a considerar una región crítica (o de rechazo) para este problema de la forma

$$\frac{|\bar{X} - 90|}{\sqrt{100}/\sqrt{10}} > C$$

donde C es una cierta constante que determinaremos en función del nivel de significación α deseado. Puesto que hemos elegido $\alpha = 0.05$, se pretende que la región de rechazo tenga probabilidad 0.05 supuesto que la hipótesis nula es cierta. Podemos preguntarnos cuál es la probabilidad de tomar una muestra de tamaño 10 de una distribución $N(90, 100)$ y obtener una media muestral tan distinta de 90 como es $\bar{x} = 84.02$. Si esa probabilidad es pequeña –entiéndase, inferior a 0.05– rechazaremos la hipótesis nula $H_0 : \mu = 90$. Puesto que la región

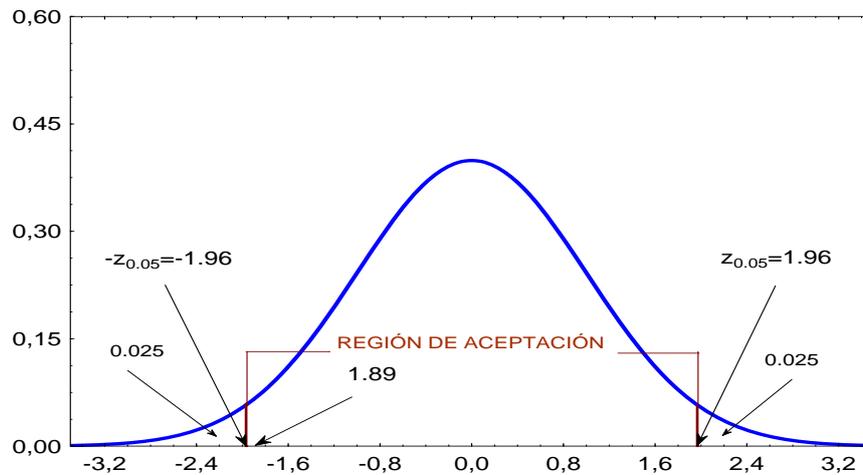


Figura 3.3: Región de aceptación del test

de rechazo tiene dos colas (es decir, tanto valores grandes como pequeños del estadístico de contraste Z nos sugieren rechazar la hipótesis nula), exigiremos que cada una de esas colas tenga probabilidad $\alpha/2 = 0.025$; las tablas de la distribución normal de media 0 y varianza 1 (tabla III.1), que es la distribución del estadístico de contraste Z si la hipótesis nula es cierta, nos proporcionan el valor $C = z_{\alpha} = 1.96$ como el número a la derecha del cual queda un área igual a $\alpha/2 = 0.025$ bajo la curva de esa distribución normal; por simetría, el área bajo la curva a la izquierda de $-z_{\alpha} = -1.96$ es de 0.025. Así pues, la región de aceptación (ver Figura 3.3) es el intervalo $[-1.96, 1.96]$ y la región de rechazo su complementario. Como el valor experimental $z_{\text{exp}} = -1.89$ (es decir, el valor del estadístico de contraste Z para la muestra obtenida) queda dentro de la región de aceptación, concluimos que los datos obtenidos no presentan evidencias en contra de la hipótesis nula, lo que nos lleva a aceptar (siempre con las debidas reservas) H_0 . Se dice también que el test no es significativo al nivel 0.05.

Observación 3.21. Los temas siguientes están esencialmente dedicados a la resolución de ciertos problemas particulares de inferencia estadística, en general, y de contraste de hipótesis, en particular. ◁

Observación 3.22. Hemos notado anteriormente de pasada que, fijado el tamaño n de la muestra, no es posible controlar los dos tipos de error simultáneamente. Como hemos indicado, α se fija siempre pequeño y el error de tipo I queda controlado por ese valor.

Otra posibilidad que no consideraremos aquí, salvo en un caso muy simple, consiste en fijar números α y β pequeños e intentar después determinar el tamaño de muestra n apropiado para que las probabilidades de error de tipo I y tipo II queden acotadas por esos valores, respectivamente. El deseo de hacer fiable también la decisión por H_0 puede tener como consecuencia un tamaño de muestra excesivamente grande, lo que aumentaría el coste de la experimentación. \triangleleft

Observación 3.23. En general, si μ_0 es un número real dado, para el problema de contrastar la hipótesis nula $H_0 : \mu = \mu_0$ contra $H_1 : \mu \neq \mu_0$ al nivel de significación α a partir de una muestra de tamaño n de una distribución normal $N(\mu, \sigma_0^2)$ de media desconocida μ y varianza conocida σ_0^2 , utilizaremos la región crítica

$$\frac{|\bar{X} - \mu_0|}{\sigma_0/\sqrt{n}} > z_\alpha$$

\triangle

3.3.3. Valor P o nivel mínimo de significación de un test

Existe otra manera de presentar los resultados de un test de hipótesis, que explicamos a continuación. Hemos visto que la constante $C = z_\alpha$ que determina la región crítica del test (llamada, por ello, valor crítico o valor teórico del test) depende del nivel de significación α elegido. Si α disminuye, las colas de la figura disminuyen, $C = z_\alpha$ aumenta, la región de aceptación crece y el test seguirá siendo no significativo. Por el contrario, si α aumenta, $C = z_\alpha$ disminuye, la región de aceptación se estrecha y el valor experimental obtenido llega a quedar fuera de la región de aceptación, con lo que el test pasaría a ser significativo. Esa reflexión nos lleva a la siguiente definición.

DEFINICIÓN 3.24. (Nivel mínimo de significación, probabilidad de significación o valor P) Llamaremos nivel mínimo de significación P , probabilidad de significación P o valor P al más pequeño valor de α para el cual se rechaza la hipótesis nula con los datos obtenidos.

Observación 3.25. Debe notarse que la probabilidad de significación P depende del resultado (muestra) obtenido en el experimento y debe entenderse como una medida de la disconformidad de los datos con la hipótesis nula. El valor P para un test de hipótesis es la probabilidad de obtener, cuando H_0 es cierta, un valor del estadístico de contraste tan extremo o más que el actualmente obtenido. Añadir a un trabajo de investigación el valor P es más informativo que decir simplemente si la hipótesis nula es o no rechazada al nivel de significación 0.05. En el ejemplo anterior la probabilidad de significación P para el valor

experimental $z_{\text{exp}} = -1.89$ obtenido se obtiene calculando la probabilidad que queda fuera del intervalo $[-1.89, 1.89]$ y es igual a 0.0588 (valor que se obtiene de las tablas de una distribución normal de media 0 y varianza 1, que es la distribución del estadístico de contraste Z si la hipótesis nula es cierta). \triangleleft

Observación 3.26. Conocido el nivel mínimo de significación P y dado el nivel de significación deseado α se decidirá por H_0 si $\alpha < P$ y por H_1 si $\alpha \geq P$. En ocasiones se gradúa la posible significación de un test del siguiente modo: si $0.01 \leq P \leq 0.05$ el test se dice significativo, si $0.001 \leq P < 0.01$ el test se dice muy significativo y si $P < 0.001$ el test se dice altamente significativo; si $P > 0.05$ el test es no significativo (aunque, a veces, para valores P entre 0.05 y 0.1 se hace notar una tendencia a la significación). \triangleleft

3.3.4. Tests unilaterales

El test considerado en el ejemplo anterior es un test bilateral o test de dos colas debido a que hay alternativas a ambos lados de la hipótesis nula;⁽²⁾ ello tiene como consecuencia, en este caso, que la región crítica ocupa las dos colas de la distribución del estadístico de contraste. Un test de hipótesis puede ser también unilateral o de una cola en el caso de que todas las alternativas se sitúen a un mismo lado de la hipótesis nula; en el caso considerado, la región crítica ocupa sólo una cola de la distribución del estadístico de contraste. La naturaleza de la cuestión planteada por el investigador determinará si se usa el test bilateral o el unilateral. Si la hipótesis nula es rechazada tanto por grandes como por pequeños valores del estadístico de contraste, el test es bilateral. Si sólo grandes o sólo pequeños valores del estadístico de contraste rechazan la hipótesis, el test es unilateral. El ejemplo siguiente pretende aclarar cuál debe ser la hipótesis nula en el caso unilateral.

Ejemplo 3.27. En el caso unilateral, se suele fijar como alternativa aquella hipótesis sobre la que queremos fiabilidad en caso de ser aceptada. Por ejemplo, si μ representa el número medio de microorganismos por cm^3 en el agua para beber y si 90 es el número a partir del cual el agua es declarada no potable, interesa considerar como hipótesis alternativa $H_1 : \mu < 90$. Con ello nos aseguramos de que sea pequeña la probabilidad de cometer un error -que será de tipo I- al declarar potable el agua; un error de tipo I de declarar potable un agua que no lo es puede tener consecuencias dramáticas y por eso conviene asegurarse

² De cualquier valor del parámetro que se encuentre en la hipótesis alternativa H_1 se dice que es una alternativa a la hipótesis nula.

de que la probabilidad de cometerlo es pequeña. Sin embargo, si la decisión es aceptar la hipótesis nula de que el agua no es potable, la probabilidad de cometer un error -que en este caso será de tipo II- no está controlada; pero las consecuencias de un error de este tipo quedan limitadas al rechazo de un agua que bien pudiera haber sido declarada potable. Conviene notar en este punto que los tests que aquí se describirán suelen proporcionar una probabilidad de error de tipo II que, a tamaño de muestra fijo, es lo más pequeña posible y que disminuye a medida que el tamaño de la muestra tiende a ∞ , según tendremos ocasión de comprobar más adelante en un caso especialmente sencillo. \triangleleft

Supongamos que la cuestión de interés para el investigador hubiera sido si $H_0 : \mu \geq 90$ o, por el contrario, $H_1 : \mu < 90$. Sigamos los mismos pasos del caso anterior:

- 1) Estructura estadística: La misma del caso anterior.
- 2) Hipótesis:

$$H_0 : \mu \geq 90 \quad , \quad H_1 : \mu < 90$$

- 3) Estadístico de contraste. Su distribución bajo la hipótesis nula:

$$Z = \frac{\bar{X} - \mu_0}{\sqrt{100}/\sqrt{10}}$$

y el valor experimental es $z_{\text{exp}} = \frac{84.02-90}{\sqrt{100}/\sqrt{10}} = -1.89$. La distribución del estadístico de contraste Z es normal de media 0 y varianza 1 si la hipótesis nula es cierta.

- 4) Regla de decisión: En primer lugar comprobaremos si \bar{x} es o no compatible con H_1 . Así, si la hipótesis alternativa es $H_1 : \mu < \mu_0$, debe suceder que $\bar{x} < \mu_0$, pues, en otro caso, se acepta H_0 directamente. Si $\bar{x} < \mu_0$, fijaremos el nivel de significación $\alpha = 0.05$. Necesitamos también determinar la región crítica. Para ello observemos la definición del estadístico de contraste. Puesto que la media muestral \bar{x} es una estimación de la media desconocida μ , valores demasiado pequeños del estadístico de contraste Z deberían hacernos pensar que la hipótesis nula no es cierta, pues, si la verdadera media poblacional fuese ≥ 90 , la media muestral $\bar{x} = 84.02$ debería ser ≥ 90 o estar próxima a 90. Es decir, valores de Z mucho más pequeños que 0 nos obligarán a rechazar la hipótesis. Cuánto más pequeño que 0 debe ser el valor experimental es algo que depende del nivel de significación elegido. Lo dicho nos lleva a elegir una región crítica o de rechazo para este problema de la forma

$$\frac{\bar{X} - 90}{\sqrt{100}/\sqrt{10}} < C$$

donde C es una cierta constante que determinaremos de acuerdo con el nivel de significación α deseado. Puesto que hemos elegido $\alpha = 0.05$ se pretende que la región de rechazo tenga probabilidad 0.05 supuesto que la hipótesis es cierta; puesto que la región de rechazo tiene una cola (es decir, sólo valores pequeños del estadístico de contraste nos obligan a rechazar la hipótesis), exigiremos que esa cola tenga probabilidad $\alpha = 0.05$; las tablas de la distribución normal de media 0 y varianza 1 (tabla III.1), que es la distribución del estadístico de contraste Z si la hipótesis nula es cierta, nos proporcionan el valor $C = -z_{2\alpha} = -1.645$ como el número a la izquierda del cual queda un área igual a 0.05 bajo la curva de esa distribución normal. Así pues, la región de aceptación es el intervalo $[-1.645, +\infty[$, y la región de rechazo es su complementario. Como el valor experimental $z_{\text{exp}} = -1.89$ (es decir, el valor del estadístico de contraste para la muestra obtenida) queda fuera de la región de aceptación, debemos concluir que los datos obtenidos no son compatibles con la hipótesis nula, lo que nos lleva a rechazar H_0 . Nótese que, en este caso, el valor P es 0.0294, que es la probabilidad de que $Z \leq -1.89$ supuesto conocido que la hipótesis nula H_0 es cierta.

Observación 3.28. Nótese como, a pesar de haber trabajado con los mismos valores, las decisiones en los casos unilateral (rechazar la hipótesis nula) y bilateral (aceptarla) son diferentes. Por ello, hay que decidir si se va a llevar a cabo un test bilateral o unilateral antes de obtener los datos. \triangleleft

Observación 3.29. (Extensión al caso no normal para muestras grandes) Lo dicho anteriormente es válido debido a la hipótesis de normalidad de las variables que estamos observando, pues se usa que la media muestral \bar{X} tiene distribución $N(\mu, \sigma_0^2/n)$. Cuando la distribución de la variable observada no es normal pero el tamaño n de la muestra es suficientemente grande (a efectos prácticos suele bastar $n \geq 60$), el teorema del límite central garantiza que la media muestral \bar{X} se distribuye aproximadamente como una normal, y todo lo dicho anteriormente es aproximadamente válido. \triangleleft

Observación 3.30. (★) (Tamaño de muestra necesario para detectar alternativas no muy próximas a la hipótesis nula) Hemos advertido anteriormente que no es posible controlar simultáneamente las probabilidades α y β de los dos tipos de error a un tamaño de muestra n dado, pero que aumentando este podemos hacer también fiable una decisión por H_0 . Veamos de qué forma en el problema unilateral de contrastar la hipótesis nula $H_0 : \mu \leq \mu_0$

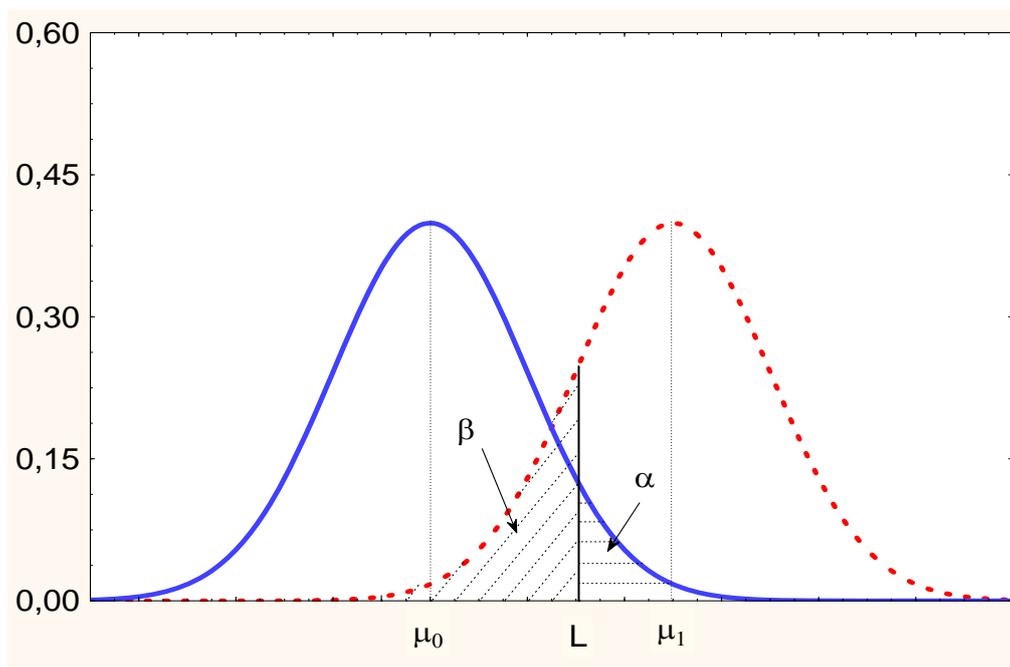


Figura 3.4: Errores α y β en un test de hipótesis

contra $H_1 : \mu > \mu_0$. Nótese que, para cada alternativa concreta $\mu_1 > \mu_0$, podemos hablar de una probabilidad β de error de tipo II (probabilidad de aceptar la hipótesis nula cuando el verdadero valor del parámetro es μ_1). De acuerdo con lo dicho anteriormente, si el valor experimental \bar{X} cae fuera del intervalo $] -\infty, \mu_0 + z_{2\alpha}\sigma_0/\sqrt{n}]$, que es lo mismo que decir que $\sqrt{n}(\bar{X} - \mu_0)/\sigma_0 > z_{2\alpha}$, entonces se rechaza la hipótesis nula. Sea μ_1 una alternativa a H_0 (es decir $\mu_1 > \mu_0$) y hagamos $L = \mu_0 + z_{2\alpha}\sigma_0/\sqrt{n}$. Entonces, la probabilidad β de cometer un error de tipo II para esa alternativa μ_1 es la rayada en la figura siguiente. Se deduce de ello que también $L = \mu_1 - z_{2\beta}\sigma_0/\sqrt{n}$. Igualando ambas expresiones para L y despejando n se obtiene

$$n = \left(\frac{z_{2\alpha} + z_{2\beta}}{\delta} \right)^2 \sigma_0^2$$

donde $\delta = \mu_1 - \mu_0$. Así pues, tomando un tamaño de muestra n como ése podemos garantizar, no sólo que la probabilidad de cometer un error de tipo I es inferior o igual a α , sino que, además, las probabilidades de error de tipo II para alternativas μ_1 que distan de la hipótesis nula μ_0 una cantidad superior o igual a δ son inferiores o iguales a β (fijado de antemano).

La fórmula obtenida para n muestra claramente que el tamaño de muestra aumenta cuando δ disminuye, es decir, que nos veremos obligados a aumentar el tamaño de la muestra si queremos distinguir de la hipótesis nula alternativas que se aproximan más y más a ella.

Por otra parte, la tabla siguiente y la Figura 3.5 se refieren al problema de contrastar $H_0 : \mu \leq 0$ contra $H_1 : \mu > 0$ a partir de una muestra de tamaño n de una distribución normal de varianza conocida $\sigma_0^2 = 1$.

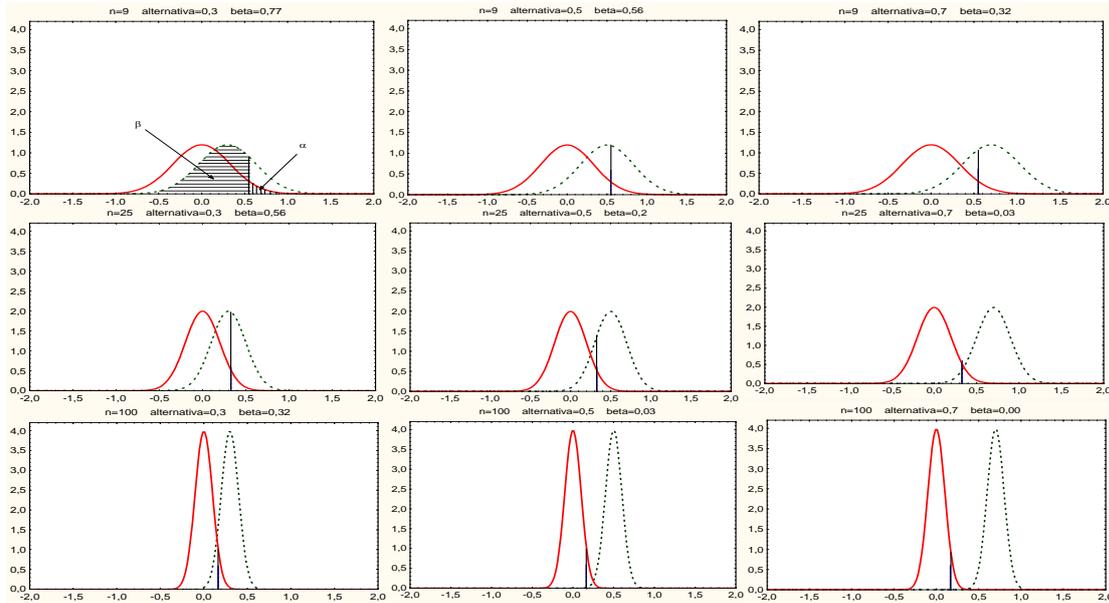


Figura 3.5: Potencia del test unilateral sobre una media contra diferentes alternativas para diferentes tamaños muestrales

$\beta(n, \mu_1)$	$\mu_1 = 0.3$	$\mu_1 = 0.5$	$\mu_1 = 0.7$
$n = 9$	0.77	0.56	0.32
$n = 25$	0.56	0.19	0.03
$n = 100$	0.09	0.00	0.00

En ellas se presentan la probabilidad de error β de tipo II para diferentes alternativas a la hipótesis nula (concretamente, se consideran las alternativas $\mu_1 = 0.3$, $\mu_1 = 0.5$ y $\mu_1 = 0.7$) y diferentes tamaños muestrales (concretamente, se consideran los tamaños muestrales $n = 9$, $n = 25$ y $n = 100$). En los distintos gráficos, se representa con una línea continua la densidad de la media muestral \bar{X} cuando $\mu = 0$ y con línea punteada la densidad de la media muestral \bar{X} cuando $\mu = \mu_1$. En cada caso, la probabilidad de error de tipo I se ha fijado igual a 0.05 y viene dada por el área que queda a la derecha de la línea vertical bajo la curva de Gauss continua, mientras que la probabilidad β de error de tipo II depende del tamaño de muestra y de la alternativa, y coincide con el área que queda bajo la campana de Gauss punteada a la izquierda de la línea vertical. Nótese como, en general, β disminuye a medida que la alternativa se aleja de cero y a medida que el tamaño de muestra aumenta. \triangleleft

Observación 3.31. (Intervalos de confianza y contraste de hipótesis) Los intervalos de confianza resultan útiles a la hora de matizar las conclusiones obtenidas por un test de hipótesis. Consideremos, por ejemplo, el problema de decidir si la media μ desconocida es o no igual a 90. Se trata, en concreto, de contrastar la hipótesis nula $H_0 : \mu = 90$

contra $H_1 : \mu \neq 90$. Si, realizado el experimento, el test descrito anteriormente proporciona significación, cabe preguntarse cómo de distinta es de 90 la media desconocida. Un intervalo de confianza para μ arrojaría algo de luz sobre esa cuestión. Incluso, en el caso de que se aceptase la hipótesis nula H_0 , sería conveniente presentar un intervalo de confianza para μ , pues ya sabemos que una decisión por H_0 es menos fiable que una decisión por H_1 y que debe entenderse más bien como que no hemos podido rechazar la hipótesis nula.

Por otra parte, el procedimiento descrito para realizar un test de hipótesis consiste en elegir un estadístico de contraste y, para este, un intervalo de aceptación bajo H_0 , comprobando después si el valor experimental cae o no dentro de ese intervalo de aceptación. Un procedimiento alternativo, que no estudiaremos en detalle, consistiría en construir un intervalo de confianza para el parámetro desconocido y decidir si el valor propuesto por la hipótesis nula cae o no dentro de él. \triangleleft

Capítulo **4**

**Problemas de Inferencia Estadística
sobre una o dos Muestras**

4.1. Problemas de Inferencia Estadística sobre una Muestra

4.1.1. Inferencia sobre la media de una distribución normal

Ya hemos comentado anteriormente que, en el problema de contrastar la media de una distribución normal, no es razonable suponer desconocida la media y conocida la varianza. Esa suposición sólo nos ha servido para simplificar el primer ejemplo de intervalo de confianza y de contraste de hipótesis, y para introducir conceptos como los de nivel mínimo de significación y de test o intervalo de confianza. Ahora supondremos que la varianza es también desconocida.

Consideremos en primer lugar el problema de contrastar la hipótesis nula $H_0 : \mu = \mu_0$ contra la hipótesis alternativa bilateral $H_1 : \mu \neq \mu_0$ a partir de una muestra de tamaño n de una distribución normal $N(\mu, \sigma^2)$ de media y varianza desconocidas. En este caso no podemos usar el estadístico de contraste

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

pues depende del parámetro desconocido σ , y un estadístico sólo es función de las observaciones; utilizaremos, sin embargo, el estadístico de contraste

$$t = \sqrt{n} \frac{\bar{X} - \mu_0}{S}$$

que se obtiene del anterior reemplazando la varianza desconocida σ^2 por la varianza muestral $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$. Éste es un estadístico de contraste diferente del anterior y tiene una distribución diferente bajo la hipótesis $\mu = \mu_0$. La región crítica para el problema bilateral de contrastar $H_0 : \mu = \mu_0$ contra $H_1 : \mu \neq \mu_0$ es de la forma

$$|t| > C$$

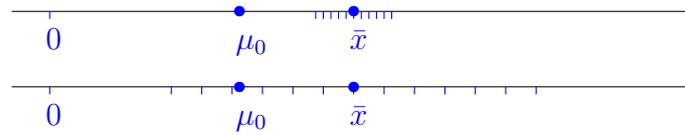


Figura 4.1: Importancia de la dispersión para el contraste de la media

pues $|t|$ es el cociente de la distancia de \bar{X} (estimación de la verdadera media poblacional μ) a μ_0 y un múltiplo de la desviación típica muestral S (que nos da una idea sobre si la diferencia observada entre \bar{X} y μ_0 es debida al azar o a que la hipótesis nula no es cierta). Podemos justificar intuitivamente el uso de ese estadístico de contraste. Puesto que \bar{X} es un estimador de μ , podemos pensar en rechazar la hipótesis nula $\mu = \mu_0$ si $|\bar{X} - \mu_0|$ es grande. Si esa distancia es, por ejemplo, igual a 3, nos encontramos con el problema de decidir si 3 es o no grande. Concretamente, cabe preguntarse si esa discrepancia de 3 unidades puede quedar explicada por el azar (como ocurre en la parte de abajo de la Figura 4.1) o no (como ocurre en la parte de arriba de esa misma figura).

Es por ello que la distancia $|\bar{X} - \mu_0|$ debe compararse con una medida de la variabilidad de los datos, como hace el estadístico t , que la compara con un múltiplo de la desviación típica muestral. Si ese cociente es grande (es decir, mayor que una cierta constante C) rechazaríamos la hipótesis nula y si no la aceptaríamos. La constante C queda también determinada por el nivel de significación α elegido y por la distribución del estadístico de contraste cuando la hipótesis nula es cierta. Esa distribución es la llamada distribución t de Student con $n - 1$ grados de libertad. Por tanto, la constante C es igual a $t_\alpha(n - 1)$, siendo este el valor de la tabla de la distribución t con $n - 1$ grados de libertad (tabla IV) para el que la región crítica tiene probabilidad α . Para resumir lo dicho, consideremos un nuevo ejemplo.

Ejemplo 4.1. Supongamos que se dispone de 10 determinaciones de glucosa en suero para otros tantos individuos extraídos al azar de la población; estamos interesados en decidir si el valor medio de glucosa en suero para esa población es o no diferente de 90. El planteamiento sugiere el uso de un test bilateral. Los cuatro pasos del procedimiento son en este caso los siguientes:

- 1) Estructura estadística: Los datos obtenidos se han resumido en dos: la media muestral

$\bar{x} = 84.02$ mg/dl y la desviación típica muestral $s = 10$ mg/dl. Las 10 determinaciones constituyen una muestra aleatoria de una población en la que el nivel de glucosa se supone normalmente distribuido. La varianza de esa distribución se supone también desconocida.

- 2) Hipótesis:

$$H_0 : \mu = 90 \quad , \quad H_1 : \mu \neq 90$$

- 3) Estadístico de contraste. Su distribución bajo la hipótesis nula: Puesto que la varianza es desconocida, el estadístico de contraste es

$$t = \sqrt{10} \frac{\bar{X} - 90}{S}$$

y el valor experimental es $t_{\text{exp}} = \frac{84.02 - 90}{10/\sqrt{10}} = -1.89$. La distribución del estadístico de contraste t es la distribución t de Student $t(9)$ con 9 grados de libertad, si la hipótesis nula es cierta.

- 4) Regla de decisión: Fijaremos el nivel de significación $\alpha = 0.05$. La región crítica es, según hemos dicho anteriormente,

$$|t| > t_{0.05}(9)$$

donde $t_{0.05}(9) = 2.26$ (véase la tabla IV de la distribución t de Student). Puesto que el valor experimental $t_{\text{exp}} = -1.89$ tiene valor absoluto menor que 2.26, podemos aceptar la hipótesis nula al nivel de significación $\alpha = 0.05$. Debe notarse aquí que cualquier hipótesis nula puede ser aceptada con sólo fijar un nivel de significación α lo bastante pequeño (recuérdese nuestro compromiso de elegir normalmente $\alpha = 0.05$). Según se ha advertido anteriormente, es siempre conveniente proporcionar el nivel mínimo de significación P ; en nuestro ejemplo, la tabla IV de la distribución t no es lo bastante precisa como para proporcionarnos el valor exacto del P . Cuando eso ocurre, proporcionaremos una cota inferior y una cota superior para ese valor P ; en el ejemplo, a la vista de la tabla, se tiene $0.05 < P < 0.1$, y conviene señalar indicios de significación. En este ejemplo, como en el ejemplo que en la sección 10.2 nos sirvió para introducir los cuatro pasos del procedimiento de ejecución de un test de hipótesis, se han considerado 10 determinaciones de glucosa que dan lugar a una media muestral $\bar{x} = 84.02$ y se desea contrastar la misma hipótesis nula $\mu = 90$; la diferencia entre ambos casos es que en aquel la varianza se supone conocida e igual a 100 mientras que en éste la varianza se supone desconocida y se estima por la varianza muestral, que hemos supuesto igual a 100; ello tiene como consecuencia que el valor absoluto del valor experimental del test (1.89 en ambos casos) se compara en aquel caso con el valor $z_\alpha = 1.96$ y en éste con el valor $t_\alpha(9) = 2.26$, debido a que los estadísticos de contraste son diferentes y tienen bajo la hipótesis nula distribuciones diferentes; de hecho, si el valor absoluto del valor experimental hubiera sido igual a 2.1 (como hubiéramos obtenido si la media muestral hubiera sido $\bar{x} = 83.36$), la decisión hubiera

sido rechazar la hipótesis nula en el caso de varianza conocida y aceptarla en el caso de varianza desconocida. Podemos decir que el hecho de suponer desconocida la varianza nos hace ser más cautos a la hora de considerar significativa la discrepancia observada entre la media muestral y el valor de la media poblacional propuesto por la hipótesis nula. \triangleleft

Observación 4.2. (Test unilateral: varianza desconocida) Si en el ejemplo anterior hubiéramos considerado el problema de contrastar la hipótesis unilateral $H_0 : \mu \geq 90$ contra $H_1 : \mu < 90$, el procedimiento establecido nos conduciría a la región crítica unilateral $t < -t_{2\alpha}(n-1)$ pues toda la región crítica está en la cola inferior y debe dar lugar a un área α bajo la densidad de la distribución del estadístico de contraste. Análogamente, para el problema de contrastar la hipótesis unilateral $H_0 : \mu \leq 90$ contra $H_1 : \mu > 90$, el procedimiento nos conduce a la región crítica unilateral $t > t_{2\alpha}(n-1)$. \triangleleft

Observación 4.3. (Intervalo de confianza para la media de una distribución normal de varianza desconocida) Para determinar un intervalo de confianza para la media desconocida μ de una distribución normal de varianza desconocida σ^2 haremos uso del hecho que, supuesto que μ es el verdadero valor de la media, $\sqrt{n}(\bar{X} - \mu)/S$ tiene distribución $t(n-1)$. Entonces, dado $0 < \alpha < 1$, la probabilidad de que

$$-t_{\alpha}(n-1) \leq \sqrt{n} \frac{\bar{X} - \mu}{S} \leq t_{\alpha}(n-1)$$

es $1 - \alpha$. Se deduce de ello que

$$\left[\bar{X} - t_{\alpha}(n-1) \frac{S}{\sqrt{n}}, \bar{X} + t_{\alpha}(n-1) \frac{S}{\sqrt{n}} \right]$$

es un intervalo de confianza para la media desconocida μ al nivel de confianza $1 - \alpha$. \triangleleft

Observación 4.4. (Extensión al caso no normal para muestras grandes) Lo dicho anteriormente (tanto para el test como para el intervalo de confianza para una media) es válido debido a la hipótesis de normalidad de las variables que estamos observando, pues se usa que la media muestral \bar{X} tiene distribución $N(\mu, \sigma^2/n)$. Cuando la distribución de la variable observada no es normal pero el tamaño n de la muestra es suficientemente grande (a efectos prácticos suele bastar $n \geq 60$), el teorema del límite central garantiza que la media muestral \bar{X} se distribuye aproximadamente como una normal y todo lo dicho en el apartado anterior es aproximadamente válido, con tal de que las cantidades $t_{\alpha}(n-1)$ sean reemplazadas por las z_{α} de la tabla de la distribución normal (tabla III.1). El argumento intuitivo que lo justifica es que el estadístico de contraste $Z = \sqrt{n}(\bar{X} - \mu_0)/\sigma$ tiene distribución aproximada $N(0, 1)$ si n es grande; como σ es desconocida se estima por la desviación típica muestral S . \triangleleft

Observación 4.5. (Intervalos de normalidad o tolerancia en el caso de parámetros desconocidos) En el tema 6 habíamos considerado el intervalo $[\mu - z_{\alpha}\sigma, \mu + z_{\alpha}\sigma]$ como intervalo de normalidad o tolerancia para una variable con distribución $N(\mu, \sigma^2)$ de parámetros conocidos, en el sentido de que ese intervalo contiene el $100 \cdot (1 - \alpha)\%$ de los individuos normales

de la población (declarando no normales los individuos que quedan fuera de ese intervalo y cometiendo el error de declarar no normales al $100 \cdot \alpha$ % de los individuos normales). Cuando la media y la varianza son desconocidas, podemos pensar en sustituir sus valores por sus estimaciones \bar{X} y S^2 para obtener un intervalo de la forma $[\bar{X} - KS, \bar{X} + KS]$. La constante K se obtiene de la tabla XIV en función de tres cantidades: el tamaño n de la muestra, la proporción π de individuos que queremos caiga dentro del intervalo y la confianza $1 - \alpha$ con que se desea hacer esa afirmación. \triangleleft

4.1.2. Inferencia sobre la varianza de una distribución normal

Si observamos una muestra de tamaño n de una distribución normal $N(\mu, \sigma^2)$ de media y varianza desconocidas, podemos estar interesados en contrastar una hipótesis del tipo $H_0 : \sigma = \sigma_0$ contra la alternativa bilateral $H_1 : \sigma \neq \sigma_0$; un problema de este tipo se presenta, por ejemplo, cuando buscamos información sobre la precisión de un aparato de medida, pues nos permite confirmar si la dispersión de las observaciones tomadas con ese aparato son del orden de σ_0^2 o no. Puesto que buscamos información sobre la varianza desconocida σ^2 , parece razonable basar nuestra decisión en el estadístico de contraste S^2 (varianza muestral): valores demasiado grandes o demasiado pequeños del cociente S^2/σ_0^2 nos aconsejarían rechazar la hipótesis nula. Cómo de grande o pequeño ha de ser ese cociente para que la decisión sea el rechazo de H_0 es algo que, como siempre, dependerá del nivel de significación α deseado para el test y de la distribución de S^2 bajo la hipótesis nula. Sabemos que, supuesta cierta H_0 (es decir, supuesto que la varianza es igual a σ_0^2), la distribución de $(n-1)S^2/\sigma_0^2$ es la distribución $\chi^2(n-1)$ (chi-cuadrado con $(n-1)$ grados de libertad). Dado $\alpha \in]0, 1[$, denotemos por $\chi_{\alpha}^2(n-1)$ el cuantil de orden $1 - \alpha$ de la distribución $\chi^2(n-1)$ (es decir, el número real que deja a su izquierda un área igual a $1 - \alpha$ bajo la densidad de la distribución $\chi^2(n-1)$); estos valores aparecen tabulados en la tabla V. Entonces, el test al nivel α para el problema considerado queda definido por la región de aceptación

$$\frac{\sigma_0^2 \chi_{1-\alpha/2}^2(n-1)}{n-1} \leq S^2 \leq \frac{\sigma_0^2 \chi_{\alpha/2}^2(n-1)}{n-1}$$

Por otra parte, para determinar un intervalo de confianza para la varianza desconocida σ^2 , usaremos el hecho de que, dado un valor σ^2 de la varianza, $(n-1)S^2/\sigma^2$ tiene distribución

$\chi^2(n-1)$; en ese caso, dado $\alpha \in]0, 1[$, la probabilidad de que $(n-1)S^2/\sigma^2$ esté en el intervalo $[\chi_{1-\alpha/2}^2(n-1), \chi_{\alpha/2}^2(n-1)]$ es $1 - \alpha$ y

$$\left[\frac{(n-1)S^2}{\chi_{\alpha/2}^2(n-1)}, \frac{(n-1)S^2}{\chi_{1-\alpha/2}^2(n-1)} \right]$$

es un intervalo de confianza para σ^2 al nivel $1 - \alpha$.

Observación 4.6. Nótese que para los parámetros de una distribución normal los tests y los intervalos de confianza que hemos presentado están relacionados del siguiente modo: se acepta la hipótesis nula considerada para uno de esos parámetros si el valor propuesto por la hipótesis nula cae dentro del intervalo de confianza que hemos presentado para ese parámetro. Esa relación entre el test y el intervalo de confianza no tendrá lugar en el caso de una proporción que estudiamos posteriormente. \triangleleft

4.1.3. (★) Comprobando la normalidad:

Test de normalidad D'Agostino

La hipótesis de normalidad de una variable es muy frecuente en los métodos clásicos de la inferencia estadística y queda, a menudo, justificada probabilísticamente por el teorema del límite central. No obstante, se han propuesto diversos tests y procedimientos gráficos en la literatura que pretenden comprobar esa hipótesis a partir de una muestra. Tradicionalmente se usaba un test χ^2 de bondad de ajuste para ello (véase pág. 171). Aquí presentaremos un test debido a D'Agostino que presenta una potencia razonable y una relativa facilidad de cálculos. Sean x_1, \dots, x_n los valores observados y

$$x_{(1)} \leq \dots \leq x_{(n)}$$

esos mismos valores escritos en orden creciente. El test de D'Agostino se apoya en el estadístico de contraste

$$D = \frac{\sum_i i x_{(i)} - \frac{n+1}{2} \sum_i x_i}{n \sqrt{n [\sum_i x_i^2 - (\sum_i x_i)^2/n]}}$$

cuyo valor se compara con los límites de significación que proporciona la tabla publicada por D'Agostino (tabla VII). Valores grandes o pequeños del estadístico de D'Agostino sugieren

rechazar la hipótesis de normalidad. Concretamente, la tabla de D'Agostino proporciona en función del tamaño n de la muestra y del nivel de significación α dos valores $a_1(n, \alpha) < a_2(n, \alpha)$, y la hipótesis de normalidad se rechazará si el valor D experimental del test queda fuera del intervalo abierto determinado por ellos.

Observación 4.7. Los programas estadísticos habituales proporcionan otros tests para contrastar la hipótesis de normalidad, como pueden ser el test de bondad de ajuste de Kolmogorov-Smirnov que permite describir en qué medida una distribución concreta -por ejemplo, una distribución normal de parámetros conocidos- ajusta bien a un cierto conjunto de datos; o la modificación de Lilliefors del test de Kolmogorov-Smirnov para el caso de parámetros desconocidos que se utiliza para contrastar la hipótesis de normalidad; o el test de Shapiro-Wilks, que resulta algo más potente que el de D'Agostino contra cierto tipo de alternativas, pero que queda seriamente afectado en el caso de datos repetidos. Zar (1999) propone un test para contrastar la hipótesis de normalidad, llamado de D'Agostino-Pearson, basado simultáneamente en los llamados coeficientes de asimetría y de curtosis y que rechaza la hipótesis de normalidad cuando la simetría y/o la curtosis del conjunto de datos difieren bastante de las que cabría esperar en el caso normal. \triangleleft

Observación 4.8. Entre los métodos gráficos que se usan para comprobar “a ojo” la normalidad se encuentran el histograma, que ya estudiamos en el tema 1, y el “gráfico de probabilidad normal”, en el que se representan los cuantiles muestrales en el eje X frente a los correspondientes cuantiles de la distribución normal $N(0, 1)$, debiendo quedar los puntos así obtenidos sobre una línea recta si el ajuste por la normal es aceptable; este último gráfico resulta útil también a la hora de detectar posibles datos anómalos. \triangleleft

4.1.4. Inferencia sobre una proporción

A continuación consideramos una v.a. X con distribución binomial $B(n, p)$ de parámetro p desconocido, de la que se ha tomado una observación que denotaremos por x . Fijamos un valor concreto p_0 entre 0 y 1 y consideramos la hipótesis nula $H_0 : p = p_0$. Para resolver ese problema, describiremos un test basado en la aproximación de la binomial por la normal, para lo cual exigiremos la siguiente condición de validez: $np_0q_0 \geq 5$, donde $q_0 = 1 - p_0$. En ese caso cabe admitir que el estadístico $\hat{p} = X/n$ (\hat{p} es la proporción muestral) tiene, supuesta cierta la hipótesis nula, distribución aproximada normal $N(p_0, p_0q_0/n)$. El problema considerado se convierte en el problema de contrastar la hipótesis $\mu = p_0$ sobre la media de una distribución normal, problema que ya hemos estudiado anteriormente.

Si consideramos la alternativa bilateral $H_1 : p \neq p_0$, la región crítica al nivel α del test será

$$\frac{|\hat{p} - p_0|}{\sqrt{p_0q_0/n}} > z_\alpha$$

A efectos de cálculo, la región crítica suele escribirse en la forma

$$\frac{|X - np_0|}{\sqrt{np_0q_0}} > z_\alpha$$

Observación 4.9. (★) (Corrección por continuidad) En realidad, se suele utilizar en la literatura la región crítica

$$\frac{|\hat{p} - p_0| - \frac{1}{2n}}{\sqrt{p_0q_0/n}} > z_\alpha \quad \text{o, lo que es lo mismo,} \quad \frac{|X - np_0| - 0.5}{\sqrt{np_0q_0}} > z_\alpha$$

donde el 0.5 del numerador es una corrección por continuidad que queda justificada por el hecho de que estamos aproximando una distribución discreta (como es la binomial) por una continua, como es la distribución normal, según habíamos señalado en una observación de la página 100. No obstante, por simplicidad, en un libro básico de bioestadística, preferimos expresar la región crítica tal y como lo hemos hecho. ◀

En el caso unilateral, comprobaremos en primer lugar si \hat{p} es o no compatible con H_1 . Así, si la hipótesis alternativa es $H_1 : p < p_0$, debe suceder que $\hat{p} < p_0$, pues, en otro caso, se acepta H_0 directamente. En el caso considerado, rechazaremos la hipótesis nula cuando

$$\frac{X - np_0}{\sqrt{np_0q_0}} < -z_{2\alpha}$$

donde se ha tomado el nivel crítico $-z_{2\alpha}$ para que la única cola que queda tenga exactamente área α .

Si queremos obtener un intervalo de confianza para el parámetro p haciendo uso de la aproximación a la normal, exigiremos las condiciones de validez siguientes: $n\hat{p}\hat{q} \geq 5$. En esas condiciones la v.a. X tiene distribución aproximada $N(np, npq)$ y, dividiendo por n , \hat{p} tiene distribución aproximada $N(p, pq/n)$. De acuerdo con lo obtenido en el caso de la media de una normal, un intervalo de confianza aproximado para p al nivel $1 - \alpha$ es

$$\left[\hat{p} - z_\alpha \sqrt{\frac{\hat{p}\hat{q}}{n}}, \hat{p} + z_\alpha \sqrt{\frac{\hat{p}\hat{q}}{n}} \right]$$

donde se han sustituido los valores desconocidos p y q en la desviación típica $\sqrt{pq/n}$ por sus estimaciones \hat{p} y \hat{q} .

Ejemplo 4.10. Con el fin de comparar dos tratamientos A y B para una enfermedad, se eligen al azar 102 individuos que han recibido ambos tratamientos y se constata que 58 prefieren el tratamiento A y 44 el B. Estamos interesados en decidir si esos datos presentan evidencias de que un tratamiento sea preferido al otro. Si p es la probabilidad desconocida de que un individuo de la población prefiera el tratamiento A, la v.a. X que describe el número de individuos entre los 102 seleccionados que prefieren ese tratamiento tiene distribución binomial $B(102, p)$ y deseamos contrastar la hipótesis nula $p = 0.5$ contra $p \neq 0.5$; se comprueba que estamos en las condiciones exigidas para usar la aproximación normal, que el valor experimental es 1.386, lo que da un valor P superior a 0.16, y no hay evidencias de que un tratamiento sea preferido al otro. Por lo demás, un estimador de p es $\hat{p} = 58/102 = 0.5686$ y un intervalo al 95 % de confianza para p es $[0.473, 0.665]$. \triangleleft

Observación 4.11. (★) (Bondad de ajuste a una distribución dada de la distribución de una variable cualitativa) Supongamos ahora que X es una v.a. cualitativa que toma un número finito de valores a_1, \dots, a_m y que estudios precedentes afirman que la distribución de X en una población viene dada por las probabilidades $P(X = a_1) = p_{01}, \dots, P(X = a_m) = p_{0m}$; obviamente, debe ocurrir que $\sum_i p_{0i} = 1$. Queremos confirmar si las probabilidades actuales p_1, \dots, p_m siguen siendo esas (o si las probabilidades p_1, \dots, p_m en otra población coinciden con esas), es decir, se trata de contrastar la hipótesis nula $H_0 : p_1 = p_{01}, \dots, p_m = p_{0m}$ contra que alguna de esas igualdades no es cierta. Para ello, seleccionamos N individuos al azar y determinamos los valores de la variable para cada uno de ellos (es decir, los clasificamos en alguna de las clases disjuntas y excluyentes a_1, \dots, a_m). Dado $1 \leq i \leq m$, denotaremos por O_i el número de individuos observados en la clase a_i ; nótese que O_i tiene distribución binomial $B(N, p_i)$ y que $\sum_{i=1}^m O_i = N$. Si la hipótesis nula fuese cierta, la proporción de individuos de la clase a_i sería p_{0i} y cabría esperar que en esa clase a_i hubiera $E_i = Np_{0i}$ individuos; también las cantidades esperadas verifican que $\sum_i E_i = N$. La discrepancia entre los valores observados O_i y los esperados E_i si H_0 se supone cierta nos ayudará a tomar la decisión de aceptar o rechazar esa hipótesis. Para medir esa discrepancia utilizaremos la llamada distancia χ^2 definida por

$$\chi^2 = \sum_{i=1}^m \frac{(O_i - E_i)^2}{E_i}$$

Valores grandes de este estadístico sugieren rechazar la hipótesis nula. Concretamente, la región crítica al nivel α es

$$\chi^2 > \chi_\alpha^2(m - 1)$$

pues la distribución del estadístico de contraste es aproximadamente la distribución $\chi^2(m - 1)$ cuando se verifican las siguientes condiciones de validez: ninguna de las cantidades esperadas E_i es estrictamente menor que 1 y no más del 20 % de ellas son menores o iguales que 5. En el caso $m = 2$, X es una variable dicotómica y se trata en realidad de contrastar la hipótesis nula $p_1 = p_{01}$ (pues si aceptamos esa hipótesis, deberemos también aceptar que $p_2 = p_{02}$ ya que $p_1 + p_2 = 1$ y $p_{01} + p_{02} = 1$). Se prueba que, en ese caso, el test coincide con el que ya hemos estudiado para contrastar una proporción. \triangleleft

Observación 4.12. (★) (Bondad de ajuste a una distribución dada de la distribución de una variable cuantitativa) El test explicado en la observación anterior se adapta sin modificación alguna al caso de una variable cuantitativa discreta. Pero también se puede utilizar para contrastar que la distribución de una variable cuantitativa sigue una distribución dada; por ejemplo, si queremos contrastar la hipótesis de que una v.a. X tiene distribución $N(0, 1)$ podemos proceder del siguiente modo: (a) dividimos la recta real en un número finito de intervalos disjuntos a_1, \dots, a_m ; (b) calculamos las probabilidades p_{01}, \dots, p_{0m} que la distribución $N(0, 1)$ asigna a cada uno de ellos; (c) se toma una muestra de tamaño N de la población y se determinan los valores x_1, \dots, x_N de la variable X para cada individuo de la muestra; (d) para $1 \leq i \leq m$, se determina la cantidad O_i de valores que caen en el intervalo a_i . No hemos hecho nada más que transformar una variable cuantitativa en otra cualitativa; a partir de aquí, se procede como en la observación anterior; si se acepta la hipótesis nula, aceptaremos que la variable X tiene distribución $N(0, 1)$. \triangleleft

4.2. Comparación de dos Muestras

4.2.1. Introducción: muestras independientes y muestras relacionadas

Buena parte de los problemas básicos que aparecen en bioestadística tienen que ver con los términos comparación y relación: comparación de un mismo parámetro en dos o más poblaciones (o la misma población cuando cambian las condiciones de experimentación) y relación entre dos o más variables determinadas para un mismo individuo. Esos dos términos realmente se pueden resumir en uno solo, relación, pues la comparación de los valores de una misma variable en dos poblaciones no es más que el estudio de la relación entre esa variable y la variable cualitativa que asigna a cada individuo la población a la que pertenece. Pero mantendremos la terminología habitual.

En este tema abordamos el problema de comparar dos muestras extraídas de sendas poblaciones. Cabe, por ejemplo, preguntarse si dos tipos de medicamentos son igualmente efectivos, o si dos técnicas de medida son igualmente precisas, o si la presión sanguínea sistólica cambia tras la administración de una cierta droga, etc. Usualmente, el investigador estará interesado en comprobar si las medias de las dos poblaciones son o no las mismas (ejemplo de la presión sanguínea sistólica); también puede estar interesado en la comparación de dos proporciones (por ejemplo, si la proporción de personas satisfechas con la atención sanitaria que reciben es o no la misma en España que en Portugal) o en la comparación de dos varianzas (ejemplo de la comparación de la precisión de dos aparatos de medida). Interesa especialmente el problema de comparar dos medias o dos proporciones.

Un ejemplo nos ayudará a introducir una técnica frecuentemente utilizada en bioestadística con el fin de eliminar fuentes de variabilidad en los resultados del experimento. Imaginemos que deseamos contrastar la eficacia de un fármaco en el tratamiento de la

hipertensión en una cierta población de individuos hipertensos. La presión sanguínea sistólica (PSS) media en esa población es una cierta cantidad desconocida μ_1 y cabe pensar que si a esa población se le administrase el citado fármaco la PSS media de la población quedaría, en principio, modificada y pasaría a ser una cierta cantidad μ_2 . Se pretende contrastar la hipótesis nula $H_0 : \mu_1 = \mu_2$ (que viene a significar que el fármaco no tiene efectos significativos sobre la PSS de la población) contra la hipótesis alternativa $H_1 : \mu_1 \neq \mu_2$. Con ese objetivo seleccionamos al azar 20 individuos de la población. A partir de aquí podemos considerar dos posibles diseños del experimento:

- 1) Muestras independientes: La primera consiste en aplicar el fármaco a un subgrupo de 10 individuos (que llamaremos grupo de tratamiento) mientras que con los 10 restantes se forma el que llamaremos grupo de control y no recibirán el tratamiento; al cabo de un tiempo prudencial se tomará la PSS de los 20 individuos y se compararán los resultados obtenidos para ambos grupos. Se trata, en realidad, de dos muestras independientes extraídas de la población una de las cuales recibió el tratamiento y la otra no.
- 2) Muestras relacionadas: Una segunda planificación consiste en determinar la PSS a los 20 individuos antes y después de aplicarles el tratamiento a todos ellos. En este segundo caso los datos aparecen por parejas: PSS antes y PSS después para cada individuo. Se habla entonces de muestras relacionadas o muestras apareadas.

Bajo el primer diseño, si el test resulta significativo admitiremos que $\mu_1 \neq \mu_2$, pero puede quedarnos la duda sobre si las diferencias de tensión observadas pueden deberse a diferencias existentes a priori entre los grupos de tratamiento y control (debidas quizás a la influencia de factores como edad, sexo, . . .); el investigador siempre confiará en que el azar distribuya los individuos de la muestra de forma análoga a como se distribuyen de hecho en la población general (y eso es tanto más así cuanto mayor es la muestra seleccionada), pero en tamaños pequeños la presencia de otros factores ajenos al tratamiento que nos interesa puede influir en el resultado. En cambio, en el experimento relacionado la variabilidad es menor puesto que los datos aparecen por pares que han sido obtenidos del mismo individuo con lo cual, factores no considerados como edad, sexo, . . . , no repercuten en el resultado: la

única diferencia entre las presiones sanguíneas sin y con tratamiento es el tratamiento en sí, con lo cual la significación del test, caso de producirse, sólo puede ser debida al tratamiento. Por ello, el experimento con muestras independientes suele ser discutible si es posible realizar un experimento relacionado.

Observación 4.13. (Apareamiento artificial) Un apareamiento como el considerado se llama natural por que los datos de las muestras se aparean por su pertenencia a un mismo bloque natural. En otras ocasiones el apareamiento natural no es posible y se efectúa un apareamiento artificial. Por ejemplo, si queremos comparar dos fármacos contra la hipertensión, aplicar ambos al mismo individuo puede oscurecer los resultados de la investigación por la superposición de los efectos de ambos fármacos; puesto que la edad influye en la presión sanguínea sistólica (y con el objetivo de reducir la variabilidad de los datos debida al factor edad), podríamos pensar en elegir n individuos al azar de la población objeto de estudio y, para cada uno de ellos, elegir un individuo con su misma edad para construir un par artificial. \triangleleft

Observación 4.14. Debe tenerse en cuenta que los dos datos de un mismo par no se pueden suponer independientes por ser mediciones realizadas en un mismo individuo; pero los pares se pueden suponer independientes entre sí si la extracción de los n primeros individuos se ha hecho realmente al azar. \triangleleft

Observación 4.15. (Efecto placebo y aleatorización) En el caso de dos muestras independientes y cuando trabajamos con seres humanos, para evitar posibles efectos psicológicos en la variable respuesta conviene aplicar a los individuos del grupo de control (sin que estos los sepan) un tratamiento similar al que reciben los del grupo de tratamiento, pero eliminando cualquier principio activo en la medicación utilizada; esto es lo que se conoce como efecto placebo. Por otra parte, para evitar posibles preferencias por parte del investigador, si éste ya conoce a los $m + n$ individuos que formarán parte del estudio, conviene asignar los individuos aleatoriamente a uno u otro grupo. \triangleleft

4.2.2. Comparación de dos varianzas

Más interesante es el problema de comparación de dos medias, pero comenzaremos abordando el problema de contrastar dos varianzas debido, además del interés que tiene el problema en sí mismo, a que en el problema de comparar dos medias uno de los métodos que describiremos (concretamente, el test de Student de comparación de dos medias) hace uso de la hipótesis de igualdad de varianzas de las dos poblaciones consideradas. Esa hipótesis

podría quedar justificada si al menos, en un test de comparación de varianzas, no hemos podido rechazar la hipótesis de homogeneidad de las mismas. ⁽¹⁾

Nos situaremos en el caso de dos muestras independientes. Para describir los pasos del procedimiento descrito en el tema anterior, nos apoyaremos en el siguiente ejemplo.

Ejemplo 4.16. Se desea comparar la variabilidad de dos métodos de medida de la catalasa en úteros de ratas tras dos tratamientos diferentes A y B. Para ello se han seleccionado dos grupos de 15 y 13 ratas, respectivamente, y se ha determinado la catalasa con el primer método en el primer grupo y con el segundo método en el segundo. Los datos del primer grupo se han resumido en su media $\bar{x} = 46.09$ y su desviación típica $S_1 = 8.16$, y los del segundo en su media $\bar{y} = 20.26$ y su desviación típica $S_2 = 4.9$. Nos preguntamos si las varianzas tras ambos tratamientos se pueden considerar iguales. ◁

Describimos a continuación el procedimiento que seguiremos para resolver el problema:

- 1) Estructura estadística: Disponemos de dos muestras x_1, \dots, x_m e y_1, \dots, y_n de dos poblaciones. Los datos se resumen en las medias muestrales \bar{x} e \bar{y} y en las desviaciones típicas muestrales S_1 y S_2 . Supondremos que las dos muestras observadas son independientes, que la primera muestra se extrae de una distribución normal de media μ_1 y varianza σ_1^2 y que la segunda se extrae de una distribución normal de media μ_2 y varianza σ_2^2 .

- 2) Hipótesis: Consideremos primero el problema bilateral

$$H_0 : \sigma_1^2 = \sigma_2^2 \quad , \quad H_1 : \sigma_1^2 \neq \sigma_2^2$$

- 3) Estadístico de contraste. Su distribución bajo la hipótesis nula: Usaremos el estadístico de contraste

$$F = \frac{S_1^2}{S_2^2}$$

Su distribución supuesta cierta la hipótesis nula, es la distribución $F(m-1, n-1)$ de Fisher con $(m-1, n-1)$ grados de libertad.

¹ Aunque algunos autores prefieren el uso de un método alternativo al test de Student –como pueden ser el test de Welch o el test no paramétrico de suma de rangos de Wilcoxon, que veremos más adelante– antes que realizar dos test consecutivos –primero un test de comparación de varianzas y, después, si éste es no significativo, el test de comparación de dos medias–, con la correspondiente propagación del error de tipo I.

- 4) Regla de decisión: Fijaremos el nivel de significación $\alpha = 0.05$. De acuerdo con lo dicho, la región de aceptación del test será de la forma

$$C_1 \leq F \leq C_2$$

pues tanto valores demasiado grandes como demasiado pequeños de F indican una diferencia significativa entre las varianzas muestrales de ambas muestras. Las constantes C_1 y C_2 dependen del nivel de significación α elegido y de la distribución de F bajo la hipótesis nula. Éste es un test bilateral y para determinar esas constantes asignaremos área $\alpha/2$ a cada una de las colas, con lo que obtenemos $C_1 = F_{1-\alpha/2}(m-1, n-1) = F_{\alpha/2}(n-1, m-1)^{-1}$ y $C_2 = F_{\alpha/2}(m-1, n-1)$, donde $F_{\alpha}(m-1, n-1)$ denota el punto que deja a su derecha un área igual a α en la distribución $F(m-1, n-1)$, valores estos que se obtienen de la tabla VI de la distribución F . Téngase en cuenta que esa tabla no proporciona los valores $F_{1-\alpha/2}$, pues pueden calcularse haciendo uso de la relación

$$F_{1-\alpha/2}(m, n) = \frac{1}{F_{\alpha/2}(n, m)}$$

El test puede simplificarse obligando a la mayor varianza muestral a estar en el numerador (si fuera preciso, la primera muestra pasaría a ser la segunda y la segunda pasaría a ser la primera, intercambiando también los papeles de m y n), con lo cual el valor experimental F_{exp} se compara solamente con $F_{\alpha/2}(m-1, n-1)$.

Ejemplo 4.17. (Ejemplo 4.16, continuación) En el ejemplo considerado, $m = n = 15$, y $F_{\text{exp}} = (8.16/4.9)^2 = 2.77$ y $F_{\alpha/2}(14, 12) = F_{0.025}(14, 12) = 3.21$; puesto que $F_{\text{exp}} \leq F_{0.5/2}(14, 12)$, el test no es significativo al nivel 0.05. Se comprueba que sí es significativo al nivel 0.1, con lo cual debemos señalar indicios de significación. \triangleleft

Observación 4.18. Las modificaciones a hacer en el test en el caso unilateral son obvias. \triangleleft

Observación 4.19. (Intervalo de confianza para el cociente de dos varianzas) Si deseamos un intervalo de confianza para el cociente σ_1^2/σ_2^2 de las varianzas de dos distribuciones normales notemos que, dados σ_1^2 y σ_2^2 , el cociente

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2}$$

tiene distribución $F(m - 1, n - 1)$; dado $\alpha \in]0, 1[$, la probabilidad de que ese cociente esté en el intervalo $[F_{1-\alpha/2}(m - 1, n - 1), F_{\alpha/2}(m - 1, n - 1)]$ es $1 - \alpha$. Se sigue de ahí que un intervalo de confianza al nivel $1 - \alpha$ para el cociente σ_1^2/σ_2^2 es:

$$\frac{S_1^2}{S_2^2} \frac{1}{F_{\alpha/2}(m - 1, n - 1)} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{S_1^2}{S_2^2} \frac{1}{F_{1-\alpha/2}(m - 1, n - 1)}$$

Puesto que $1/F_{1-\alpha/2}(m - 1, n - 1) = F_{\alpha/2}(n - 1, m - 1)$, ese intervalo de confianza se escribe en la forma siguiente:

$$\frac{S_1^2}{S_2^2} \frac{1}{F_{\alpha/2}(m - 1, n - 1)} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{S_1^2}{S_2^2} F_{\alpha/2}(n - 1, m - 1)$$

△

4.2.3. Test t de Student de comparación de dos medias: Muestras independientes y varianzas iguales

Estudiamos a continuación el problema de comparación de medias para el caso de muestras independientes y el de muestras relacionadas. Comenzamos considerando el caso de muestras independientes en el caso de que las varianzas de las dos poblaciones se suponen iguales, aunque desconocidas. Después modificaremos el test para el caso general en que las varianzas no tienen por qué ser iguales. En cualquier caso puede ser conveniente realizar previamente el test de comparación de varianzas de la sección anterior antes de decidir qué suposición vamos a hacer sobre las varianzas.

Un ejemplo nos ayudará a fijar las ideas.

Ejemplo 4.20. La tabla siguiente contiene los niveles de potasio en útero para dos grupos de ratas: un grupo de 9 ratas que ha recibido un cierto tratamiento (T) y otro de 10 ratas que permanece como grupo de control (C).

T	3786	4556	3564	4121	3890	4478	4876	3719	4324	
C	3990	4655	5112	4987	4233	4354	4223	5207	4887	4775

Tabla: Niveles de potasio en dos grupos de ratas

Disponemos, pues, de $m + n = 9 + 10$ datos $x_1, \dots, x_m, y_1, \dots, y_n$ que representan los niveles de potasio en útero de las ratas consideradas y se resumen como sigue: $\bar{x} = 4146$, $\bar{y} = 4642.3$, $S_1 = 441.55$ y $S_2 = 420.04$. Suponemos que los datos constituyen dos muestras independientes de dos distribuciones normales con medias μ_1 y μ_2 y varianzas desconocidas.

Las varianzas se suponen iguales a σ^2 (una aplicación del test de comparación de varianzas de la sección anterior a los datos obtenidos resuelve que no podemos rechazar la hipótesis nula al nivel de significación $\alpha = 0.2$). Se trata de contrastar la hipótesis nula $H_0 : \mu_1 = \mu_2$ contra $H_1 : \mu_1 \neq \mu_2$. Puesto que, bajo la hipótesis nula, las varianzas se suponen iguales, consideraremos la varianza muestral combinada

$$S^2 = \frac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-2} = \frac{\sum_i (X_i - \bar{X})^2 + \sum_j (Y_j - \bar{Y})^2}{m+n-2}$$

como estimación de la varianza común desconocida y basaremos nuestra decisión en el estadístico de contraste

$$t = \frac{\bar{X} - \bar{Y}}{S \sqrt{\frac{1}{m} + \frac{1}{n}}}$$

que tiene, bajo H_0 , distribución $t(m+n-2)$ (t de Student con $m+n-2$ grados de libertad). Nótese que no bastaría con tomar $|\bar{X} - \bar{Y}|$ como medida de la discrepancia entre ambas medias, sino que esa cantidad debe compararse con la variabilidad S presentada por los datos para decidir si la diferencia observada entre las dos medias muestrales es sólo debida al azar o al hecho de que las medias poblacionales son diferentes. En nuestro caso el valor experimental es $t_{\text{exp}} = -2.51$. Valores grandes del estadístico de contraste deben hacernos sospechar que la diferencia entre las dos medias es grande. La región crítica al nivel $\alpha = 0.05$ es, entonces

$$|t| > t_\alpha(m+n-2)$$

La tabla de la distribución t de Student para el caso de 17 grados de libertad (tabla IV) proporciona un valor $0.01 < P < 0.05$, es decir, la diferencia entre los niveles de potasio comparados es significativa. \triangleleft

4.2.4. Test de Welch de comparación de dos medias: Muestras independientes y varianzas distintas

Si el test de comparación de varianzas es significativo, las varianzas no se pueden considerar iguales y no se conoce ninguna solución exacta al problema; pero sí disponemos de soluciones aproximadas. El test de Welch que describimos a continuación es una de ellas. Ese test propone el estadístico de contraste

$$t = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}}$$

Si hubiésemos supuesto las varianzas σ_1^2 y σ_2^2 conocidas, el estadístico de contraste natural hubiera sido

$$\frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$$

y el propuesto por Welch lo único que hace es reemplazar las varianzas desconocidas por las varianzas muestrales. La región crítica propuesta por Welch es

$$|t| > t_\alpha(f)$$

donde el número f de grados de libertad a considerar es

$$f = \frac{(S_1^2/m + S_2^2/n)^2}{\frac{(S_1^2/m)^2}{m-1} + \frac{(S_2^2/n)^2}{n-1}}$$

valor que no tiene por qué ser un entero; suele aproximarse f por el entero más próximo.

Ejemplo 4.21. (Ejemplo 4.20, continuación) Supongamos que en una situación análoga a la del ejemplo precedente, tuviésemos los siguiente valores: $m = 30$, $n = 40$, $\bar{x} = 4146$, $\bar{y} = 4642.3$, $S_1 = 441.55$ y $S_2 = 250.04$. En este caso, el test de comparación de varianzas resulta muy significativo (necesitamos las tablas de la distribución $F(29, 39)$ y sólo aparecen las tablas de las distribuciones $F(29, 30)$ y $F(29, 40)$, pero en ambos casos el valor experimental $F_{exp} = 3.11$ es mayor que el valor crítico $F_{0.005}$, por lo que también $F_{exp} > F_{0.005}(29, 39)$). El valor experimental del test de Welch es $t = -2.97$; con las notaciones precedentes, $f = 12.37$ y el valor P correspondiente está entre 0.01 y 0.02 (haciendo uso de la tabla IV de la distribución t , eso es lo que ocurre tanto para 12 grados de libertad como para 13). Si se desea calcular el valor teórico $t_{0.05}(f)$, mejor que aproximar por el entero más próximo a f (12, en nuestro ejemplo), es hacer una interpolación lineal para aproximar $t_{0.05}(f)$ como se explica en la página 97 a partir de los valores $t_{0.05}(12)$ y $t_{0.05}(13)$. \triangleleft

Observación 4.22. (Tests unilaterales) Sólo hemos descrito los tests de Student y de Welch en el caso bilateral. Las modificaciones que deben hacerse en los mismos en el caso unilateral son obvias. Por ejemplo, si bajo la hipótesis de igualdad de varianzas se pretende contrastar la hipótesis nula $H_0 : \mu_1 \leq \mu_2$ contra la alternativa unilateral $H_1 : \mu_1 > \mu_2$, la región crítica del test de Student sería

$$t = \frac{\bar{X} - \bar{Y}}{S\sqrt{\frac{1}{m} + \frac{1}{n}}} > t_{2\alpha}(m + n - 2)$$

El punto crítico es ahora $t_{2\alpha}(m + n - 2)$ pues sólo consideramos una cola a la derecha que debe tener área α . \triangleleft

Observación 4.23. (Intervalos de confianza para la diferencia de medias) Si alguno de los tests precedentes de comparación de dos medias resulta significativo, cabe preguntarse cómo de diferentes son las medias μ_1 y μ_2 . Si las varianzas fuesen conocidas, el hecho de

que $\bar{X} - \bar{Y}$ tiene distribución normal de media $\mu_1 - \mu_2$ y varianza $(\sigma_1^2/m) + (\sigma_2^2/n)$ prueba que un intervalo de confianza para $\mu_1 - \mu_2$ al nivel de confianza $1 - \alpha$ es

$$\mu_1 - \mu_2 \in (\bar{X} - \bar{Y}) \pm z_\alpha \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}$$

Puesto que la hipótesis de suponer conocidas las varianzas no es razonable, un intervalo de confianza para $\mu_1 - \mu_2$ al nivel de confianza $1 - \alpha$ es

$$\mu_1 - \mu_2 \in (\bar{X} - \bar{Y}) \pm t_\alpha(m+n-2)S\sqrt{\frac{1}{m} + \frac{1}{n}}$$

en el caso de que las varianzas se supongan iguales, y

$$\mu_1 - \mu_2 \in (\bar{X} - \bar{Y}) \pm t_\alpha(f)\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}$$

si las varianzas no se suponen iguales, donde f es el número de grados de libertad que proporciona el test de Welch. \triangleleft

Observación 4.24. (Variables no normales) Los métodos estadísticos anteriores han sido obtenidos bajo la hipótesis de normalidad. Si las variables observadas no siguen una distribución normal, pero los tamaños m y n de las muestras son suficientemente grandes (valores ≥ 60 suelen ser aceptables en la práctica) se prueba que las medias muestrales siguen aproximadamente una distribución normal y los métodos descritos son aproximadamente válidos, sólo que deben reemplazarse las cantidades t_α de las tablas de la distribución t de Student por las z_α correspondientes a la distribución normal (tabla III.1). \triangleleft

4.2.5. (★) Test no paramétrico de suma de rangos de Mann-Whitney-Wilcoxon de comparación de dos muestras independientes

Los tests anteriores de comparación de dos medias en el caso de muestras independientes están basados en la hipótesis de normalidad (tanto en el caso de varianzas iguales como en el de varianzas desiguales). Describimos a continuación un test no paramétrico para el problema de comparación de dos muestras independientes en el caso de que la hipótesis de normalidad no esté bien justificada (es incluso un buen competidor de los tests precedentes en el caso normal): es el test de suma de rangos de Wilcoxon, también llamado test de

Mann-Whitney. Se habla de test no paramétrico para aludir al hecho de que, a diferencia de lo que ocurría en el caso normal, el test no está relacionado con parámetros concretos de las poblaciones consideradas.

Se dispone de $m + n$ números reales $x_1, \dots, x_m, y_1, \dots, y_n$, que suponemos realizaciones concretas de dos muestras independientes X_1, \dots, X_m e Y_1, \dots, Y_n de la distribución de una misma característica numérica en dos poblaciones (o en una misma población cuando cambian las condiciones en que se realiza el experimento). Se trata de contrastar la hipótesis nula de que la distribución común de las X 's coincide con la distribución común de las Y 's.

Observación 4.25. (★) Para concretar algo más la hipótesis nula, necesitaremos el concepto de función de distribución: Si X es una v.a.r. en un espacio de probabilidad (Ω, P) , se define su función de distribución F como la aplicación $F : x \in \mathbb{R} \mapsto F(x) = P(X \leq x)$; se prueba que dos v.a.r. tienen la misma función de distribución si, y sólo si, tienen la misma distribución de probabilidad. En el problema de contraste de hipótesis que consideramos, suponemos que las distribuciones de las X 's y las Y 's sólo se diferencian, quizá, en la posición; es decir: si la función de distribución común a las X_i es F_1 y la función de distribución común a las Y_j es F_2 , se supone que existe $\Delta \in \mathbb{R}$ tal que $F_1(x) = F_2(x + \Delta)$, para cada $x \in \mathbb{R}$, con lo que F_2 se obtiene desplazando la función de distribución F_1 una cantidad Δ . Finalmente, la hipótesis nula H_0 se puede escribir en la forma $H_0 : F_1 = F_2$, o, lo que es equivalente, $H_0 : \Delta = 0$. ◁

Supondremos en lo que sigue que $m \leq n$ (cambiéense las muestras de orden en otro caso). Esos $m + n$ valores se ordenan de menor a mayor para obtener la muestra combinada en orden creciente

$$u_1 \leq \dots \leq u_m \leq u_{m+1} \leq \dots \leq u_{m+n} \quad (1)$$

Supongamos en primer lugar que los valores obtenidos son dos a dos distintos (se dice entonces que no hay empates). Se asignan rangos a cada uno de los elementos de la primera muestra del siguiente modo: el rango R_1 de x_1 es la posición que ocupa el valor x_1 en la muestra combinada (1), ..., el rango R_m de x_m es la posición que ocupa el valor x_m en la muestra combinada (1). Denotaremos por R la suma de los rangos correspondientes a los valores de la primera muestra, es decir, $R = R_1 + \dots + R_m$. Que R sea demasiado grande significa que las x_i tienen tendencia a ser mayores que las y_j pues tienden a aparecer a la derecha en la sucesión reordenada (1); que R sea demasiado pequeño significa que las x_i

tienen tendencia a ser menores que las y_j pues tienden a aparecer a la izquierda en la sucesión reordenada (1). En ambos casos deberíamos rechazar la hipótesis nula. Cómo de grande o de pequeño tiene que ser R es algo que depende, como siempre, del nivel de significación que deseamos para el test y de la distribución del estadístico de contraste bajo la hipótesis nula. Esa distribución aparece tabulada para valores pequeños de m y n : la tabla VIII proporciona dos valores $v_i(\alpha)$ y $v_s(\alpha)$ separados por un guión para $m + n \leq 30$ y para diferentes valores de α ; si el valor de R obtenido se encuentra en el intervalo abierto $]v_i(\alpha), v_s(\alpha)[$ determinado por esos dos valores, se acepta la hipótesis nula al nivel de significación α , y se rechaza en otro caso. En el caso $m + n > 30$ no podemos usar la tabla VIII de la distribución de Wilcoxon, pero podemos utilizar la aproximación normal del siguiente modo: la media de R bajo la hipótesis nula es $E(R) = (m + n + 1)m/2$ y su varianza es $\text{Var}(R) = (m + n + 1)mn/12$, y la hipótesis nula se rechaza si

$$\frac{|R - E(R)|}{\sqrt{\text{Var}(R)}} > z_\alpha$$

Lo dicho es válido en el caso de que no haya empates. Supongamos ahora que hay r grupos de empates en la muestra combinada y reordenada (1) con t_1 empates en el primer grupo, ..., t_r empates en el r -ésimo grupo. A cada dato en un grupo de empates se le asignará el rango medio de los rangos que corresponderían a los datos de ese grupo si estos fueran distintos; por ejemplo, si $x_1 = 2, x_2 = 5, x_3 = 1, x_4 = 2, x_5 = 4, y_1 = 1, y_2 = 6, y_3 = 2$ e $y_4 = 3$, la sucesión reordenada es $1 \leq 1 \leq 2 \leq 2 \leq 2 \leq 3 \leq 4 \leq 5 \leq 6$, y hay dos grupos de empates, uno con dos elementos y otro con 3 elementos; el rango del valor 1 es $(1 + 2)/2 = 1.5$ y el rango del valor 2 es $(3 + 4 + 5)/3 = 4$, con lo que $R_1 = 4, R_2 = 8, R_3 = 1.5, R_4 = 4, R_5 = 7$.⁽²⁾ Se determina la suma R de los rangos de los elementos de la primera muestra. Si $m + n \leq 30$ se busca en la tabla VIII de la distribución de Wilcoxon y, si $m + n > 30$, se utiliza la aproximación normal como lo hemos hecho antes, pero teniendo en cuenta que ahora

$$\text{Var}(R) = \frac{(m + n)[(m + n)^2 - 1] - \sum_{i=1}^r T_i}{12(m + n)} \frac{mn}{m + n - 1}$$

² En general, si el primer grupo de empates en (1) comienza con el dato u_{k_1} , entonces $u_{k_1} = u_{k_1+1} = \dots = u_{k_1+t_1-1}$ y a ese valor común se le asigna rango $[k_1 + (k_1 + 1) + \dots + (k_1 + t_1 - 1)]/t_1 = k_1 + \frac{1}{2}(t_1 - 1)$.

donde $T_i = t_i^3 - t_i$, $1 \leq i \leq r$ ($E(R)$ es ahora igual que antes).

Ejemplo 4.26. (Ejemplo 4.20, continuación) Con los datos de la tabla 12.1, los rangos de los datos de la primera muestra son 3,12,1,6,4,11,15,2 y 9, respectivamente, con lo que se obtiene una suma de rangos $R = 63$. Para $\alpha = 0.05$, la tabla VIII muestra que el intervalo $]v_i(\alpha), v_s(\alpha)[$ es igual a $]65, 115[$ y, para $\alpha = 0.01$, es $]58, 122[$. Luego $0.01 < P < 0.05$. \triangleleft

Observación 4.27. (★) (Tests unilaterales) El test de suma de rangos de Wilcoxon se puede utilizar para contrastar la hipótesis unilateral de que los datos de la primera muestra tienen tendencia a ser menores que los de la segunda, o viceversa. Para ello reemplazaremos α por 2α en la forma habitual. Así, para contrastar la hipótesis nula unilateral de que la primera variable tiene tendencia a tomar valores menores que la segunda al nivel α , rechazaremos dicha hipótesis si $R \geq v_s(2\alpha)$ en el caso $m + n \leq 30$, y, en el caso $m + n > 30$, si

$$\frac{R - E(R)}{\sqrt{\text{Var}(R)}} > z_{2\alpha}$$

Para contrastar la hipótesis nula unilateral de que la primera variable tiene tendencia a tomar valores mayores que la segunda al nivel α , rechazaremos dicha hipótesis si $R \leq v_i(2\alpha)$ en el caso $m + n \leq 30$, y si

$$\frac{R - E(R)}{\sqrt{\text{Var}(R)}} < -z_{2\alpha}$$

en el caso $m + n > 30$. \triangleleft

Observación 4.28. (★) (Corrección por continuidad) En realidad, se suele utilizar también en este caso una corrección por continuidad debido a que la distribución de R es discreta y se está aproximando por una distribución continua como es la normal; la región crítica quedaría expresada en la forma

$$(|R - E(R)| - 0.5) / \sqrt{\text{Var}(R)} > z_\alpha$$

aunque los argumentos utilizados en una observación de la página 169 son también aplicados aquí para justificar la supresión de esa corrección. \triangleleft

4.2.6. Test t de Student de comparación de dos medias en el caso de muestras relacionadas

Supongamos que elegimos n individuos alcohólicos al azar y que medimos sus presiones sanguíneas sistólicas antes (variable X) y después (variable Y) de dos meses de haber dejado la bebida. Se desea contrastar la hipótesis nula de que no varían las presiones medias antes

y después de dejar el alcohol contra la hipótesis bilateral de que sí varían. A partir de los n pares (x_i, y_i) , $1 \leq i \leq n$, obtenidos calcularemos los valores diferencia $d_i = x_i - y_i$, $1 \leq i \leq n$, que serán resumidos en su media \bar{D} y en su varianza muestral S_D^2 . La media de la variable diferencia $D = X - Y$ es $\mu_1 - \mu_2$; supondremos que la diferencia D tiene distribución normal con media $\mu_D = \mu_1 - \mu_2$ y varianza σ_D^2 . La hipótesis a contrastar $H_0 : \mu_1 = \mu_2$ se puede escribir también en la forma $H_0 : \mu_D = 0$ con lo que, considerando la variable diferencia, el problema queda reducido al de contrastar la hipótesis de que la media de una distribución normal de varianza desconocida es nula, problema que ya fue resuelto en el tema anterior. El test se basará en el estadístico de contraste

$$t = \frac{\bar{D}}{S_D/\sqrt{n}} = \sqrt{n} \frac{\bar{D}}{S_D}$$

donde S_D^2 es la varianza muestral calculada a partir de d_1, \dots, d_n , y la región crítica al nivel α será

$$|t| > t_\alpha(n-1)$$

Éste es el llamado test de Student para muestras relacionadas. El caso unilateral se trata de la forma habitual. Un intervalo de confianza para la diferencia de medias es

$$\mu_D = \mu_1 - \mu_2 \in \bar{D} \pm t_\alpha(n-1)S_D/\sqrt{n}$$

Observación 4.29. (Caso no normal) Si la hipótesis de normalidad de la variable diferencia D no se verifica, pero el número de pares es mayor o igual que 60, podemos asumir que \bar{D} tiene distribución aproximadamente normal y todo lo dicho es aproximadamente válido. ◀

4.2.7. (★) Test no paramétrico de rangos con signo de Wilcoxon de comparación de dos muestras relacionadas

Dadas m parejas de datos como en la sección anterior, si no podemos asumir normalidad de las diferencias, se puede utilizar el test de rangos con signo de Wilcoxon. Para ello se obtienen las m diferencias entre las componentes primera y segunda de cada par, se rechazan

las diferencias nulas (supongamos que hay n diferencias no nulas) y se ordenan de menor a mayor los valores absolutos de las diferencias restantes; se asignan después rangos a las diferencias no nulas en la forma habitual y se calculan las sumas $R(+)$ y $R(-)$ de los rangos correspondientes a las diferencias positivas y negativas, respectivamente. Nótese que basta calcular una de esas sumas, pues $R(+) + R(-) = n(n+1)/2$.

Supongamos primero que $n \leq 100$. Para el test bilateral, rechazaremos la hipótesis nula si $R(+)$ o $R(-)$ es menor o igual que el valor crítico bilateral $T_\alpha(n)$ que proporciona la tabla IX de la distribución T de Wilcoxon para muestras relacionadas. En el problema unilateral de contrastar la hipótesis nula de que los valores de la variable X son menores o iguales que los de la variable Y contra que aquellos son mayores que estos, rechazaremos la hipótesis nula si $R(-) \leq T_{2\alpha}(n)$.

En el problema unilateral de contrastar la hipótesis nula de que los valores de la variable X son mayores o iguales que los de la variable Y contra que aquellos son menores que estos, rechazaremos la hipótesis nula si $R(+) \leq T_{2\alpha}(n)$.

Si $n > 100$, utilizaremos que $R(+)$ (el resultado sería exactamente el mismo si utilizamos $R(-)$ en lugar de $R(+)$) tiene aproximadamente distribución normal de media $E(R(+)) = n(n+1)/4$ y varianza

$$\text{Var}(R(+)) = n(n+1)(2n+1)/24$$

si no hay empates, y

$$\text{Var}(R(+)) = \frac{2n(n+1)(2n+1) - \sum_{i=1}^r T_i}{48}$$

si hay r grupos de empates con t_i datos en el i -ésimo grupo de empates, donde $T_i = t_i^3 - t_i$.

Se compara el estadístico

$$Z = \frac{|R(+)-E(R(+))|}{\sqrt{\text{Var}(R(+))}}$$

con z_α en el caso bilateral; en el caso unilateral se compara el valor de Z con $z_{2\alpha}$ en la forma habitual.

Observación 4.30. (★) (Corrección por continuidad) También ahora deberíamos incluir una corrección por continuidad (no lo haremos por el motivo expuesto en la observación de la página 169) en el estadístico de contraste, que quedaría expresado en la forma

$$Z = \frac{|R(+)-E(R(+))| - 0.5}{\sqrt{\text{Var}(R(+))}}$$

△

Ejemplo 4.31. La tabla siguiente contiene los niveles de colesterol de catorce hombres mayores de 30 años tomados antes y dos horas después de desayunar.

Antes	180	195	175	190	192	185	188
Después	180	185	190	210	192	200	208
Antes	184	183	187	192	198	193	210
Después	202	203	182	207	193	199	219

Las diferencias entre los datos antes y después son 0, 10, -15, -20, 0, -15, -20, -18, -20, 5, -15, 5, -6 y -9. Escribiendo las 12 diferencias no nulas en orden creciente de sus valores absolutos obtenemos

$$5, 5, -6, -9, 10, -15, -15, -15, -18, -20, -20, -20$$

Entonces $R(+)$ = 1.5 + 1.5 + 5 = 8 y $R(-)$ = 3 + 4 + 7 + 7 + 7 + 9 + 11 + 11 + 11 = 70. En la tabla IX se observa que $T_{0.05}(12)$ = 13. Puesto que $R(+)$ ≤ 13, debemos admitir que hay una diferencia significativa entre los niveles de colesterol antes y después del desayuno. Puesto que $T_{0.01}(12)$ = 9 y $T_{0.001}(12)$ = 4, esa diferencia es muy significativa ($0.001 < P < 0.01$). Si a esos datos aplicamos el test t de Student para muestras relacionadas obtenemos un valor P entre 0.01 y 0.02. ◁

4.2.8. Comparación de dos proporciones: muestras independientes

En esta sección consideramos dos observaciones x_1 y x_2 de sendas variables X_1 y X_2 independientes con distribuciones respectivas $B(n_1, p_1)$ y $B(n_2, p_2)$ de parámetros p_1 y p_2 desconocidos, y pretendemos contrastar la hipótesis nula de que esos dos parámetros son iguales. Consideremos un ejemplo concreto.

Ejemplo 4.32. En dos centros de salud A y B de la ciudad de Badajoz se han seleccionado al azar $n_1 = 173$ y $n_2 = 86$ individuos varones mayores de 14 años, respectivamente, a los que se ha pasado un cuestionario diseñado por Altisent y otros con el fin de decidir quiénes de ellos son bebedores de riesgo, habiéndose obtenido que $x_1 = 50$ del centro de salud (CS) A y $x_2 = 10$ del centro de salud B lo son. Nos preguntamos si estos datos son compatibles con la hipótesis nula de que la proporción de bebedores de riesgo es la misma en ambos centros de salud. Las proporciones de bebedores de riesgo para los centros A y B son parámetros desconocidos p_1 y p_2 (de los que lo único que conocemos a priori es que son números entre 0 y 1). x_1 (respectivamente, x_2) se considera una realización concreta de una v.a. X_1

(respectivamente, X_2) con distribución binomial $B(173, p_1)$ (respectivamente, $B(86, p_2)$). Las proporciones muestrales son $\hat{p}_1 = x_1/n_1 = 50/173 = 0.29$ y $\hat{p}_2 = 10/86 = 0.12$, y cabe preguntarse si ambas son distintas por azar o porque en un centro de salud hay más bebedores de riesgo que en el otro. Conviene poner los datos en forma de una tabla 2×2 como sigue:

	Bebedor de riesgo	No bebedor de riesgo	Totales
CS A	$x_1 = 50$	$n_1 - x_1 = 123$	$n_1 = 173$
CS B	$x_2 = 10$	$n_2 - x_2 = 76$	$n_2 = 86$
Totales	$x_1 + x_2 = 60$	$n_1 + n_2 - x_1 - x_2 = 199$	$N = 259$

Tabla 2×2 ◁

Consideraremos el problema bilateral de contrastar la hipótesis nula $H_0 : p_1 = p_2$ contra la alternativa $H_1 : p_1 \neq p_2$ en el caso de que la aproximación a la normal es válida. En ese caso sabemos que \hat{p}_1 tiene distribución aproximada $N(p_1, p_1q_1/n_1)$ y que \hat{p}_2 tiene distribución aproximada $N(p_2, p_2q_2/n_2)$, donde $q_i = 1 - p_i$, $i = 1, 2$. La independencia de las muestras prueba que la distribución aproximada de $\hat{p}_1 - \hat{p}_2$ es

$$N\left(p_1 - p_2, \frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}\right)$$

con lo cual, bajo $H_0 : p_1 = p_2$, $\hat{p}_1 - \hat{p}_2$ tiene distribución aproximada

$$N(0, pq(1/n_1 + 1/n_2))$$

donde $p = p_1 = p_2$ y $q = 1 - p$. Por analogía con lo dicho para el caso de una muestra de la distribución normal, para comparar p_1 con p_2 cabe pensar en la cantidad $|\hat{p}_1 - \hat{p}_2|/\sqrt{pq(1/n_1 + 1/n_2)}$. Siendo el parámetro p desconocido, lo sustituiremos por una estimación del mismo. Bajo la hipótesis nula, disponemos de dos estimaciones posibles, \hat{p}_1 y \hat{p}_2 , para p . Para no perder información, será preferible usar una media ponderada de ambos, es decir,

$$\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{x_1 + x_2}{n_1 + n_2}$$

Así pues, la región crítica del test es:

$$\frac{|\hat{p}_1 - \hat{p}_2|}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} > z_\alpha$$

Observación 4.33. (★) (Corrección por continuidad) Teniendo en cuenta la corrección por continuidad (que no consideraremos por los motivos expuestos en una observación de la página 169), la región crítica se expresaría en la forma

$$\frac{|\hat{p}_1 - \hat{p}_2| - \frac{1}{2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}{\sqrt{\hat{p}\hat{q} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} > z_\alpha$$

△

El test obtenido es doblemente aproximado, pues se usa la aproximación de la binomial por la normal y la sustitución de p por su estimación \hat{p} , y, sin embargo, el test funciona bastante bien en la práctica. Las condiciones de validez del test son las siguientes: $n_1\hat{p}\hat{q}, n_2\hat{p}\hat{q} \geq 5$.

Por analogía con lo dicho en otras ocasiones, un intervalo de confianza para la diferencia de proporciones es

$$p_1 - p_2 \in (\hat{p}_1 - \hat{p}_2) \pm z_\alpha \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

Se exigen las condiciones de validez $n_1\hat{p}_1\hat{q}_1, n_2\hat{p}_2\hat{q}_2 \geq 5$ para que la aproximación por la normal sea aceptable.

Ejemplo 4.34. (Ejemplo 4.32, continuación) Aplicado el test al ejemplo considerado, se obtiene significación ($P = 0.0026$), es decir, los datos no son compatibles con la hipótesis nula de que las proporciones de bebedores de riesgo son iguales en ambos centros de salud. Por otra parte, un intervalo al 95 % de confianza para la diferencia de proporciones es $[0.07, 0.27]$. Se comprueba que se verifican las condiciones de validez exigidas en ambos casos. ◁

4.2.9. Test de McNemar de comparación de dos proporciones: muestras relacionadas

Consideramos ahora el problema de comparación de dos proporciones relacionadas. Este problema consiste en lo siguiente: se eligen al azar n individuos de la población y se exige de cada uno de ellos un par de respuestas entre dos valores posibles, que pueden ser 0 y 1, o SÍ y NO, etc. Se trata de decidir si los porcentajes de respuestas de un tipo (0, pongamos por caso) coinciden para las dos cuestiones formuladas. Veamos un ejemplo aclaratorio.

Ejemplo 4.35. Con el fin de comparar las preferencias de los individuos de una población sobre la sanidad pública y la sanidad privada, se diseña el siguiente experimento relacionado: se seleccionan al azar 125 individuos de la población y se les invita a pronunciarse a favor (SÍ) o en contra (NO) de cada uno de esos dos modelos sanitarios. Se obtienen de ese modo 125 pares de síes y/o noes. Como sólo hay cuatro tipos de respuestas posibles ((SÍ,SÍ),(SÍ,NO),(NO,SÍ),(NO,NO)) es más apropiado recoger los datos como se hace en la tabla siguiente:

		Sanidad Pública		
		SÍ	NO	Totales
Sanidad Privada	SÍ	$n_{11} = 27$	$n_{12} = 35$	62
	NO	$n_{21} = 43$	$n_{22} = 20$	63
Totales		70	55	125

Tabla. Entrada de datos para un test de McNemar <

Si p_1 y p_2 denotan las proporciones poblacionales (desconocidas) de respuestas SÍ a la primera cuestión (a favor de la sanidad privada, en el ejemplo) y a la segunda (a favor de la sanidad pública, en el ejemplo) respectivamente, se trata de contrastar la hipótesis nula $H_0 : p_1 = p_2$ contra $H_1 : p_1 \neq p_2$. Pero p_1 es suma de p_{11} (proporción de respuestas SÍ-SÍ) y p_{12} (proporción de respuestas SÍ-NO), y p_2 es suma de p_{11} (proporción de respuestas SÍ-SÍ) y p_{21} (proporción de respuestas NO-SÍ). Entonces $p_1 - p_2 = p_{12} - p_{21}$, y el problema anterior es equivalente al de contrastar $H'_0 : p_{12} = p_{21}$ contra $H'_1 : p_{12} \neq p_{21}$. Esas proporciones no contemplan a los individuos con respuesta homogénea (SÍ-SÍ o NO-NO), los cuales no ayudan a distinguir entre los dos tipos de atención sanitaria considerados. Retendremos por tanto, los $n_{12} + n_{21}$ individuos con respuesta dispar. Denotemos por p la probabilidad de que un individuo con respuesta dispar sea del tipo SÍ-NO, y por X_{12} la v.a. que describe el número de individuos SÍ-NO entre los que tienen respuesta dispar; entonces X_{12} tiene distribución $B(n_{12} + n_{21}, p)$. Pero si $H'_0 : p_{12} = p_{21}$ es cierta, será $p = 1/2$. Por tanto, el test anterior es equivalente al test para una proporción $H''_0 : p = 1/2$ contra $H''_1 : p \neq 1/2$. En el caso de que la aproximación a la normal sea aceptable (exigiremos para ello que $n_{12} + n_{21} \geq 20$), y la región crítica al nivel α del test viene a ser

$$\frac{\left|n_{12} - \frac{n_{12} + n_{21}}{2}\right|}{\sqrt{(n_{12} + n_{21})(1/2)(1/2)}} > z_\alpha$$

o, multiplicando numerador y denominador del estadístico de contraste por 2,

$$\frac{|n_{12} - n_{21}|}{\sqrt{n_{12} + n_{21}}} > z_{\alpha}$$

El test recibe el nombre de test de McNemar. Compruébese que, en nuestro ejemplo, no hay evidencias de que un modelo sanitario sea preferido al otro en la población de referencia.

Observación 4.36. (★) (Corrección por continuidad) Con la correspondiente corrección de continuidad, la región crítica se escribiría en la forma

$$\frac{|n_{12} - n_{21}| - 1}{\sqrt{n_{12} + n_{21}}} > z_{\alpha}$$

pero los motivos expuestos en una observación de la página 169 también son considerados aquí. ◁

Un intervalo de confianza usado habitualmente para la diferencia de proporciones en el caso relacionado es:

$$p_1 - p_2 \in \frac{1}{n} \left[(n_{12} - n_{21}) \pm z_{\alpha} \sqrt{(n_{12} + n_{21}) - \frac{(n_{12} - n_{21})^2}{n}} \right]$$

válido si $n_{12}, n_{21} > 5$. En el ejemplo, el intervalo al 95% de confianza que se obtiene es $[-0.2, 0.07]$.

¿Verdadero o Falso? Capítulos 3 y 4

Decidir si cada una de las proposiciones siguientes es verdadera (V) o falsa (F), justificando la respuesta.

- 3-4.1** En un test de hipótesis, un error de tipo I se comete cuando se rechaza incorrectamente la hipótesis nula.
- 3-4.2** En un contraste sobre la media de una distribución normal de varianza conocida, si la varianza aumenta el valor P del test aumenta.
- 3-4.3** Supongamos conocido que para contrastar la hipótesis $\mu = \mu_0$ contra $\mu \neq \mu_0$ a partir de una muestra de una distribución normal se ha obtenido un nivel mínimo de significación $P = 0.07$. Entonces, el nivel mínimo de significación para el problema de contrastar $\mu \leq \mu_0$ contra $\mu > \mu_0$ es igual a 0.035.
- 3-4.4** Si realizamos un test de hipótesis a un nivel $\alpha \in]0, 1[$ y la decisión es aceptar la hipótesis nula H_0 , podemos asegurar con una confianza del $(1 - \alpha) \cdot 100\%$ que la decisión tomada es correcta.
- 3-4.5** Si deseamos realizar un test de hipótesis a un nivel $\alpha \in]0, 1[$ y conocemos el nivel mínimo de significación P , aceptaremos la hipótesis nula al nivel α si $\alpha < P$.
- 3-4.6** En un contraste sobre la media de una distribución normal de varianza conocida, si el tamaño muestral aumenta y la media muestral permanece constante, el nivel mínimo de significación del test disminuye.
- 3-4.7** En un test de hipótesis, un error de tipo II se comete cuando se acepta incorrectamente la hipótesis alternativa.
- 3-4.8** La longitud del intervalo de confianza para la media de una distribución normal aumenta con el tamaño de la muestra.
- 3-4.9** Un estadístico es una función de las observaciones que, eventualmente, puede depender también del parámetro.

- 3-4.10** La longitud del intervalo de confianza para la media de una distribución normal de varianza conocida decrece a medida que aumenta la varianza.
- 3-4.11** En estadística, una hipótesis es una afirmación sobre la distribución desconocida de la variable que observamos.
- 3-4.12** Diferentes estructuras estadísticas para un mismo problema real y unos mismos datos conducen siempre a una misma decisión estadística.
- 3-4.13** Un estimador de un parámetro desconocido es una función de los datos y del parámetro en cuestión.
- 3-4.14** El nivel mínimo de significación de un test es la probabilidad de obtener un valor de la variable que observamos tan extremo o más que el realmente obtenido supuesta cierta la hipótesis nula.
- 3-4.15** Se observa una muestra de tamaño n de una distribución normal con varianza conocida. Cualquier diferencia, por pequeña que sea, entre la media muestral y el valor de la media propuesto por la hipótesis nula puede terminar siendo significativa si n es lo suficientemente grande (y el valor de la media muestral permanece constante).
- 3-4.16** Si un test de hipótesis rechaza la hipótesis nula, queda demostrado que esta es falsa.
- 3-4.17** Si en el problema de comparación de dos medias de dos distribuciones normales la hipótesis nula es $H_0 : \mu_1 \leq \mu_2$, la hipótesis nula se rechazará siempre que la media muestral de la primera muestra sea estrictamente mayor que la de la segunda.
- 3-4.18** La longitud del intervalo de confianza para el cociente de dos varianzas de dos distribuciones normales en el caso de muestras independientes aumenta con el nivel de confianza $1 - \alpha$ deseado.
- 3-4.19** Con el objetivo de comparar dos medicamentos A y B para el tratamiento de una enfermedad, se seleccionan al azar 100 enfermos de los cuales 60 reciben el medicamento A y los 40 restantes el B. La comparación se llevará a cabo a partir del número

de enfermos que sanan en cada caso. Se trata, en realidad, de un problema de comparación de dos medias en el caso de muestras independientes.

- 3-4.20** Para comparar las presiones sanguíneas de los brazos izquierdo y derecho de las personas de una cierta población es posible y preferible un experimento relacionado.
- 3-4.21** (★) El estadístico de contraste del test de Mann-Whitney-Wilcoxon de comparación de dos muestras independientes no utiliza exactamente los valores de los datos obtenidos: sólo hace uso de las posiciones que ocupan los datos de la primera muestra en la muestra combinada y reordenada.
- 3-4.22** Un test de McNemar no significativo a un nivel de significación $\alpha \in]0, 1[$, puede ser significativo al nivel de significación $\alpha/2$.
- 3-4.23** Si el test de comparación de dos varianzas de dos distribuciones normales resulta significativo a un nivel de significación $\alpha \in]0, 1[$, entonces sigue siendo significativo para cualquier nivel de significación superior a α .
- 3-4.24** (★) En el test de Wilcoxon de homogeneidad de dos muestras relacionadas, la varianza del estadístico $R(+)$ (suma de los rangos de las diferencias positivas en la muestra reordenada de los valores absolutos de las diferencias no nulas) disminuye a medida que aumenta el número de empates.
- 3-4.25** Supongamos que en un test de McNemar se tiene que $n_{12} > n_{21}$, y que n_{12} (número de respuestas (SI,NO)) aumenta en la misma cantidad en que disminuye n_{21} (número de respuestas (NO,SI)). Entonces el valor P del test también aumenta.
- 3-4.26** (★) En un problema de comparación de dos muestras relacionadas en ausencia de normalidad, la suma de los rangos de las diferencias positivas del test de Wilcoxon toma el valor 25. Sabiendo que el número de diferencias no nulas es igual a 18, el test resulta no significativo.
- 3-4.27** Cuanto mayor es la longitud del intervalo de confianza sobre un parámetro real desconocido de una distribución, menor es la información que proporciona sobre el mismo.

Problemas de los Capítulos 3 y 4

3-4.1 Se desea decidir si existen diferencias entre las presiones sanguíneas de las personas cuando están sentadas en una silla y cuando están tumbadas en una cama.

a) ¿Qué diseño consideras preferible para tomar esa decisión: el de muestras independientes o el de muestras relacionadas? Justificar la respuesta.

b) Toma finalmente la decisión asumiendo normalidad y teniendo en cuenta que, tras elegir diez personas al azar y determinar sus presiones sanguíneas en ambas posiciones, se han obtenido los siguientes resultados:

Persona	1	2	3	4	5	6	7	8	9	10
Sentado	142	100	112	92	104	100	108	94	104	98
Tumbado	154	100	110	100	112	100	120	90	104	114

3-4.2 La anemia por deficiencia de hierro es un importante problema de salud nutricional. Un estudio ha sido realizado con 51 chicos de entre 9 y 12 años cuyas familias están por debajo del nivel de pobreza. El consumo medio diario de hierro obtenido para esos chicos es de 12.50 mg con una desviación estándar de 4.75 mg. Supongamos conocido que el consumo medio diario de hierro en una población más grande de chicos de entre 9 y 12 años es de 14.44 mg.

a) Contrasta al nivel $\alpha = 0.05$ la hipótesis de que el consumo medio de hierro en el grupo de bajo consumo es menor que el de la población general, aclarando las suposiciones que se hacen para resolver esa cuestión.

b) ¿Cuál es el valor P para ese test?

3-4.3 Se formula la hipótesis de que la probabilidad de que una mujer haya tenido su primer hijo con 30 o más años es superior entre las mujeres que han desarrollado un cáncer de mama que entre las que no lo han desarrollado, lo que vendría a establecer una asociación positiva entre el hecho de desarrollar un cáncer de mama y la edad con que la mujer tiene su primer hijo. Se ha realizado un estudio internacional para contrastar esa hipótesis y se han obtenido los siguientes datos: 683 de un total de 3220 mujeres

con cáncer de mama seleccionadas al azar tuvieron su primer hijo con una edad ≥ 30 , y 1498 de un total de 10245 mujeres sin cáncer de mama seleccionadas al azar tuvieron su primer hijo con una edad ≥ 30 . ¿Avalan estos datos la hipótesis formulada? Realiza un test apropiado para justificar la respuesta. **SOLUCIÓN: PÁG. 342**

- 3-4.4** Con el objeto de estudiar la posible influencia del uso de anticonceptivos en la presión sanguínea sistólica (PSS) se seleccionan al azar 10 mujeres y se determinan sus PSS un mes antes de empezar a tomar anticonceptivos (variable X) y seis meses después de ello (variable Y), obteniéndose los siguientes datos:

Mujer núm.	1	2	3	4	5	6	7	8	9	10
PSS Antes (X)	115	112	107	119	115	138	126	105	104	115
PSS Después (Y)	128	115	106	128	122	145	132	109	102	117

- a) ¿Puede suponerse que las PSS “antes” siguen una distribución normal? Justifica la respuesta aplicando un test apropiado añadiendo el valor p .
- b) Asumiendo normalidad, ¿contienen esos datos evidencia de que el tratamiento influye realmente en la PSS? Presenta un intervalo al 95 % de confianza para la diferencia de PSS medias “antes” y “después”.
- 3-4.5** Con el objeto de estudiar la relación entre el uso de anticonceptivos y la malformación del niño al nacer se llevó a cabo un estudio caso-control seleccionando 1430 madres de niños malformados; para cada una de esas mujeres (casos) se seleccionó como control una madre de un niño sin malformaciones de la misma edad maternal, obteniendo de este modo 1430 pares de mujeres en las que la primera es madre de un niño con malformaciones y la segunda tuvo un hijo sin malformaciones a la misma edad que la primera. Todas esas mujeres respondieron a la cuestión de si habían tomado anticonceptivos antes de la concepción del niño en cuestión. Los resultados obtenidos se recogen en la siguiente tabla

Control

		SI	NO
Caso	SI	39	154
	NO	138	1099

¿Existe alguna relación entre el uso de anticonceptivos y la malformación de los niños al nacer?

- 3-4.6** Se desean comparar dos tratamientos quirúrgicos para el cáncer de mama: mastectomía simple (S) y mastectomía radical (R). Se han formado pares de mujeres con cáncer de mama dentro de la misma década y con similares condiciones clínicas, y cada una de ellas ha recibido un tratamiento diferente. Se ha monitorizado su supervivencia dentro de los 5 años siguientes, habiéndose obtenido los resultados siguientes (V=vivió al menos 5 años, F=falleció antes de los cinco años):

Par	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
S	F	V	V	V	V	F	V	V	V	V	F	V	V	V	V
R	V	F	V	V	F	V	V	F	F	F	F	F	V	F	F
Par	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
S	F	V	V	V	V	V	V	V	F	V	V	V	V	V	V
R	F	F	V	F	F	V	F	F	V	F	V	F	F	V	F

¿Qué test deberíamos utilizar para analizar estos datos? ¿Qué conclusiones nos proporciona ese test sobre los dos tratamientos comparados? **SOLUCIÓN: PÁG. 342**

- 3-4.7** La tabla siguiente contiene determinaciones en útero de catalasa (CAT) y malondialdehído (MDA) de 14 ratas.

Rata	1	2	3	4	5	6	7
CAT	51.6	48.7	42.6	38	50.1	60.2	41.3
MDA	24.8	25	23.8	23.5	27.2	21.3	24.2
Rata	8	9	10	11	12	13	14
CAT	47.6	40.6	49.5	42	55.1	46.2	52.1
MDA	22.8	22.9	20.6	24	24.7	23.2	25.1

- a) Calcula media, mediana y varianza de cada una de las variables CAT y MDA.
 b) Usa el test de D'Agostino para contrastar su normalidad.

3-4.8 Una muestra aleatoria de 50 personas presenta un contenido medio de glucosa en suero de 100 mg/dl. Supongamos conocido que el contenido de glucosa sigue una distribución normal con varianza conocida $\sigma^2 = 100$.

- a) Contrasta la hipótesis nula $H_0 : \mu \leq 90$ contra $\mu > 90$.
 b) Determina un intervalo al 95 % de confianza para la media poblacional.
 c) Determina el tamaño de muestra necesario para reducir la longitud del intervalo de confianza al 95 % a la tercera parte.

3-4.9 La tabla siguiente contiene las presiones sanguíneas sistólicas (PSS) de un grupo de ratas normotensas y otro de hipertensas.

Normotensas	114	104	97	102	108	112	100	121	99	109
	111	115	114	105	104	107				
Hipertensas	128	135	130	120	126	140	133	132	139	150

- a) ¿Podemos asumir normalidad en ambos grupos?
 b) ¿Podemos asumir la igualdad de varianzas de los dos grupos?
 c) Asumiendo normalidad e igualdad de varianzas, ¿son las PSS de los dos grupos de ratas significativamente diferentes?
 d) Proporciona un intervalo al 99 % de confianza para la diferencia de medias.
 e) (★) ¿Un test no paramétrico modificaría la decisión tomada en el apartado c)?

SOLUCIÓN: PÁG. 342

3-4.10 Compara las calificaciones en los grupos $S = 0$ y $S = 1$ de la tabla 0.1 (página 12).

3-4.11 Se ha llevado a cabo un estudio dental en niños de 4 años en dos centros de salud de una ciudad, uno situado en el centro de la misma y otro en una zona suburbana. Se han tomado dos muestras de tamaños 110 y 85, respectivamente, de esas poblaciones

habiéndose observado que 90 entre los del primer grupo y 80 de los del segundo no poseían ninguna caries.

- a) ¿Son significativas las diferencias observadas entre los dos grupos?
- b) Facilita un intervalo de confianza para la diferencia de proporciones.

3-4.12 En un estudio sobre 193 pares de mellizos se observaron 56 casos en los que ambos eran varones y 65 casos en que ambos eran hembras.

(a) ¿Avalan esos datos la hipótesis de que observar una pareja de distinto sexo en parejas de mellizos es menos probable que observar una pareja del mismo sexo? Facilita en este caso un intervalo de confianza adecuado al resultado obtenido.

(b) Entre las parejas de mellizos del mismo sexo, ¿alguno de los dos sexos es más frecuente que el otro? Comenta el P -valor del test utilizado. **SOLUCIÓN: PÁG. 342**

3-4.13 Este problema está basado en la tabla 0.24 del tema 1. Asumiremos que las calificaciones de los 100 alumnos considerados provienen de una distribución normal.

a) Haz uso de una tabla de números aleatorios para extraer una muestra de tamaño 15 de la clase.

b) A partir de los datos de esa muestra, estima la media y la desviación típica de la variable N suponiendo que ésta sigue una distribución normal, y proporciona intervalos a 95 % de confianza para esos parámetros.

c) Estima a partir de esa misma muestra las proporciones de hombres y mujeres en la clase y proporciona intervalos al 95 % de confianza para esos parámetros.

d) ¿Sostienen esos 15 datos la hipótesis de que la calificación media es 6.13? ¿Y la de que la desviación típica es 2.14?

e) Compara las calificaciones medias de chicos y chicas en ese examen parcial y proporciona un intervalo de confianza para la diferencia de medias.

f) Contrasta la hipótesis de normalidad de las calificaciones de esos alumnos.

g) Da una estimación del porcentaje de aprobados en ese primer parcial.

3-4.14 Para comparar dos antibióticos A y B en el tratamiento de una cierta enfermedad, se seleccionan 79 parejas de personas con esa enfermedad procurando que los dos componentes de un mismo par sean similares en dos aspectos que se han considerado relevantes para el estudio: no se diferencien en más de cinco años y tengan el mismo sexo. Tras una semana de tratamiento se obtuvieron los siguientes datos: para 40 parejas ambos tratamientos fueron exitosos, para 20 parejas el tratamiento A fue efectivo pero el B no, para 16 parejas fue efectivo el B pero no el A, y para las 3 restantes no fueron efectivos ninguno de los dos tratamientos.

- a) Contrasta la hipótesis nula de que ambos tratamientos son igualmente efectivos.
- b) Proporciona un intervalo de confianza para completar la conclusión del test anterior.

3-4.15 De 65 muertes ocurridas entre trabajadores varones de centrales nucleares con edad comprendida entre 55 y 64 años de edad, 25 fueron debidas a algún tipo de cáncer. (a) Estima la proporción de muertes por cáncer para ese tipo de trabajadores y proporciona un intervalo al 95 % de confianza para ella; interpreta los resultados obtenidos. (b) Si es conocido, por las estadísticas de mortalidad de la población general, que el 20 % de las muertes de varones entre 55 y 64 años de edad son atribuibles al cáncer, ¿podemos afirmar que el hecho de trabajar en una central nuclear incrementa el porcentaje de muertes por cáncer? Justifica la respuesta realizando un test apropiado y comentando la conclusión. **SOLUCIÓN: PÁG. 342**

3-4.16 Se lleva a cabo un estudio para contrastar el efecto del etanol en la dilatación de los vasos sanguíneos de la retina tras la aplicación de metanol. Se aplica metanol (1 g/Kg) por vía nasogástrica a 9 caballos y se cuentan las venas de la retina que se han dilatado. Dos semanas después se repite el experimento con la diferencia de que, 30 minutos después de la aplicación del metanol, se aplica una dosis oral de etanol (1.4 g/Kg). La tabla siguiente recoge los datos del estudio. Aplica un test paramétrico para decidir si el etanol produce algún efecto en el número de venas retinales dilatadas. Facilita también un intervalo al 95 % de confianza adecuado y comenta las conclusiones del

estudio.

Núm. caballo	1	2	3	4	5	6	7	8	9
Metanol	10	9	7	10	7	4	8	4	10
Metanol+etanol	6	5	5	1	6	7	5	5	7

Capítulo **5**

Comparación de Varias Muestras

5.1. Comparación de Varias Muestras Cuantitativas: Modelo de Clasificación Simple en Análisis de la Varianza

5.1.1. Comparación de varias medias: muestras independientes y varianzas iguales

Consideremos un ejemplo que nos ayude a introducir el que llamaremos modelo de clasificación simple en análisis de la varianza, o modelo de análisis de la varianza de una vía.

Ejemplo 5.1. Se sospecha que cierto tipo de dolores tras una pequeña intervención quirúrgica pueden aliviarse con sólo administrar un medicamento, potente o no. Un grupo de 15 pacientes a los que se acaba de extirpar las amígdalas son divididos en 3 grupos de 5 individuos: al primero de ellos se le aplica un analgésico A, a otro un analgésico B y al tercero un placebo. Se determina el tiempo en horas hasta que el paciente se queja de dolor. La tabla siguiente contiene los resultados de la experiencia:

Placebo	2.2	0.3	1.1	2.0	3.4
A	2.8	1.4	1.7	4.3	4.3
B	1.1	4.2	3.8	2.6	0.5

Se pretende decidir si el tiempo medio transcurrido es el mismo para los 3 tratamientos considerados o existen diferencias significativas entre al menos dos de ellos. \triangleleft

En general, el problema es el siguiente: se supone que estamos interesados en comparar el efecto producido por k tratamientos sobre el valor medio μ (desconocido desde el punto de vista estadístico) de una cierta característica numérica de una población; la media μ quedaría afectada por el tratamiento aplicado, pasando a ser un valor μ_i tras haber recibido el i -ésimo tratamiento, $1 \leq i \leq k$. Estas nuevas medias poblacionales son parámetros desconocidos pero, en principio, diferentes unos de otros. Precisamente, el objetivo del

problema es contrastar la hipótesis nula

$$H_0 : \mu_1 = \cdots = \mu_k$$

contra la alternativa

$$H_1 : \mu_i \neq \mu_j, \text{ para algún } i \text{ y } j$$

de que al menos dos de las medias son diferentes. Rechazar la hipótesis nula indicaría que no todos los tratamientos considerados son iguales, en tanto que una decisión por la hipótesis nula nos indicaría que, a la luz de los datos, no es posible rechazar la hipótesis de homogeneidad de las medias. Con ese objetivo se eligen al azar N individuos de la población que se dividen aleatoriamente en k grupos de tamaños n_1, \dots, n_k (que obviamente deben sumar N), aplicando a cada subgrupo un tratamiento diferente.

Observación 5.2. Los términos una vía o clasificación simple se refieren al hecho de que sólo se estudia la influencia sobre la media de un factor que actúa a k niveles diferentes (en el ejemplo, el único factor que se considera en el estudio es el tratamiento contra el dolor aplicado, factor que actúa a tres niveles; ningún otro factor, tal como las condiciones en que se realiza la intervención, sexo u otras que pudieran parecer relevantes para el problema, es tenido en cuenta). El modelo de clasificación simple en análisis de la varianza se conoce también como diseño completamente aleatorio con efectos fijos. El término diseño completamente aleatorio se refiere al hecho de que no se ha hecho ningún intento de aparear unidades experimentales de las distintas muestras. El término efectos fijos alude al hecho de que el experimentador selecciona específicamente los niveles del factor implicado, pues son estos los niveles que interesan, y no se seleccionan al azar de un grupo más grande de niveles posibles. \triangleleft

Los datos obtenidos de la experimentación se denotarán por x_{ij} , $1 \leq j \leq n_i$, $1 \leq i \leq k$, cantidades que se suponen observaciones concretas de variables X_{ij} que supondremos independientes y normalmente distribuidas con varianza común σ^2 (varianza desconocida que se supone no afectada por el nivel del factor); se supone que X_{ij} tiene media μ_i (desconocida y dependiente del nivel i del factor). Podemos pensar también que se disponen de k muestras independientes

Nivel del factor	1	\cdots	k
Observaciones	$X_{11} \dots X_{1n_1}$	\cdots	$X_{k1} \dots X_{kn_k}$

Necesitaremos los siguientes estadísticos:

$$\bar{X}_{i.} = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij} = \text{media muestral para el nivel } i\text{-ésimo, } 1 \leq i \leq k$$

$$\bar{X}_{..} = \frac{1}{N} \sum_{i=1}^k \sum_{j=1}^{n_i} X_{ij} = \text{media muestral combinada}$$

Nótese que en las notaciones precedentes, el punto sustituye al índice sobre el que se calcula la media. La siguiente identidad (que llamaremos identidad de la suma de cuadrados para el modelo de clasificación simple), y que se prueba sin dificultad, es clave en la resolución del problema:

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_{..})^2 = \sum_{i=1}^k n_i (\bar{X}_{i.} - \bar{X}_{..})^2 + \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_{i.})^2$$

Cada uno de los términos de esa identidad tiene una interpretación práctica interesante. El primer término $\sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_{..})^2$ es una medida de la variabilidad de todos los datos alrededor de la media general $\bar{X}_{..}$ que llamaremos “variación total” y denotaremos por V_T . El sumando $\sum_{i=1}^k n_i (\bar{X}_{i.} - \bar{X}_{..})^2$ es una suma ponderada de las desviaciones de las medias de cada muestra respecto a la media general y se interpreta como la parte de la variación total debida a la diferencia entre los k tratamientos (niveles) considerados; se llamará “variación entre” y se denotará por V_e . Para cada $i \in \{1, \dots, k\}$, el sumando $\sum_{j=1}^{n_i} (X_{ij} - \bar{X}_{i.})^2$ mide la desviación de los valores observados en la muestra i -ésima respecto a su media $\bar{X}_{i.}$, o variación dentro de la muestra i -ésima, y el sumando $\sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_{i.})^2$ se llamará “variación dentro”, se denotará por V_d y se interpreta como la parte de la variación total debida a las fluctuaciones aleatorias entre sujetos dentro del mismo nivel del factor. Es esa descomposición de la variación total en dos componentes lo que justifica el nombre de análisis de la varianza. Con esas notaciones se tiene

$$V_T = V_e + V_d$$

Esas interpretaciones sugieren el uso del cociente V_e/V_d , pues compara la parte de la variación total debida a la diferencia entre tratamientos con la parte de la variación total debida al azar con el fin de decidir si el azar puede explicar también la diferencia observada

entre los tratamientos (en cuyo caso no hay motivos para rechazar la hipótesis nula de homogeneidad de las medias) o si, por el contrario, la diferencia observada entre los tratamientos es demasiado grande en comparación con la variación debida al azar (en cuyo caso habría que rechazar la hipótesis de homogeneidad de las medias). En lugar de ese cociente, basaremos nuestra decisión en el estadístico de contraste F múltiplo del anterior y cociente de las medias cuadráticas “entre” $mc_e = V_e/(k - 1)$ y “dentro” $S_d^2 = mc_d = V_d/(N - k)$ (que se obtienen dividiendo las variaciones correspondientes por sus grados de libertad); concretamente

$$F = \frac{V_e/(k - 1)}{V_d/(N - k)}$$

La distribución de F bajo la hipótesis nula de homogeneidad de las medias es la distribución $F(k - 1, N - k)$ de Fisher con $(k - 1, N - k)$ grados de libertad. La región crítica al nivel α para el test es

$$F > F_\alpha(k - 1, N - k)$$

donde el valor crítico $F_\alpha(k - 1, N - k)$ se obtiene de la tabla VI de la distribución de Fisher y es el punto que deja a su derecha una cola de área α en la distribución $F(k - 1, N - k)$.

La tabla siguiente resume los cálculos necesarios para llevar a cabo el análisis de la varianza:

Variación	Suma Cuadrados	G.L.	Media Cuadrática	F
Entre	$V_e = \sum_{i=1}^k n_i [\bar{x}_i - \bar{x}_{..}]^2$	$k - 1$	$mc_e = V_e/(k - 1)$	$F = mc_e/mc_d$
Dentro	$V_d = \sum_{i=1}^k \sum_{j=1}^{n_i} [x_{ij} - \bar{x}_i]^2$	$N - k$	$S_d^2 = V_d/(N - k)$	
Total	$V_T = \sum_i \sum_j [x_{ij} - \bar{x}_{..}]^2$	$N - 1$		

Tabla de análisis de la varianza

Observación 5.3. (Cálculo de V_T , V_d y V_e) A la hora de calcular expresamente las variaciones total, dentro y entre con una calculadora son más sencillas las siguientes expresiones:

$$V_T = \sum_{i=1}^k \sum_{j=1}^{n_i} x_{ij}^2 - \frac{1}{N} T_{..}^2$$

$$V_d = \sum_{i=1}^k \frac{1}{n_i} T_i^2 - \frac{1}{N} T_{..}^2$$

$$V_e = V_T - V_d$$

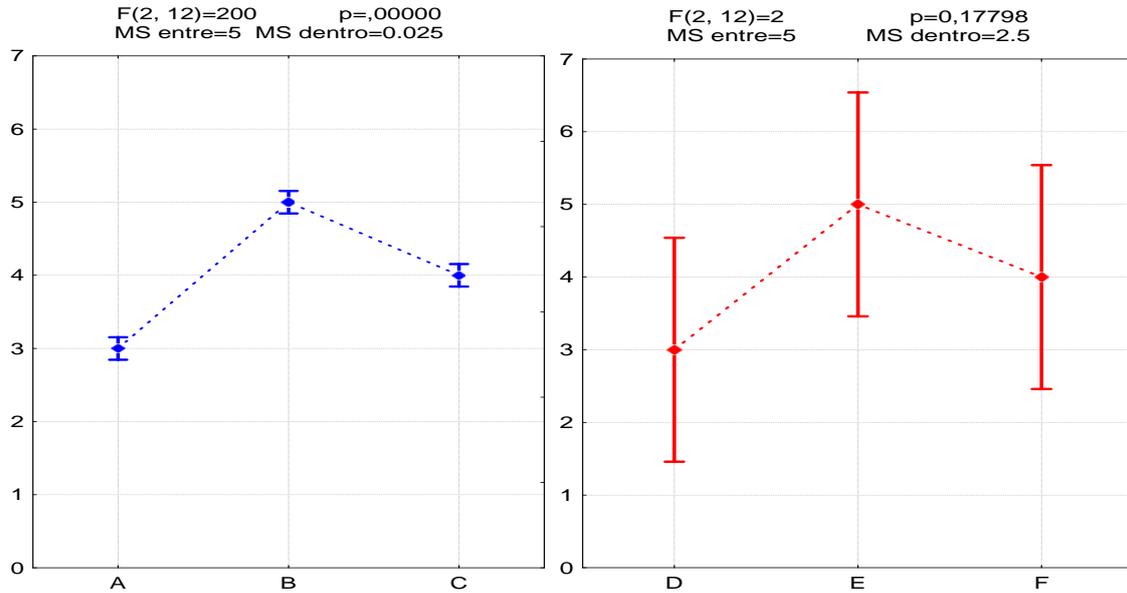


Figura 5.1: Importancia de la dispersión en el test F

donde se ha denotado

$$T_{..} = \sum_{i=1}^k \sum_{j=1}^{n_i} x_{ij}, \quad T_{i.} = \sum_{j=1}^{n_i} x_{ij}. \quad \Delta$$

Observación 5.4. (Estimación en el modelo de clasificación simple en análisis de la varianza) Como estimador de la media μ_i usaremos la media muestral \bar{X}_i del grupo i -ésimo, mientras que el denominador $S_d^2 = V_d/(N - k)$ del estadístico de contraste F es un estimador insesgado de la varianza σ^2 común a los k grupos. Un intervalo de confianza para μ_i al nivel $1 - \alpha$ es

$$\bar{X}_i \pm t_{\alpha}(N - k)S_d/\sqrt{n_i} \quad \Delta$$

Ejemplo 5.5. (Interpretación gráfica del test F) Consideremos los dos conjuntos de datos siguientes, dividido cada uno de ellos en 3 grupos (A,B,C y D,E,F, respectivamente) con 5 datos cada uno.

A	2.8	2.9	3	3.1	3.2	D	1	2	3	4	5
B	4.8	4.9	5	5.1	5.2	E	3	4	5	6	7
C	3.8	3.9	4	4.1	4.2	F	2	3	4	5	6

Las medias de los grupos A y D son iguales a 3, las de los grupos B y E son iguales a 5 y las de los grupos C y F son iguales a 4; sin embargo, el test F señala diferencias significativas entre las medias del primer conjunto, pero no entre las del segundo, debido a la diferente dispersión de los datos. \triangleleft

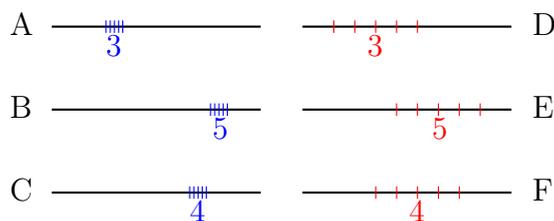


Figura 5.2: Importancia de la dispersión en la realización de un test F

Volvamos al ejemplo considerado en la introducción.

Ejemplo 5.6. (Ejemplo 5.1, continuación) Para el problema de comparación de los tiempos medios de espera hasta que el paciente se queja de dolor se tiene que $k = 3$, $n_1 = 5$, $n_2 = 5$, $n_3 = 5$, $\bar{X}_1. = 1.8$, $\bar{X}_2. = 2.9$, $\bar{X}_3. = 2.44$, $\bar{X}.. = 2.38$, $V_e = 3.052$, $V_d = 23.652$ y $S_d^2 = 23.652/12 = 1.971$. El valor experimental es

$$F_{\text{exp}} = \frac{3.052/(3 - 1)}{23.652/(15 - 3)} = 0.774$$

La tabla VI de la distribución F de Fisher proporciona los valores $F_{0.05}(2, 12) = 3.89$ y $F_{0.1}(2, 12) = 2.81$. Por tanto, el nivel mínimo de significación P verifica $0.1 < P$ y el test es no significativo; no existen, por tanto, diferencias significativas entre los tiempos de espera hasta que el paciente se queja de dolor para los 3 tratamientos considerados.⁽¹⁾

La tabla siguiente resume el análisis de la varianza realizado, en la que se ha incluido también el valor P del test:

Variación	Suma Cuadrados	G.L.	Media Cuadrática	F	P
Entre	3.052	2	1.526	0.774	0.483
Dentro	23.652	12	1.971		
Total	26.704	14			

Tabla. Análisis de la varianza. ◁

La Figura 5.3 presenta la posición relativa de los tiempos medios de los 3 grupos, junto con los intervalos al 95 % de confianza para cada una de ellas. ◁

Veamos un segundo ejemplo.

¹ El cociente $r^2 = V_e/V_T$ se llama coeficiente de determinación, y mide la fracción de variabilidad de la variable de interés explicada por la influencia del factor considerado. En el ejemplo ese cociente es igual a 0.1143, lo que viene a decir que el 11.43 % de la variabilidad en el tiempo de espera hasta que el paciente se queja de dolor queda explicada por el tratamiento analgésico.

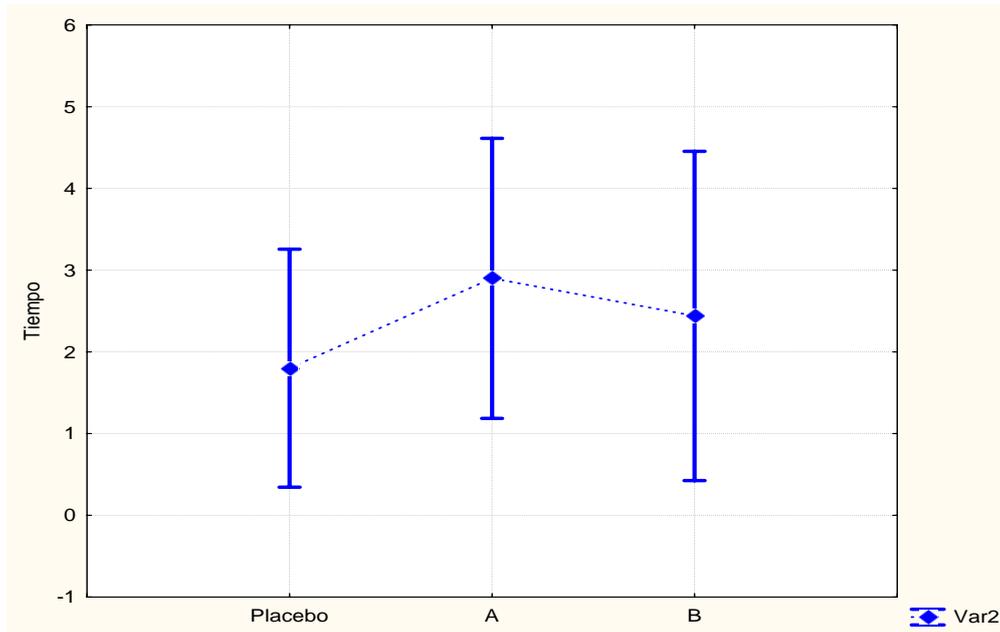


Figura 5.3: Test F : $F(2, 12) = 0,774$, $P = 0,483$

Ejemplo 5.7. Para estudiar el efecto que produce el ruido en la eficiencia en el trabajo, se han elegido 21 jóvenes al azar y se han dividido en tres grupos de 7 jóvenes cada uno. Se ha determinado para cada uno de esos individuos el número de errores cometidos en ciertas tareas aritméticas mientras las realizaban bajo tres diferentes niveles de ruido: ruido bajo continuo (RBC), ruido intermedio continuo (RIC) y ruido intermedio intermitente (RII). La tabla siguiente contiene los resultados de la experiencia:

RBC	7	8	11	10	9	8	7
RIC	9	10	8	9	11	8	9
RII	8	14	10	12	14	13	16

Se pretende decidir si el número medio de errores cometidos es el mismo para los 3 niveles de ruido considerados o existen diferencias significativas entre al menos dos de ellos. En este caso, se tiene que $k = 3$, $n_1 = 7$, $n_2 = 7$, $n_3 = 7$, $\bar{X}_1 = 8.57$, $\bar{X}_2 = 9.14$, $\bar{X}_3 = 12.43$, $\bar{X}_{..} = 10.05$, $V_e = 60.67$, $V_d = 64.29$ y $S_d^2 = 64.29/18 = 3.57$. El valor experimental es

$$F_{\text{exp}} = \frac{60.67/(3-1)}{64.29/(21-3)} = 8.49$$

La tabla VI de la distribución F de Fisher proporciona los valores $F_{0.05}(2, 18) = 3.55$, $F_{0.01}(2, 18) = 6.01$ y $F_{0.001}(2, 18) = 10.39$. Por tanto, el nivel mínimo de significación P verifica $0.001 < P < 0.01$ y el test es muy significativo; existen, por tanto, diferencias

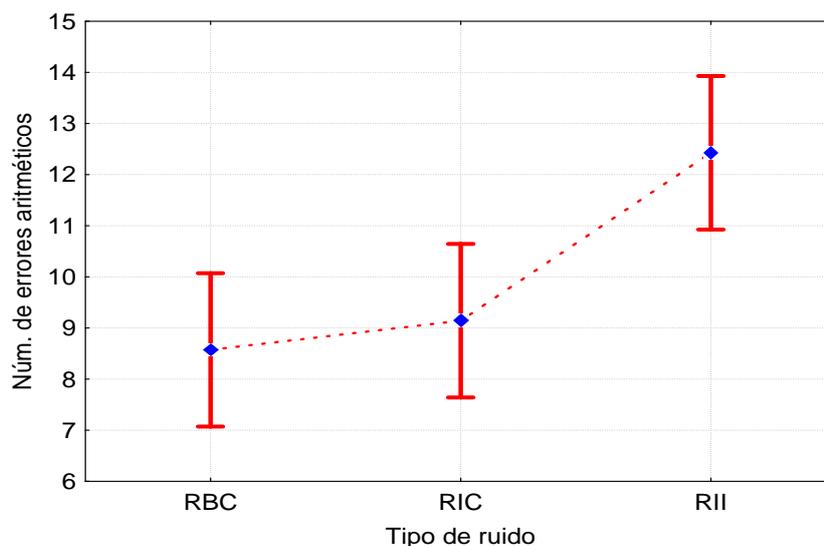


Figura 5.4: Test F : $F(2, 18) = 8,55$, $P = 0,0025$

significativas entre el número medio de errores cometidos bajo los 3 diferentes niveles de ruido considerados.⁽²⁾

La tabla siguiente resume el análisis de la varianza realizado, en la que se ha incluido también el valor P del test:

Variación	Suma Cuadrados	G.L.	Media Cuadrática	F	P
Entre	60.67	2	30.33	8.49	0.0025
Dentro	64.29	18	3.57		
Total	124.96	20			

Tabla. Análisis de la varianza ◁

La Figura 5.4 presenta la posición relativa de las medias de los 3 grupos, junto con los intervalos al 95 % de confianza para cada una de ellas. ◁

Observación 5.8. (★) Los programas estadísticos habituales suelen tener implementados diversos métodos estadísticos que nos permiten comprobar las suposiciones de independencia, normalidad e igualdad de varianzas que se hacen en el modelo de clasificación simple en análisis de la varianza: la independencia quedará garantizada por la elección al azar de los individuos; se utilizará algún método gráfico (histograma, gráfico de probabilidad normal) o algún test de hipótesis (como los de D’Agostino, Kolmogorov-Smirnov o la modificación

² El cociente $r^2 = V_e/V_T$ se llama coeficiente de determinación, y mide la fracción de variabilidad de la variable de interés explicada por la influencia del factor considerado. En el ejemplo ese cociente es igual a 0.4823, lo que viene a decir que el 40.23 % de la variabilidad en el número de errores aritméticos queda explicada por el nivel de ruido.

del mismo propuesta por Lilliefors, o el de Shapiro-Wilks) para comprobar la normalidad de los residuos $r_{ij} = x_{ij} - x_{i.}$; la suposición de igualdad de varianzas se puede contrastar con el test de Bartlett (véase la observación siguiente), el test de Levene o el test C de Cochran. Si la suposición de normalidad o de igualdad de varianzas no se verifica, podemos pensar en realizar un test no paramétrico como el de Kruskal-Wallis que veremos en la sección 13.3. \triangleleft

Observación 5.9. (★) (Test de Bartlett de comparación de varias varianzas) Se dispone de r muestras independientes de tamaños n_1, \dots, n_r de otras tantas distribuciones normales con varianzas desconocidas σ_i^2 , $1 \leq i \leq r$, y se desea contrastar la hipótesis $H_0 : \sigma_1^2 = \dots = \sigma_r^2$ de igualdad de esas varianzas. Sean s_1^2, \dots, s_r^2 sus correspondientes varianzas muestrales. Los tests que explicamos a continuación son válidos cuando la mayoría de los n_i son ≥ 6 . Distinguiremos dos casos:

- Si todos los n_i son iguales (digamos, a m) entonces la región crítica del test de Bartlett es:

$$\chi_{\text{exp}}^2 = \frac{(m-1)[r \log s^2 - \sum \log s_i^2]}{1 + \frac{r+1}{3r(m-1)}} > \chi_{\alpha}^2(r-1),$$

donde $s^2 = \sum s_i^2 / r$.

- Si los n_i no son todos iguales, la región crítica es

$$\chi_{\text{exp}}^2 = \frac{[\sum (n_i - 1) \log s^2 - \sum n_i \log s_i^2]}{1 + \frac{\sum \frac{1}{n_i - 1} - \frac{1}{\sum (n_i - 1)}}{3(r-1)}} > \chi_{\alpha}^2(r-1),$$

donde $s^2 = [\sum (n_i - 1) s_i^2] / [\sum (n_i - 1)]$.

\triangle

5.1.2. Comparaciones múltiples

El análisis de la varianza es sólo un primer paso en el análisis de los datos. Si el test descrito resulta significativo, el grupo de k medias se declara no homogéneo y cabe preguntarse por las causas de la significación examinando las medias de cada muestra y la magnitud de las diferencias entre ellas. Son varios los métodos de comparaciones múltiples propuestos. Describimos a continuación el test de Tukey-Kramer, que es una modificación de Kramer de un test similar propuesto por Tukey en el caso de que todas las muestras tengan el mismo tamaño. El procedimiento se resume en los pasos siguientes:

- 1) Sean $m_1 \leq \dots \leq m_k$ las k medias muestrales ordenadas de menor a mayor, μ'_1, \dots, μ'_k las medias poblacionales correspondientes y n'_1, \dots, n'_k los tamaños muestrales. Sea $N = n'_1 + \dots + n'_k$. Hagamos $i = 1$ y $j = k$.

- 2) Se comparan las medias μ'_i y μ'_j utilizando el estadístico de rango estudentizado

$$q_{ij} = \frac{m_j - m_i}{\sqrt{\frac{S_d^2}{2} \left(\frac{1}{n'_j} + \frac{1}{n'_i} \right)}}$$

en cuyo denominador, llamado error estándar para la comparación de μ'_i y μ'_j , aparece S_d^2 como estimador de la varianza común desconocida σ^2 . Si $n'_1 = \dots = n'_k = n$, el test de Tukey-Kramer se conoce como test de Tukey (llamado también test de la diferencia honestamente significativa), y q_{ij} admite la expresión más simple

$$q_{ij} = \sqrt{n} \frac{m_j - m_i}{S_d}$$

- 3) Las medias μ'_i y μ'_j se declaran homogéneas si $q_{ij} \leq q_\alpha(N - k, k)$ (véase la tabla X de la distribución q de Tukey). Se declaran igualmente homogéneas todas las medias comprendidas entre ellas, y no es necesario comparar dos medias dentro de un grupo homogéneo.
- 4) Las medias μ'_i y μ'_j se declaran no homogéneas si $q_{ij} > q_\alpha(N - k, k)$. En ese caso, se vuelve al paso 2 reemplazando i por $i + 1$ para comparar las medias μ'_{i+1} con μ'_j .
- 5) Si μ'_j se declara no homogénea a μ'_i para todo $i < j$, concluiremos que existen diferencias significativas al nivel α entre μ'_j y todas las medias menores que ella. Se reemplaza j por $j - 1$ y se vuelve al paso 2.
- 6) El proceso continúa hasta que no quede nada que comparar.

Ejemplo 5.10. (Ejemplo 5.7, continuación) En este ejemplo se verifica que $m_1 = \bar{X}_1. = 8,57$, $m_2 = \bar{X}_2. = 9,14$ y $m_3 = \bar{X}_3. = 12,43$. La tabla X.2 proporciona el valor crítico $q_{0,05}(18, 3) = 3,61$.

- (a) m_1 frente a m_3 :

$$q_{13} = \frac{m_3 - m_1}{S_d/\sqrt{7}} = \frac{12,43 - 8,57}{\sqrt{3,57/7}} = 5,41 > 3,61 \Rightarrow \mu_1 \neq \mu_3$$

- (b) m_2 frente a m_3 :

$$q_{23} = \frac{m_3 - m_2}{S_d/\sqrt{7}} = \frac{12,43 - 9,14}{\sqrt{3,57/7}} = 4,61 > 3,61 \Rightarrow \mu_2 \neq \mu_3$$

(c) m_1 frente a m_2 :

$$q_{12} = \frac{m_2 - m_1}{S_d/\sqrt{7}} = \frac{9,14 - 8,57}{\sqrt{3,57/7}} = 0,8 \leq 3,61 \Rightarrow \mu_1 = \mu_2$$

Se concluye pues que las dos primeras medias se pueden declarar homogéneas, mientras que la tercera es significativamente diferente de ellas. \triangleleft

Observación 5.11. El método de comparaciones múltiples descrito está basado en las mismas suposiciones de normalidad e igualdad de varianzas del modelo de clasificación simple en análisis de la varianza. \triangleleft

Observación 5.12. En general, se acostumbra a sugerir el uso del test de Tukey-Kramer una vez que el test F ha resultado significativo. En ese caso, conviene a posteriori comparar todas las medias dos a dos tal y como se ha explicado anteriormente. \triangleleft

Observación 5.13. Una vez realizadas las comparaciones múltiples, el grupo original de medias queda dividido en subgrupos homogéneos que no tienen por qué ser disjuntos. Por ejemplo, si tenemos un total de 3 medias μ_1 , μ_2 y μ_3 , puede ocurrir que terminemos aceptando que $\mu_1 = \mu_2$ y $\mu_2 = \mu_3$, pero que $\mu_1 \neq \mu_3$; eso no debe entenderse como una contradicción, pues, según hemos dicho anteriormente, no debemos considerar la aceptación de la hipótesis nula como una demostración de la misma, y bien pudiera ocurrir que las discrepancias observadas entre la primera y la segunda medias muestrales y entre la segunda y la tercera no sean estadísticamente significativas, pero sí lo sea la discrepancia entre la primera y la tercera. Podemos decir en ese caso que existen dos subgrupos homogéneos de medias entre las 3 medias consideradas, uno que contiene a la primera media y otro a la tercera, mientras que los datos obtenidos no nos permiten decidir a cuál de ellos pertenece la segunda media. \triangleleft

Observación 5.14. (Intervalos de confianza simultáneos) Los $k(k-1)/2$ intervalos siguientes son intervalos de confianza para las diferencias entre medias con un nivel de confianza simultáneo igual a $1 - \alpha$:

$$\mu_i - \mu_j \in x_i - x_j \pm q_\alpha(N - k, k) \sqrt{\frac{S_d^2}{2} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}, \quad 1 \leq i < j \leq k$$

\triangle

Observación 5.15. (★) (Método de Bonferroni) Si el test de Tukey es usualmente aceptado en el caso de que las muestras tengan el mismo tamaño cuando se quieren comparar todas las medias dos a dos, otros autores sugieren para el caso de tamaños de muestra desiguales el uso del método de Bonferroni que describimos a continuación, método que también podemos utilizar cuando sólo queremos comparar algunos pares de medias. El método de Bonferroni (llamado así por estar basado en la desigualdad de Bonferroni que vimos en el tema 1), propone lo siguiente: para realizar s tests de hipótesis simultáneos (se desea contrastar que las s hipótesis nulas son ciertas contra que alguna de ellas no lo es) a un nivel global de significación α , cada uno de los tests individuales se realizará a un nivel de significación

α/s ; de ese modo, si las hipótesis nulas de los s tests se suponen ciertas, la desigualdad de Bonferroni garantiza que la probabilidad de error de tipo I es menor o igual que α , pues la probabilidad de rechazar incorrectamente la hipótesis nula global de que todas las hipótesis nulas individuales son ciertas coincide con la probabilidad de que alguna de esas hipótesis nulas sea falsa, probabilidad que es menor o igual que la suma de las probabilidades de error de tipo I particulares, en virtud de la propiedad (P9) de la página 48. Este método lo podemos aplicar para realizar un cierto número de comparaciones de medias planificadas de antemano (bien entendido que, si se desean comparar todas las medias dos a dos, es preferible el método de Tukey en el caso de que todos los tamaños muestrales sean iguales). Como hemos dicho anteriormente, se puede utilizar también como alternativa al test de Tukey en el caso de que no todas las muestras tengan el mismo tamaño, teniendo en cuenta que en ese caso hay que comparar $s = k(k-1)/2$ pares de medias. La tabla XII contiene los valores críticos $b_\alpha(N-k, s)$ de la distribución de Bonferroni para los niveles de significación globales $\alpha = 0,1, 0,05$ y $0,01$; la primera columna de cada tabla contiene el valor $N-k$, mientras que en la primera fila leeremos el número s de comparaciones a realizar. También ahora utilizaremos el estadístico de rango estudentizado

$$q_{ij} = \frac{m_j - m_i}{\sqrt{\frac{S_d^2}{2} \left(\frac{1}{n'_j} + \frac{1}{n'_i} \right)}}$$

para comparar las medias m_i y m_j . En este caso, a diferencia del método de Tukey-Kramer, no se recomienda declarar homogéneas dos medias situadas entre dos que ya han sido declaradas homogéneas pues, debido a que los tamaños muestrales son diferentes, puede ocurrir que la diferencia entre ellas sea significativa. \triangleleft

Observación 5.16. (★) (Otros métodos de comparaciones múltiples) Además del método de Tukey y el de Tukey-Kramer, indicados en el caso de que se deseen comparar las medias dos a dos, y del método de Bonferroni, método muy general indicado para realizar varios tests de hipótesis simultáneamente que se puede utilizar en particular para comparaciones múltiples como hemos visto, son diversos los procedimientos de comparaciones múltiples que se han descrito en la literatura con diferentes objetivos, como pueden ser el método de la mínima diferencia significativa, indicado en el caso de que se deseen comparar sólo algunas de las muestras habiendo seleccionado dichas muestras antes de realizar el experimento (comparaciones no sugeridas por los datos), o el método de Dunnet, especialmente diseñado para comparar $k-1$ de las medias con la restante (que podría corresponder a un grupo control). El lector puede encontrar, por ejemplo, en Woolson (1987) una exposición más detallada de estos y otros métodos. \triangleleft

Observación 5.17. (★) (Otros modelos en análisis de la varianza) En este tema hemos estudiado la influencia de un factor sobre la media de una distribución normal: un factor que actúa a diferentes niveles y que, suponemos, no afecta a la varianza de esa distribución. En la literatura se describen otros modelos de análisis de la varianza en los que, por ejemplo, se considera la influencia sobre la media de dos o más factores, aunque su estudio queda lejos del alcance de este libro. \triangleleft

5.1.3. (★) Solución no paramétrica al problema de comparación de k muestras:

Test de Kruskal-Wallis

Para comparar k muestras utilizando un método no paramétrico, se combinan todas las muestras para obtener una muestra combinada de tamaño $N = n_1 + \dots + n_k$. Se asignan rangos a cada una de las observaciones individuales (la muestra combinada se ordena de menor a mayor y el rango de una observación es el lugar que ocupa en la sucesión ordenada; se utilizará el rango medio en el caso de empates, como se explica en la sección 12.6). Para $1 \leq i \leq k$, denotaremos por R_i la suma de los rangos de las observaciones de la muestra i -ésima. Consideremos el estadístico

$$H^* = \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1)$$

Si no hay empates, se utiliza como estadístico de contraste $H = H^*$, que coincide con

$$\frac{12}{N(N+1)} \sum_{i=1}^k n_i \left(\frac{R_i}{n_i} - \frac{R}{N} \right)^2$$

donde $R = \sum_{i=1}^k R_i$. Si hay empates (digamos, r grupos de empates), se utiliza el estadístico de contraste

$$H = \frac{H^*}{C}$$

donde

$$C = 1 - \frac{1}{N^3 - N} \sum_{j=1}^r T_j$$

siendo $T_j = t_j^3 - t_j$ y t_j el número de observaciones en el j -ésimo grupo de empates, $1 \leq j \leq r$.

La región crítica del test de Kruskal-Wallis es, en cualquier caso,

$$H > \chi_{\alpha}^2(k-1)$$

El valor P del test viene a ser la probabilidad de que $\chi^2(k-1) > H_{\text{exp}}$. La aproximación χ^2 utilizada para la distribución de H es válida cuando los tamaños de las muestras no son muy pequeños: suele exigirse en la práctica que todas las muestras tengan, al menos, tamaño 5. Existen tablas para algunos casos en los que este requisito no se da.

Si el test de Kruskal–Wallis resulta significativo, podemos pensar en hacer comparaciones múltiples entre las muestras siguiendo un procedimiento análogo (método de Dunn) al método de Tukey-Kramer descrito anteriormente, pero utilizando ahora para comparar las muestras i y j el estadístico de contraste

$$q'_{ij} = \frac{\left| \frac{R_i}{n_i} - \frac{R_j}{n_j} \right|}{\sqrt{\frac{CN(N+1)}{12} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}}$$

que compararemos con el valor $Q_\alpha(k)$ proporcionado en la tabla XI si $k \leq 25$.

Ejemplo 5.18. (★) (Ejemplo 5.7, continuación) Analizamos a continuación los datos del ejemplo sobre la influencia del ruido en el número de errores aritméticos cometidos mediante el test de Kruskal-Wallis que acabamos de describir. Reordenados de menor a mayor esos datos se obtiene la sucesión:

7, 7, 8, 8, 8, 8, 8, 9, 9, 9, 9, 10, 10, 10, 10, 11, 11, 12, 13, 14, 14, 16

La tabla siguiente contiene, además de los datos obtenidos, el rango de cada uno de ellos entre paréntesis, y en la columna de la derecha la suma de los rangos del grupo correspondiente:

RBC	7(1.5)	8(5)	11(15.5)	10(13)	9(9.5)	8(5)	7(1.5)	$R_1 = 51$
RIC	9(9.5)	10(13)	8(5)	9(9.5)	11(15.5)	8(5)	9(9.5)	$R_2 = 67$
RII	8(5)	14(19.5)	10(13)	12(17)	14(19.5)	13(18)	16(21)	$R_3 = 113$

Entonces

$$H^* = \frac{1}{21 \cdot 22} \left(\frac{51^2}{7} + \frac{67^2}{7} + \frac{113^2}{7} \right) - 3 \cdot 22 = 7.69$$

y, puesto que hay 6 grupos de empates con 2, 5, 4, 3, 2 y 2 datos cada uno en la sucesión reordenada precedente,

$$C = 1 - \frac{1}{21^3 - 21} \left((2^3 - 2) + (2^3 - 2) + (5^3 - 2) + (4^3 - 2) + (3^3 - 2) + (2^3 - 2) \right) = 0.98$$

Luego,

$$H = \frac{H^*}{C} = 7.88$$

Puesto que el valor crítico es $\chi_{0.05}^2(2) = 5.99$, rechazaremos la hipótesis nula de homogeneidad de los 3 grupos considerados. Siendo $\chi_{0.01}^2(2) = 9.21$, se verifica que $0.01 < P < 0.05$. Recordemos que el test F resultó muy significativo. ⁽³⁾

Procedemos también en este caso a realizar comparaciones múltiples; necesitaremos el valor crítico $Q_{0.05}(3) = 2.394$:

³ Con la ayuda de un programa estadístico podemos calcular exactamente el valor P en ambos casos: mientras que $P = 0.0025$ para el test F , para el de Kruskal-Wallis $P = 0.0195$.

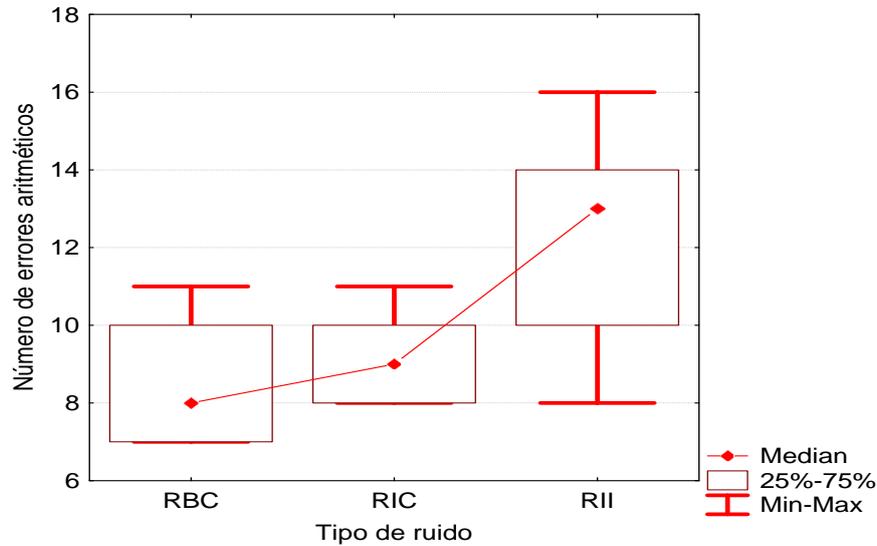


Figura 5.5: Test Kruskal-Wallis: $H = 7.88$, $P = 0.0195$

(a) Muestra 1 frente a muestra 3: son no homogéneas pues

$$q'_{13} = \frac{\left| \frac{R_1}{n_1} - \frac{R_3}{n_3} \right|}{\sqrt{\frac{CN(N+1)}{12} \left(\frac{1}{n_1} + \frac{1}{n_3} \right)}} = \frac{16.14 - 7.29}{\sqrt{3.28}} = 2.7 > 2.394$$

(b) Muestra 2 frente a muestra 3: son declaradas homogéneas pues

$$q'_{23} = \frac{\left| \frac{R_2}{n_2} - \frac{R_3}{n_3} \right|}{\sqrt{\frac{CN(N+1)}{12} \left(\frac{1}{n_2} + \frac{1}{n_3} \right)}} = \frac{16.14 - 9.57}{\sqrt{3.28}} = 2.01 < 2.394$$

(c) Muestra 1 frente a muestra 2: son declaradas homogéneas pues

$$q'_{12} = \frac{\left| \frac{R_1}{n_1} - \frac{R_2}{n_2} \right|}{\sqrt{\frac{CN(N+1)}{12} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} = \frac{9.57 - 7.29}{\sqrt{3.28}} = 0.7 < 2.394$$

Nótese como la decisión de tratar unos mismos datos con métodos estadísticos diferentes puede dar lugar a conclusiones diferentes. Aunque tanto el test F como el de Kruskal-Wallis sugieren rechazar la homogeneidad de las 3 muestras, los valores P son diferentes, y las comparaciones múltiples subsiguientes son ligeramente distintas, pues mientras en el caso paramétrico las muestras 1 y 2 constituyen un subgrupo homogéneo significativamente

distinto del formado por la tercera, en el caso no paramétrico los datos no nos permiten precisar a qué subgrupo pertenece la segunda muestra: al de la primera o al de la tercera, que desde luego están en diferentes subgrupos.

En el caso no paramétrico puede ser preferible utilizar un diagrama de caja en cada grupo como el de la figura para visualizar la diferencia entre ellos. \triangleleft

5.1.4. (★) Modelo de clasificación doble en análisis de la varianza sin interacciones y una observación por celda o diseño de bloques al azar: test F (paramétrico) y test de Friedman (no paramétrico)

Se consideran dos factores (que actúan a r y t niveles) que influyen sobre la media (pero no sobre la varianza) de una distribución normal. Se eligen al azar rt individuos y a cada uno de ellos se le aplica una de las rt posibles combinaciones de los niveles de los dos factores (cada combinación (i, j) posible se aplica a uno y sólo un individuo: de ahí lo de una observación por celda), obteniendo rt respuestas x_{ij} , $1 \leq i \leq r$, $1 \leq j \leq t$ que consideraremos como observaciones independientes de otras tantas distribuciones normales de medias μ_{ij} , $1 \leq i \leq r$, $1 \leq j \leq t$, y varianza común σ^2 (que no queda afectada por ninguno de los dos factores considerados). Supondremos además que los factores actúan aditivamente, en el sentido de que cada media μ_{ij} , aunque desconocida, se supone suma de dos componentes α'_i y β'_j , dependiente cada una de ellas de un único factor y llamadas efectos debidos a los niveles i y j en que se encuentran los factores (se dice también que no existen interacciones entre los factores). Es preferible descomponer en ese caso la media μ_{ij} como suma de una media general μ y de dos componentes α_i y β_j que conoceremos como efectos principales con las restricciones $\sum_i \alpha_i = \sum_j \beta_j = 0$ (ambos planteamientos son equivalentes).

Observación 5.19. El modelo considerado se utiliza también para situaciones en las que el diseño experimental elegido es el llamado *diseño de bloques al azar*: este diseño fue utilizado por primera vez por Fisher en 1925 cuando buscaba métodos para mejorar la experimentación en agricultura; el término diseño refleja su origen en experimentos agrícolas

donde la tierra era dividida en bloques homogéneos y los bloques eran divididos en parcelas que recibían los tratamientos bajo investigación. El diseño de bloques al azar es un diseño en el que las unidades experimentales que recibirán los distintos tratamientos son subdivididas en grupos homogéneos llamados **bloques**; el número de unidades experimentales en cada bloque es igual al número de tratamientos de interés y los distintos tratamientos son asignados al azar a las unidades experimentales dentro de cada bloque. Siendo las parcelas dentro de cada bloque homogéneas, este diseño es una generalización del diseño de muestras apareadas estudiado en el caso de comparación de dos tratamientos (en el que cada individuo recibe los dos tratamientos, como ahora cada uno de los t bloques recibe todos y cada uno de los r tratamientos considerados); podría usarse este diseño incluso cuando se dispone de t individuos (cada individuo es un bloque) y cada uno de ellos recibe todos y cada uno de los r tratamientos considerados en el caso de que no haya superposición de efectos en la aplicación de los distintos tratamientos a un mismo individuo. \triangleleft

Tenemos pues un conjunto de datos x_{ij} , $1 \leq i \leq r$, $1 \leq j \leq t$. Denotaremos por $x_i.$ la media de la fila i -ésima, por $x_{.j}$ la media de la columna j -ésima, y por $x_{..}$ la media general. Consideraremos los estadísticos siguientes:

- Suma de cuadrados “entre filas” ($r - 1$ grados de libertad):

$$V_{ef} = t \sum_{i=1}^r [x_i. - x_{..}]^2$$

- Suma de cuadrados “entre columnas” ($t - 1$ grados de libertad):

$$V_{ec} = r \sum_{j=1}^t [x_{.j} - x_{..}]^2$$

- Suma de cuadrados “dentro” o error ($(r - 1)(t - 1)$ grados de libertad):

$$V_d = r \sum_{i=1}^r \sum_{j=1}^t [x_{ij} - x_i. - x_{.j} + x_{..}]^2$$

- Variación total:

$$\sum_i \sum_j [x_{ij} - x_{..}]^2 = V_{ef} + V_{ec} + V_d$$

- Estimador de la varianza desconocida σ^2 :

$$s_d^2 = mc_d = \frac{V_d}{(r - 1)(t - 1)}$$

Consideremos ahora el problema de contrastar la hipótesis de que el primer factor no influye sobre la media, es decir, $H_0 : \alpha_1 = \dots = \alpha_r$. La región crítica del test es:

$$F_{\text{exp}} = \frac{\frac{1}{r-1} V_{ef}}{\frac{1}{(r-1)(t-1)} V_d} > F_{\alpha}(r - 1, (r - 1)(t - 1)).$$

Análogamente se contrastaría la hipótesis de que el segundo factor no influye sobre la media.

Los resultados se suelen recoger en una tabla como la siguiente:

Variación	Suma Cuadrados	G.L.	Media Cuadrática
Entre filas	$V_{ef} = t \sum_{i=1}^r [x_{i.} - x_{..}]^2$	$r - 1$	$mc_{ef} = V_{ef}/(r - 1)$
Entre columnas	$V_{ec} = r \sum_{j=1}^t [x_{.j} - x_{..}]^2$	$t - 1$	$mc_{ec} = V_{ec}/(t - 1)$
Dentro (error)	$V_d = \sum_{i=1}^r \sum_{j=1}^t [x_{ij} - x_{i.} - x_{.j} + x_{..}]^2$	$(r - 1) \cdot (t - 1)$	$mc_d = V_d/[(r - 1)(t - 1)]$
Total	$V_T = V_{ef} + V_{ec} + V_d = \sum_{ij} (x_{ij} - x_{..})^2$	$rt - 1$	

Tabla de Análisis de la Varianza

Observación 5.20. (★) (Solución no paramétrica: test de Friedman) El test de Friedman es la solución no paramétrica al modelo de clasificación doble sin interacciones en análisis de la varianza. Como habíamos explicado anteriormente, el modelo de clasificación doble sin interacciones y una observación por celda es una generalización a más de dos muestras del problema de comparación de dos muestras apareadas y, por tanto, el test de Friedman es la generalización del test de Wilcoxon para muestras apareadas. Supongamos que estamos interesados en contrastar la hipótesis nula H_0 de que la media no depende realmente del primer factor (si hablamos de tratamientos en lugar de primer factor y de bloques en lugar de segundo factor, se trata de contrastar la hipótesis de que ningún tratamiento tiende a dar valores más altos o más bajos que los otros). Los datos se disponen en un cuadro de doble entrada en el que el primer factor (r tratamientos) aparece en las filas y el segundo (t bloques) en las columnas. Se asignan rangos (teniendo en cuenta los posibles empates) a los r datos de la primera columna (anotaremos al lado de cada dato en la tabla su rango entre paréntesis); procederemos de igual modo con las restantes columnas (asignando rangos a los datos de cada columna, independientemente de los rangos asignados en otras columnas). Denotaremos por R_1, \dots, R_r las sumas de los rangos correspondientes a los tratamientos $1, \dots, r$. Debe ocurrir que $\sum_i R_i = tr(r + 1)/2$. El test de Friedman viene definido por la región crítica:

$$\frac{\frac{12}{rt(r+1)} \sum R_i - 3t(r + 1)}{1 - \frac{1}{(r-1)r(r+1)t} \sum_i \sum_j T_{ij}} > \chi^2(r - 1),$$

donde $T_{ij} = (t_{ij} - 1)t_{ij}(t_{ij} + 1)$, siendo t_{ij} el número de empates ocurridos en el j -ésimo grupo de empates del bloque i . El test descrito es válido para muestras grandes; a efectos prácticos suele exigirse que $rt > 25$. En el caso $r = 2$ este test coincide con el de Wilcoxon para muestras apareadas. ◁

5.2. Comparación de Varias Muestras Cualitativas

5.2.1. Tablas de contingencias

Consideramos en este tema el problema de la comparación de varias muestras cualitativas. Concretamente, se consideran r muestras cuyos individuos pueden clasificarse en s clases disjuntas. Los datos suelen recogerse en una tabla $r \times s$, llamada tabla de contingencias, cuya entrada O_{ij} , $1 \leq i \leq r$, $1 \leq j \leq s$, indica el número de individuos de la muestra i -ésima que están en la clase j -ésima. En los márgenes de la tabla consignaremos las cantidades $F_i = \sum_j O_{ij}$ (tamaño de la muestra i -ésima) y $C_j = \sum_i O_{ij}$ (número de individuos en la clase j). Denotaremos por T el número total de individuos seleccionados.

	1	...	j	...	s	Totales
1	O_{11}	...	O_{1j}	...	O_{1s}	F_1
\vdots						
i	O_{i1}	...	O_{ij}	...	O_{is}	F_i
\vdots						
r	O_{r1}	...	O_{rj}	...	O_{rs}	F_r
Totales	C_1	...	C_j	...	C_s	T

Tabla de contingencias $r \times s$

Estamos interesados en contrastar la hipótesis nula H_0 de que la proporción de individuos que cae en una cierta clase es la misma para todas las muestras (es decir, que no hay diferencias significativas entre las r muestras en lo que a su clasificación en alguna de las s clases consideradas se refiere). Un ejemplo nos ayudará a comprender el planteamiento del problema.

Ejemplo 5.21. Se ha llevado a cabo un estudio en 4 regiones españolas (1, 2, 3 y 4) con la intención de comparar la opinión que sus ciudadanos tienen sobre la atención sanitaria

pública que reciben. Para ello, se han seleccionado en esas regiones muestras de tamaños respectivos 510, 1305, 492 y 223. Los datos obtenidos son los de la siguiente tabla de contingencias, y nos preguntamos si, a la luz de esos datos, podemos aceptar la hipótesis nula de que la distribución de los individuos según su opinión sobre la atención sanitaria que reciben es la misma en esas 4 regiones o si, por el contrario, existen diferencias significativas en al menos dos de ellos.

	Buena	Regular	Mala	Totales
Región 1	$O_{11} = 400$	$O_{12} = 70$	$O_{13} = 40$	$F_1 = 510$
Región 2	$O_{21} = 927$	$O_{22} = 310$	$O_{23} = 68$	$F_2 = 1305$
Región 3	$O_{31} = 300$	$O_{32} = 150$	$O_{33} = 42$	$F_3 = 492$
Región 4	$O_{41} = 129$	$O_{42} = 79$	$O_{43} = 15$	$F_4 = 223$
Totales	$C_1 = 1756$	$C_2 = 609$	$C_3 = 165$	$T = 2530$

Cantidades observadas

△

5.2.2. Comparación de varias muestras cualitativas

Supongamos cierta la hipótesis nula y calculemos la cantidad E_{ij} esperada de individuos en la clase (i, j) bajo esa suposición. Consideremos, por ejemplo, la casilla $(2, 3)$, es decir, el número esperado de individuos con un interés medio por la estadística y alta capacidad para las matemáticas si no hubiera diferencias entre los 3 niveles de interés por la estadística; si H_0 es cierta, la proporción p_3 desconocida de individuos con alta capacidad para las matemáticas puede ser estimada por la proporción muestral \hat{p}_3 , cociente entre el número total de $C_3 = 100$ individuos con alta capacidad matemática por el número total de individuos $T = 2530$ y, puesto que hay $F_2 = 1305$ en la segunda muestra (individuos con interés medio por la estadística), cabe esperar que, entre ellos, $F_2\hat{p}_3 = F_2C_3/T = 85,11$ tengan alta capacidad matemática. En general, si la hipótesis nula se supone cierta, la cantidad esperada para la casilla (i, j) es $E_{ij} = F_iC_j/T$. La discrepancia entre los valores realmente observados O_{ij} y los valores esperados bajo la hipótesis nula E_{ij} se mide con el estadístico

$$\chi^2 = \sum_i \sum_j \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \sum_i \sum_j \frac{O_{ij}^2}{E_{ij}} - T$$

Ese estadístico tiene bajo ciertas condiciones y bajo la hipótesis nula distribución aproximada $\chi^2((r-1)(s-1))$, lo que sugiere el uso de la región crítica:

$$\chi^2 > \chi_\alpha^2((r-1)(s-1))$$

Las condiciones de validez del test anterior han sido fijadas por Cochran del siguiente modo:

1) ninguna de las cantidades esperadas E_{ij} es < 1 ; 2) no más del 20% de las cantidades esperadas es < 5 .

Ejemplo 5.22. (Ejemplo 5.21, continuación) La tabla siguiente, junto a las cantidades observadas O_{ij} , contiene entre paréntesis las cantidades esperadas correspondientes E_{ij} :

	Buena	Regular	Mala	Totales
Región 1	$O_{11} = 400(353,98)$	$O_{12} = 70(122,76)$	$O_{13} = 40(33,26)$	$F_1 = 510$
Región 2	$O_{21} = 927(905,76)$	$O_{22} = 310(314,13)$	$O_{23} = 68(85,11)$	$F_2 = 1305$
Región 3	$O_{31} = 300(341,48)$	$O_{32} = 150(118,43)$	$O_{33} = 42(32,09)$	$F_3 = 492$
Región 4	$O_{41} = 129(154,78)$	$O_{42} = 79(53,68)$	$O_{43} = 15(14,54)$	$F_4 = 223$
Totales	$C_1 = 1756$	$C_2 = 609$	$C_3 = 165$	$T = 2530$

Cantidades observadas y esperadas

Se deduce de ello que el valor experimental del test χ^2 es

$$\chi^2 = \sum_{i=1}^4 \sum_{j=1}^3 \frac{O_{ij}^2}{E_{ij}} - T = 66.79$$

Es claro que se verifican las condiciones de validez del test. Puesto que $\chi_{0.05}^2((4-1) \cdot (3-1)) = 12.59$, los datos son incompatibles con la hipótesis nula. Siendo $\chi_{0.001}^2((4-1) \cdot (3-1)) = 22.46$, debemos señalar diferencias altamente significativas entre las 4 regiones considerados en lo que se refiere a la opinión que sus ciudadanos tienen sobre la atención sanitaria pública recibida. \triangleleft

Observación 5.23. (El caso 2×2) En el caso especial de una tabla 2×2 , suelen exigirse las condiciones de validez siguientes: 1) los marginales F_i y C_j son todos mayores que $T/10$; 2) todas las E_{ij} son mayores que 5. La región crítica del test es, en este caso,

$$\chi^2 = \frac{(O_{11}O_{22} - O_{12}O_{21})^2}{F_1 F_2 C_1 C_2} T > \chi_\alpha^2(1)$$

Este test coincide, salvo en las condiciones de validez, con el test de comparación de dos proporciones en el caso de dos muestras independientes mediante las identificaciones $O_{11} = x_1$, $O_{12} = x_2$, $O_{21} = n_1 - x_1$ y $O_{22} = n_2 - x_2$, en las que se han utilizado las notaciones de la sección 12.8. Ello es debido a que el cuadrado de una v.a. con distribución $N(0, 1)$ tiene distribución $\chi^2(1)$. \triangleleft

Observación 5.24. (★) (Corrección por continuidad de Yates) En el caso 2×2 , muchos autores continúan utilizando la corrección por continuidad de Yates en el estadístico de contraste χ^2 descrito anteriormente; esta corrección por continuidad propone reemplazar $(O_{ij} - E_{ij})^2$ por $(|O_{ij} - E_{ij}| - 0.5)^2$ en la definición del estadístico χ^2 , o, lo que es lo mismo, considerar la región crítica

$$\chi^2 = \frac{(|O_{11}O_{22} - O_{12}O_{21}| - \frac{1}{2}T)^2}{F_1F_2C_1C_2}T > \chi_\alpha^2(1)$$

Esta corrección por continuidad proporciona, en general, valores P mayores y es discutida por muchos autores. Nosotros no haremos uso de ella en este libro. Es, sin embargo, habitual en la literatura no considerar la corrección por continuidad en tablas de contingencia de dimensión superior. ◁

Observación 5.25. (★) (Partición de tablas) Cuando el test χ^2 de comparación de varias proporciones (o, más generalmente, el test χ^2 de comparación de varias muestras cualitativas) resulta significativo, cabe preguntarse por las causas de la significación, es decir, qué proporciones o muestras cualitativas o tratamientos son diferentes y cuáles pueden considerarse iguales. Éste es un problema análogo al de las comparaciones múltiples considerado en análisis de la varianza y que, en este caso, se resuelve haciendo una partición de la tabla $r \times s$. A la hora de hacer una partición de la tabla de datos podemos apoyarnos en una tabla auxiliar $r \times s$ que contenga la contribución $(O_{ij} - E_{ij})^2/E_{ij}$ de cada casilla (i, j) al estadístico χ^2 y contenga en el margen izquierdo la suma de las contribuciones correspondientes a cada fila. Esos totales marginales pueden hacernos sospechar que una de las r filas contribuye mucho más a la significación que los demás, lo cual puede comprobarse analizando una tabla $(r - 1) \times s$ obtenida de la original eliminando la fila correspondiente. Si el test χ^2 para la tabla reducida resulta no significativo, los tratamientos correspondientes se declaran homogéneos y el que habíamos eliminado se declara discrepante del resto. Si en alguna subtabla el test χ^2 resulta significativo se procede para ella como se hizo anteriormente. Si para una subtabla el test resulta no significativo, a veces se agrupan las filas correspondientes en una sola (pues se consideran provenientes de una misma población) y se continúa adelante con la partición. Una vez concluido el proceso de partición de la tabla por filas, dejando claro qué tratamientos pueden ser considerados iguales y cuáles no, puede continuarse haciendo algo análogo con las columnas, con el fin de dejar claro en qué columnas difieren dos filas que han sido declaradas distintas. ◁

5.2.3. Comparación de varias proporciones: muestras independientes

El problema de comparación de varias proporciones en el caso de muestras independientes es un caso particular del problema de comparación de varias muestras cualitativas en

el que $s = 2$ (o $r = 2$, si las poblaciones muestreadas aparecen en la primera fila de la tabla). Las tablas $r \times 2$ y $2 \times r$ pueden considerarse como una generalización de las tablas 2×2 , pues si en éstas se comparan dos proporciones, en una tabla $r \times 2$ se comparan r proporciones simultáneamente.

Ejemplo 5.26. (Ejemplo 5.21, continuación) Supongamos que se declara como satisfactoria una opinión buena e insatisfactoria una opinión regular o mala sobre la atención sanitaria recibida. Los tabla queda en la forma:

	Satisfactoria	Insatisfactoria	Totales
Región 1	$O_{11} = 400$	$O_{12} = 110$	$F_1 = 510$
Región 2	$O_{21} = 927$	$O_{22} = 378$	$F_2 = 1305$
Región 3	$O_{31} = 300$	$O_{32} = 192$	$F_3 = 492$
Región 4	$O_{41} = 129$	$O_{42} = 94$	$F_4 = 223$
Totales	$C_1 = 1756$	$C_2 = 774$	$T = 2530$

Si denotamos por p_i la proporción de individuos que consideran satisfactoria la atención sanitaria pública que recibe en la región i -ésima ($1 \leq i \leq 4$), podemos estar interesados en contrastar la hipótesis nula $H_0 : p_1 = p_2 = p_3 = p_4$ de homogeneidad de esas 4 proporciones, que podríamos interpretar como que no hay diferencias significativas entre la opinión que los ciudadanos de esas 4 regiones sobre la atención sanitaria pública recibida, contra la hipótesis alternativa de que, al menos, 2 de esas proporciones son diferentes. Puede comprobarse que se verifican las condiciones de validez del test chi-cuadrado y que el valor experimental del mismo es $\chi_{exp}^2 = 51,69$. Puesto que $\chi_{0,05}^2(3) = 7,82$, rechazamos la hipótesis nula de homogeneidad de las 4 proporciones; incluso $p < 0,001$. \triangleleft

Observación 5.27. En el caso 2×2 , este test coincide, salvo en las condiciones de validez, con el test de comparación de dos proporciones en el caso de dos muestras independientes mediante las identificaciones $O_{11} = x_1$, $O_{12} = x_2$, $O_{21} = n_1 - x_1$ y $O_{22} = n_2 - x_2$, en las que se han utilizado las notaciones de la sección 12.8. Ello es debido a que el cuadrado de una v.a. con distribución $N(0, 1)$ tiene distribución $\chi^2(1)$. \triangleleft

5.2.4. (★) Comparación de varias proporciones: muestras relacionadas

En el problema de comparación de varias proporciones también podemos distinguir dos tipos de diseños: el de muestras independientes y el de muestras relacionadas. Como hemos comentado anteriormente, el problema de comparación de varias proporciones en el caso

de muestras independientes es un caso particular del problema de comparación de varias muestras cualitativas. Presentaremos a continuación el test de Cochran para la comparación de varias proporciones relacionadas. Se dispone de r tratamientos que queremos comparar del siguiente modo: t individuos elegidos al azar reciben todos y cada uno de los tratamientos (con la debida separación de tiempo para evitar la superposición de efectos), exigiéndoles una opinión favorable (1) o desfavorable NO (0) sobre cada tratamiento. Si p_i denota la proporción de individuos de la población favorable al tratamiento i -ésimo, pretendemos contrastar la hipótesis nula $H_0 : p_1 = \dots = p_r$ de que todos los tratamientos son igualmente preferidos contra la alternativa de que no todos los tratamientos son igualmente preferidos. Los datos pueden presentarse en una tabla $r \times t$ cuya entrada (i, j) es 1 o 0 según la respuesta del individuo i -ésimo al tratamiento j -ésimo. Denotemos por F_i la suma de la fila i -ésima, por C_j la suma de la columna j -ésima y por T la suma de las F_i (que coincide también con la de las C_j). La región crítica del test de Cochran es

$$Q = \frac{r \sum F_i^2 - T^2}{rT - \sum C_j^2} (r - 1) > \chi_\alpha^2 (r - 1)$$

Los cálculos pueden simplificarse si eliminamos todos los individuos con respuesta homogénea a todos los tratamientos (todos 1's o todos 0's). El número t' de individuos que restan debe verificar que $rt' > 25$ para poder utilizar el test de Cochran. En el caso $r = 2$ el problema se reduce al de comparar dos proporciones relacionadas que ya ha sido resuelto; se comprueba que en ese caso el test de Cochran coincide con el test de McNemar allí obtenido, sólo que ahora las condiciones de validez del test son algo más exigentes que allí.

¿Verdadero o Falso? Capítulo 5

Decidir si cada una de las proposiciones siguientes es verdadera (V) o falsa (F), justificando la respuesta.

- 5.1 Si en un ANOVA para comparar tres medias μ_1 , μ_2 y μ_3 , el test F resulta significativo, al hacer las comparaciones múltiples puede ocurrir que μ_1 y μ_2 , por una parte, y μ_2 y μ_3 , por otra, sean declaradas homogéneas.
- 5.2 En el modelo de clasificación simple en análisis de la varianza, la variación “entre” dividida por sus grados de libertad constituye un estimador de la varianza común de las variables que observamos.
- 5.3 En un problema de comparación de cinco proporciones para muestras independientes el valor observado del estadístico de contraste es igual a 17.34. Podemos concluir entonces que el test es altamente significativo.
- 5.4 En un problema de comparación de 5 proporciones a partir de otras tantas muestras independientes, el valor del estadístico de contraste es igual a 14.1. El test resulta no significativo, pero deben señalarse indicios de significación.
- 5.5 En el modelo de clasificación simple en análisis de la varianza, si la parte de la variación total explicada por la diferencia entre los distintos niveles del factor aumenta y la parte de la variación total explicada por el azar se mantiene constante, entonces el nivel mínimo de significación del test F de homogeneidad de las medias disminuye.
- 5.6 En el modelo de clasificación simple en análisis de la varianza (o ANOVA de una vía), la suma de cuadrados “dentro” se interpreta como la parte de la variación total que queda explicada por el azar.
- 5.7 (★) El test de Kruskal-Wallis se aplica en problemas como los considerados en el modelo de clasificación simple en análisis de la varianza si la suposición de normalidad es cuestionable.

Problemas del Capítulo 5

- 5.1** Con el objetivo de comparar cuatro tratamientos oculares A, B, C y D, se seleccionan 24 conejos que se dividen en cuatro grupos de 6 unidades; cada grupo recibe un tratamiento diferente. Tras la aplicación de los tratamientos se determina para cada uno de los conejos una variable que mide el poder antiinflamatorio de los mismos y cuya distribución no puede suponerse normal, obteniéndose los siguientes datos:

A	2	3	3	3	3	0
B	1	3	1	2	2	3
C	3	1	2	1	3	3
D	1	0	0	0	0	-1

¿Pueden considerarse igualmente efectivos los tratamientos en lo que a la variable considerada se refiere? Calcula y comenta la probabilidad de significación.

- 5.2** Se desean comparar los efectos de tres fármacos A, B y C en la reducción de la fiebre; para ello se seleccionan 12 pacientes con una fiebre entre 38.5°C y 39°C y se dividen en tres grupos de 4, administrando un fármaco a cada grupo. Tras cuatro horas se anotaron las siguientes reducciones en la fiebre

Fármaco A	2.0	1.6	2.1	0.6
Fármaco B	0.5	1.2	0.3	0.2
Fármaco C	1.1	-1.0	-0.2	0.2

Asumiendo normalidad, analiza esos datos y decide si existen diferencias entre los tres fármacos, comentando el nivel mínimo de significación P .

- 5.3** En un estudio llevado a cabo en varones mayores de 14 años en la ciudad de Badajoz con el objetivo de estimar el porcentaje de bebedores de riesgo (BR) y su distribución por grupos de edad y por centro de salud (CS), se ha extraído una muestra de 856 individuos y se ha aplicado a cada uno de ellos una encuesta a partir de la cual se toma la decisión de declarar o no bebedor de riesgo al individuo haciendo uso de un

criterio elaborado por Altisent y otros. Los datos obtenidos se resumen en las tablas siguientes categorizados por grupos de edad y por centro de salud:

Edad	14 a 29	30 a 64	> 64	Total
BR	62	85	11	158
Total	268	440	148	856

CS	La Paz	S. Fernando	S. Roque	Anexo	Centro	Total
BR	47	50	15	36	10	158
Total	198	173	133	266	86	856

- a) ¿Existen diferencias significativas entre las proporciones de bebedores de riesgo entre los grupos de edad 14–29, 30–64 y > 64?
- b) Responde a una cuestión análoga a la anterior para los diferentes centros de salud considerados. **SOLUCIÓN: PÁG. 342**

5.4 Para decidir si existe una estrecha relación entre la presión sanguínea sistólica (PSS) y la ocurrencia de accidentes vasculares cerebrales (AVC), se lleva a cabo un estudio de 5 años de duración con 1000 varones con edad comprendida entre los 45 y los 60 años. La tabla siguiente contiene los datos obtenidos.

		(AVC)	
		SI	NO
(PSS)	70-80	5	245
	80-90	20	230
	90-100	30	220
	> 100	65	185

Efectúa un test apropiado para tomar esa decisión y facilita alguna medida de la posible relación entre las variables consideradas.

Capítulo **6**

Relación entre Variables

6.1. Relación entre dos Variables Cuantitativas: Regresión Lineal

6.1.1. El problema de regresión

Si en temas anteriores se ha considerado una única variable cualitativa o cuantitativa cuyos valores en una, dos o más poblaciones sugieren ciertos problemas de contraste de hipótesis de comparación, estamos ahora interesados en considerar dos variables en un mismo individuo y decidir si existe o no algún tipo de relación entre ellas, es decir, si los valores que tome alguna de ellas condicionan de algún modo los que tome la otra. En este tema abordaremos este problema para el caso de que esas variables sean cuantitativas. En realidad, la comparación de los valores de una misma variable en dos o más poblaciones distintas es también un estudio de relación entre dos variables: la variable que nos interesa y la variable cualitativa que asigna a cada individuo la población a la que pertenece. De hecho, el conocimiento científico avanza siempre estableciendo nuevas relaciones entre variables, y todos los métodos estadísticos estudiados y por estudiar pueden considerarse como instrumentos útiles a la hora de evaluar esas relaciones.

Para el estudio de la relación entre dos variables cuantitativas se usan en estadística dos técnicas principalmente: la técnica de regresión y la técnica de correlación. Ambas técnicas proporcionan métodos que pretenden detectar la existencia de asociación o dependencia entre las variables consideradas, pero la técnica de regresión nos permite incluso ir un poco más lejos y nos puede ser útil a la hora de estudiar cómo varía una de las variables observando la variación de la otra y a la hora de predecir los valores de una de las variables cuando se sabe que la otra ha tomado un cierto valor. Estudiaremos en este tema el problema de regresión lineal y en el próximo las técnicas de correlación.

Consideremos dos v.a. X e Y que miden ciertas características numéricas de interés sobre los individuos de una cierta población. Nuestro objetivo primero es el estudio de la posible dependencia entre esas dos variables. Ejemplos: 1) X podría ser el peso de un individuo e Y su altura; 2) X podría ser la edad de un individuo e Y su presión sanguínea sistólica; 3) X podría ser la longitud del brazo derecho de un individuo e Y la de su pierna derecha, . . . A menudo, X es una variable que el investigador puede controlar (edad del individuo, temperatura a la que se realiza el experimento, . . .) y su objetivo es predecir los valores de Y a partir de los de X ; en este caso, la técnica de correlación es cuestionable.

Podemos empezar preguntándonos si existe entre ambas variables una dependencia funcional, es decir, si existe una cierta función f de modo que $Y = f(X)$ pues, de ser así, una vez conocido X , no nos cabría duda alguna de cual sería el valor de Y . La respuesta a esa pregunta suele ser negativa debido a la presencia, además del valor de X , de otros muchos factores, controlables o no, que influyen en el valor de Y . Así las cosas, podemos plantearnos el problema de encontrar una función f de forma que $f(X)$ aproxime lo mejor posible a la variable Y ; de ese modo, conocido el valor x que toma la variable X , aproximando el valor de la variable Y por $f(x)$ haríamos todo cuanto está en nuestra mano para resolver el problema planteado. A una función f como ésa la llamaríamos curva de regresión de Y sobre X .

6.1.2. Regresión lineal

Determinar exactamente la curva de regresión de Y sobre X no suele ser un problema sencillo. Para simplificarlo se exige de la función f que sea de una forma determinada (una recta, una parábola, una función polinómica de grado k , . . .) y el problema se replantea, por ejemplo, como sigue: ¿cuál es la recta f tal que $f(X)$ aproxima lo mejor posible a Y ? Es decir, ¿cuáles son las constantes α, β tales que $\alpha + \beta X$ aproxima a Y lo mejor posible? Esas constantes son los llamados coeficientes de regresión; concretamente, β es la llamada pendiente de regresión y α es la ordenada en el origen de la recta de regresión. Queda claro que buscamos la mejor recta y que, por tanto, puede haber una función g (que no sea una recta) de forma que $g(X)$ aproxime a Y mejor que $\alpha + \beta X$. El caso de la recta es

especialmente importante y el problema considerado se llama problema de regresión lineal.

Observación 6.1. (Regresión no lineal) Cuando la nube de puntos obtenida a partir de los datos no sugiere una relación lineal entre las dos variables, sino una relación curvilínea, el modelo lineal de regresión no es apropiado. Debido a que el modelo lineal es más simple que otros modelos de regresión no lineal y es el mejor estudiado y el más sencillo de interpretar, los investigadores tienen tendencia a adoptarlo a costa de lo que sea; eso ha tenido como consecuencia los cambios de escala no lineales que en esencia pretenden obtener, no la regresión de Y sobre X , sino la de una cierta función de Y sobre una cierta función de X . Debe cuidarse en ese caso que el cambio de escala no afecte a las otras suposiciones del modelo de regresión. Este problema de linealización no será abordado en este texto. \triangleleft

Observación 6.2. (★) (Modelos de regresión múltiple) Si disponemos de una variable dependiente Y y varias variables independientes X_1, \dots, X_k , podemos estar interesados en determinar constantes $\beta, \alpha_1, \dots, \alpha_k$ de forma que $\beta + \sum_{i=1}^k \alpha_i X_i$ aproxime a Y lo mejor posible en el sentido de los mínimos cuadrados. Se trata del llamado problema de regresión lineal múltiple que, desde un punto de vista estadístico, supone que, dados valores concretos x_1, \dots, x_k de X_1, \dots, X_k , la variable Y tiene distribución normal de media $\beta + \sum_{i=1}^k \alpha_i x_i$ y varianza σ^2 , siendo $\beta, \alpha_1, \dots, \alpha_k$ y σ^2 los parámetros desconocidos del modelo de regresión múltiple. Este problema, que queda fuera del alcance de este libro, puede resultar de utilidad, por ejemplo, si queremos expresar la presión sanguínea en función de la edad y el peso del individuo. \triangleleft

En apoyo del caso lineal se encuentra el siguiente resultado: si la distribución del par (X, Y) es normal bidimensional, entonces la recta de regresión coincide con la curva de regresión, es decir, la mejor función f coincide con la mejor recta. La distribución normal bidimensional o bivalente es una generalización de la distribución normal ya conocida al caso de una distribución bidimensional. Su función de densidad es una función de dos variables que, representada en el espacio, tiene el aspecto de una campana tridimensional o sombrero de bombero, como en la figura siguiente. Esa distribución, además de las medias y varianzas $\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2$ de ambas variables, depende de un quinto parámetro que denotaremos por ρ y llamaremos coeficiente de correlación entre X e Y ; este coeficiente toma valores entre -1 y 1 y constituye una medida del grado de dependencia entre las variables X e Y .

Desde el punto de vista estadístico, la distribución del par (X, Y) es desconocida; la hipótesis de que ese par tiene distribución normal bidimensional queda justificada en muchos casos de modo análogo a como quedaba justificada en el caso univariante, lo que convierte el problema de regresión lineal en algo más interesante de lo que pudiera parecer de entrada.

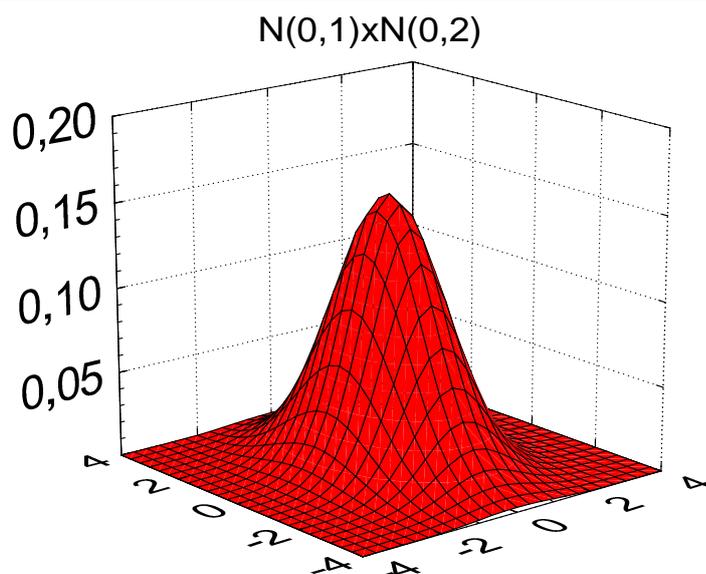


Figura 6.1: Distribución normal bivalente $N(0, 1) \times N(0, 2)$

En un problema de regresión, la terminología clásica propone llamar a la X variable independiente o controlada, y a Y , variable dependiente, aunque los papeles son, a menudo, reversibles y a veces podemos estar interesados en la recta de regresión de X sobre Y .

Desde un punto de vista descriptivo, la mejor manera de hacerse una idea sobre la distribución bidimensional desconocida del par (X, Y) consiste en obtener una muestra (x_i, y_i) , $1 \leq i \leq n$, de extensión n de esa distribución y dibujar esos puntos en el plano, para obtener lo que llamaremos una nube de puntos. La forma de esa nube de puntos en el plano puede sugerir una relación lineal entre X e Y , o una relación curvilínea, o una ausencia de relación (ver Figura 6.2).

El llamado principio de los mínimos cuadrados nos permite calcular inmediatamente la recta que mejor aproxima a esa nube de puntos: se trata en concreto de determinar las constantes a y b que hacen mínima la expresión

$$\sum_{i=1}^n (y_i - a - bx_i)^2$$

Las soluciones a este problema son los llamados coeficientes de la recta de regresión mínimo cuadrática o coeficientes de regresión muestrales (b es la pendiente de regresión muestral y

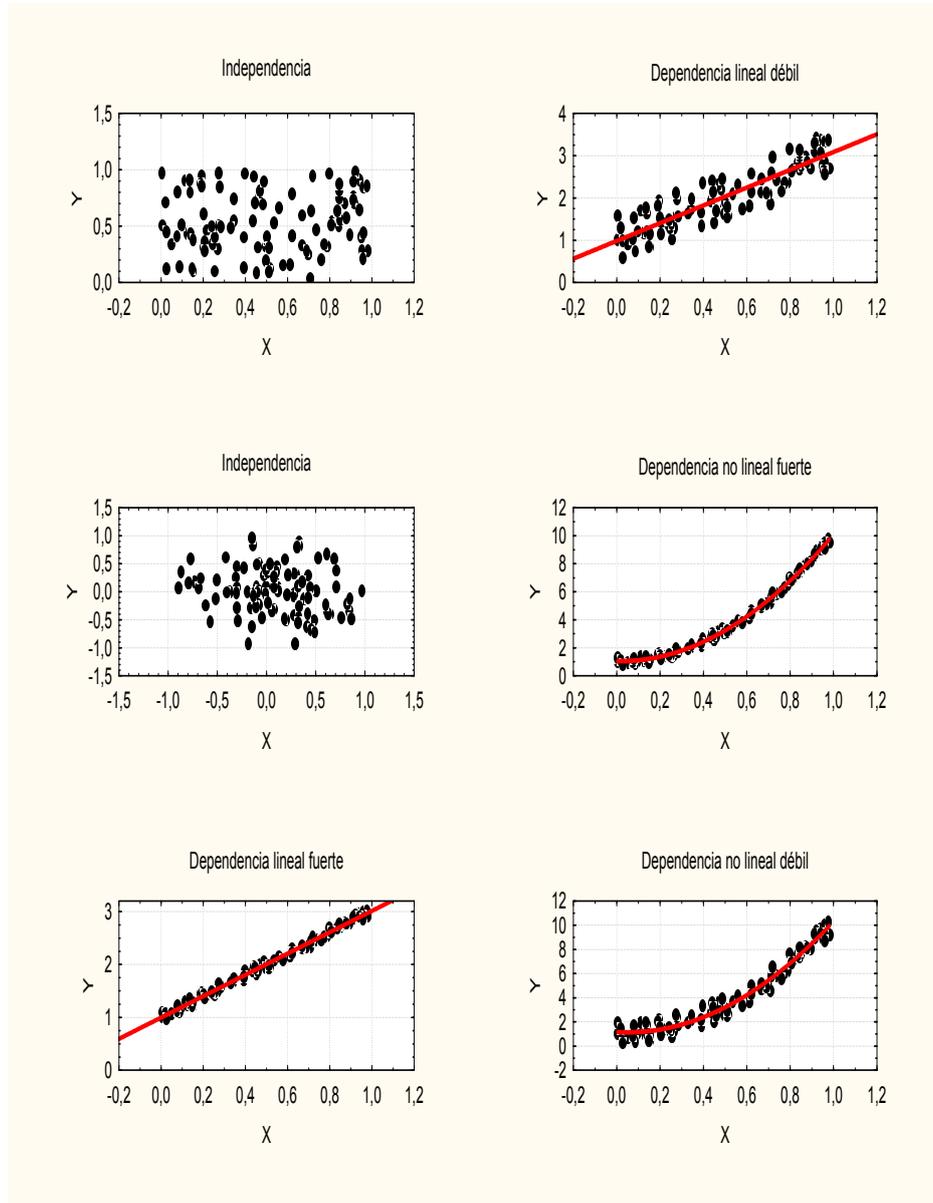


Figura 6.2: Nubes de puntos

a es la ordenada en el origen de la recta de regresión muestral) y son concretamente

$$b = \frac{S_{xy}}{S_{xx}}$$

$$a = \bar{y} - b\bar{x}$$

donde

$$S_{xy} = \sum_i (x_i - \bar{x})(y_i - \bar{y}) = \sum_i x_i y_i - \frac{1}{n} \left(\sum_i x_i \right) \left(\sum_i y_i \right)$$

y

$$S_{xx} = \sum_i (x_i - \bar{x})^2 = \sum_i x_i^2 - \frac{1}{n} \left(\sum_i x_i \right)^2$$

Análogamente, denotaremos en lo que sigue $S_{yy} = \sum_i (y_i - \bar{y})^2 = \sum_i y_i^2 - \frac{1}{n} (\sum_i y_i)^2$.

La ecuación de la recta de regresión muestral es, por tanto,

$$y - \bar{y} = b(x - \bar{x})$$

ecuación que pone de manifiesto que esta es una recta que tiene pendiente b y pasa por el punto (\bar{x}, \bar{y}) .

6.1.3. Estimación puntual en el modelo de regresión lineal

A parte del tratamiento descriptivo del problema que acabamos de considerar, interesa ver qué uso hace la inferencia estadística de los datos obtenidos para resolver los problemas de regresión antes mencionados. Para ello necesitamos describir previamente el modelo de regresión lineal, es decir, las suposiciones que admitiremos en los problemas de inferencia (estimación y test de hipótesis) que consideraremos más adelante.

El modelo de regresión lineal supone que la distribución de Y , una vez conocido que $X = x$, es una distribución normal univariante

$$N(\alpha + \beta x, \sigma^2)$$

con media $\alpha + \beta x$ dependiente linealmente de x y varianza σ^2 desconocida pero independiente del valor x de X ; los parámetros α y β son los llamados coeficientes de regresión poblacionales

y se suponen desconocidos, pues sólo tenemos acceso a una muestra de la población. El concepto matemático que define rigurosamente lo que acabamos de decir queda muy lejos del alcance de esta obra. También podemos describir el modelo diciendo que la distribución de Y dado $X = x$ es la de la variable

$$\alpha + \beta x + \varepsilon$$

donde ε es una variable con distribución normal $N(0, \sigma^2)$ que representa el error cometido en la aproximación de Y por $\alpha + \beta X$.

Con objeto de hacer inferencia estadística sobre la recta de regresión poblacional, se extrae una muestra de la distribución bidimensional. Es decir, se seleccionan al azar n individuos de la población para los que se obtienen las características X e Y ; a veces, cuando la variable X es controlable, también pueden fijarse previamente los valores de X y, para cada uno de esos valores, se obtienen uno o más valores de Y . En cualquier caso, tras realizar el experimento, dispondremos de n pares (x_i, y_i) , $1 \leq i \leq n$, en los que basaremos nuestras decisiones sobre los parámetros poblacionales desconocidos. Así, los coeficientes de regresión poblacionales α (ordenada en el origen poblacional) y β (pendiente poblacional) se estiman mediante los coeficientes de regresión muestrales a (ordenada en el origen muestral) y b (pendiente muestral) obtenidos anteriormente. Para estimar la varianza de regresión desconocida σ^2 tengamos en cuenta que, dado $X = x$, la v.a. $\varepsilon = Y - \alpha - \beta x$ sigue una distribución normal $N(0, \sigma^2)$, con lo cual, la varianza muestral obtenida de los residuos $d_i = y_i - a - bx_i$, $1 \leq i \leq n$, será un buen estimador de σ^2 ; como $\sum d_i = 0$, el estimador de la varianza de regresión σ^2 será

$$\begin{aligned} S^2 &= \frac{1}{n-2} \sum d_i^2 = \frac{1}{n-2} \sum (y_i - \hat{y}_i)^2 = \frac{1}{n-2} \sum [(y_i - \bar{y}) - b(x_i - \bar{x})]^2 \\ &= \frac{1}{n-2} [S_{yy} - 2bS_{xy} + b^2S_{xx}] = \frac{1}{n-2} [S_{yy} - bS_{xy}] \end{aligned}$$

donde se ha denotado $\hat{y}_i = a + bx_i$ (son las predicciones o valores de la recta de regresión en los x_i). El $n - 2$ del denominador (en lugar del $n - 1$ tradicional) queda justificado porque se han debido estimar dos parámetros α y β (en lugar del único tradicional μ por \bar{x}).

Ejemplo 6.3. Se ha llevado a cabo un estudio para analizar la influencia del consumo medio de sal en gramos por día (variable X) en la presión sanguínea sistólica (variable Y)

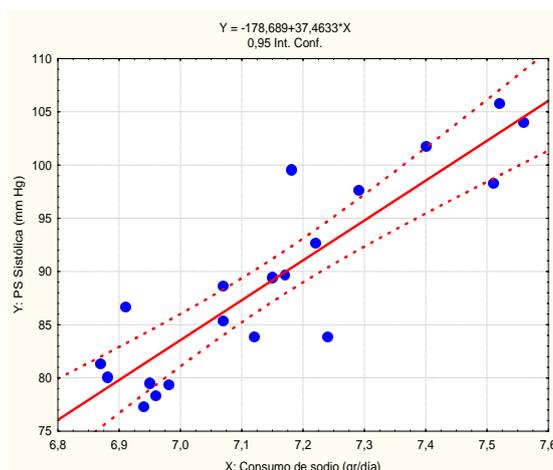


Figura 6.3: Recta de regresión de ecuación: $y = -178.69 + 37.46 \cdot x$

de los individuos adultos de una población. Para ello, se ha seleccionado una muestra de 20 individuos y se han obtenido los datos siguientes:

Ind.	1	2	3	4	5	6	7	8	9	10
x_i	6.96	7.51	7.17	7.15	6.94	6.91	7.22	6.98	7.07	6.88
y_i	78.37	98.29	89.67	89.42	77.32	86.70	92.67	79.36	85.34	80.06
\hat{y}_i	82.05	102.66	89.92	89.17	81.30	80.18	91.79	82.80	86.17	79.05
Ind.	11	12	13	14	15	16	17	18	19	20
x_i	6.87	7.24	7.12	7.07	7.29	6.95	7.56	7.40	7.52	7.18
y_i	81.33	83.88	83.86	88.65	97.63	79.50	103.97	101.73	105.79	99.53
\hat{y}_i	78.68	92.54	88.04	86.17	94.41	81.68	104.53	98.53	103.03	90.29

En este caso se tiene $\bar{x} = 143/20 = 7.15$, $\bar{y} = 1783/20 = 89.15$, $S_{xx} = 0.90$, $S_{xy} = 33.82$ e $S_{yy} = 1596.56$. La pendiente de la recta de regresión muestral es entonces $b = S_{xy}/S_{xx} = 33.82/0.9 = 37.46$ y la ordenada en el origen $a = \bar{y} - b\bar{x} = -178.689$. La varianza alrededor de la línea ajustada es $S^2 = \frac{1}{20-2}[1596.56 - 37.46 \cdot 33.83] = 18.31$. La pendiente b indica el aumento que se produce en la recta de regresión cuando el valor de X aumenta en una unidad; en nuestro caso, cabe pensar que, en término medio, un aumento de un gramo en el consumo diario de sal conlleva un incremento de 37.46 mm Hg en la presión sanguínea sistólica. Se ha denotado por \hat{y}_i el valor de la recta de regresión en el punto x_i , que llamaremos predicción del valor de la variable Y cuando X toma el valor x_i .

Nótese que en la Figura 6.3, con el fin de obtener una mejor visualización de la nube de puntos, se han utilizado diferentes escalas en los ejes de abscisas y ordenadas: uno de esos ejes comienza en el valor 6.8 y otro en 75 (ninguno de ellos en 0, como es habitual). Además, mientras que las abscisas aumentan de 0.1 en 0.1, las ordenadas aumentan de 5 en 5; como consecuencia se obtiene una falsa idea de la pendiente de la recta de regresión. Una imagen de la recta de regresión después de corregir esto aparece en la Figura 6.4, en la que, por el contrario, resulta más difícil distinguir unos puntos de otros.

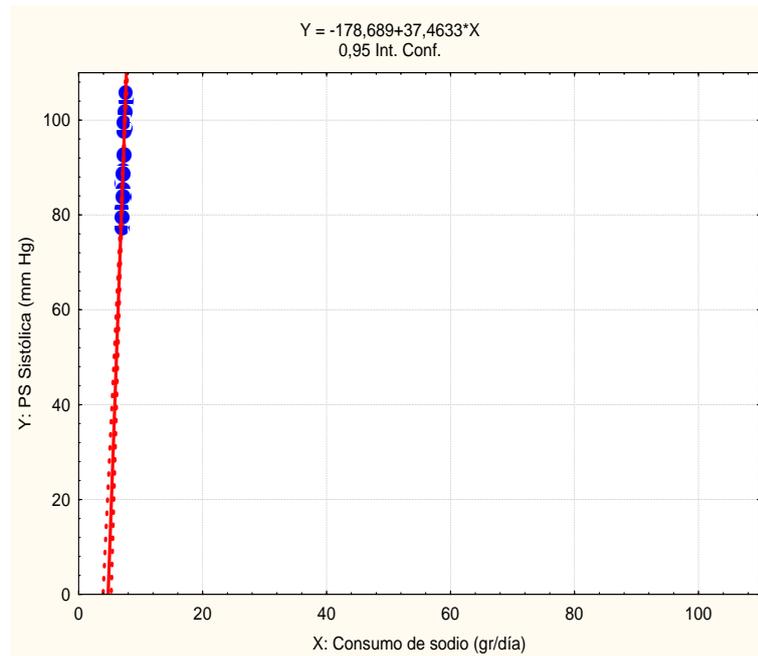


Figura 6.4: Aspecto real de la regresión de Y sobre X

En este sentido, a pesar de lo dicho en la observación de la página 34, la primera figura puede ser preferible a la segunda, debido al efecto lupa que ejerce sobre la nube de puntos y porque no altera la impresión fundamental sobre la relación entre ambas variables, es decir, una fuerte dependencia lineal directa entre ambas. \triangleleft

6.1.4. Contraste de hipótesis e intervalos de confianza sobre la pendiente de regresión

De la expresión $S^2 = \frac{1}{n-2}[S_{yy} - bS_{xy}]$ se obtiene la siguiente descomposición:

$$S_{yy} = bS_{xy} + (n-2)S^2 \quad (1)$$

Notar que, puesto que $\hat{y}_i = a + bx_i = \bar{y} - b\bar{x} + bx_i$, se tiene que

$$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \sum_{i=1}^n b^2 (x_i - \bar{x})^2 = b^2 S_{xx} = \frac{S_{xy}^2}{S_{xx}} S_{xx} = bS_{xy}$$

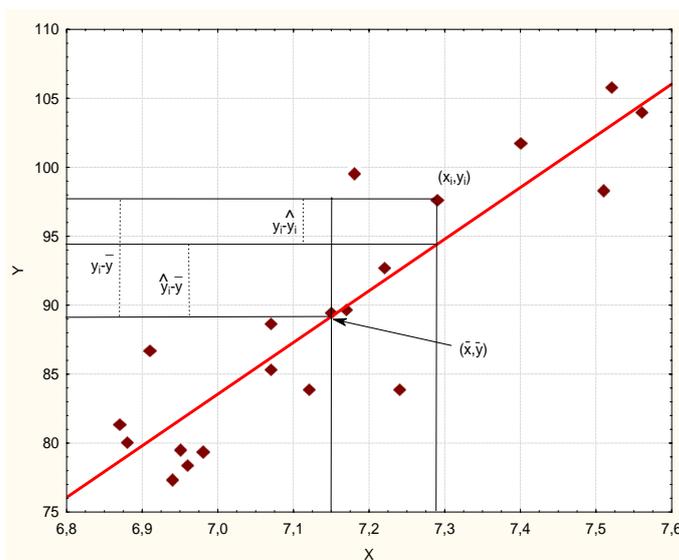


Figura 6.5: Análisis de la varianza para el modelo de regresión lineal

y (1) se puede escribir en la forma

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Esta descomposición de la variación total $S_{yy} = \sum (y_i - \bar{y})^2$ de la variable Y se conoce como análisis de la varianza para la regresión lineal. El primer sumando de esa descomposición es $bS_{xy} = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$ se interpreta como la parte de la variación total que queda explicada por la regresión lineal de Y sobre X y se llama suma de cuadrados debida a la regresión lineal. El segundo sumando $(n - 2)S^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2$ se interpreta como la parte de la variación total debida a la desviación de las observaciones respecto a la recta de regresión y se llama suma de cuadrados residual o error. Dicho de otro modo, ante la pregunta de por qué las y_i son distintas de la media \bar{y} , responderemos que por dos razones: porque las y_i tienden a aumentar con x (efecto debido a la regresión lineal) y porque en la variabilidad de Y , además del azar y otras variables no consideradas en el estudio, debemos considerar una posible dependencia no lineal de la variable X , y aparece, por tanto, una variabilidad alrededor de la recta de regresión (efecto debido a la desviación de la recta de regresión). Una medida interesante de la importancia de la v.a. X en la predicción lineal

de Y es el llamado coeficiente de determinación r^2 que estudiaremos en el próximo tema y queda definido por

$$r^2 = \frac{S_{xy}^2}{S_{xx}S_{yy}} = \frac{bS_{xy}}{S_{yy}} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Se trata de un número entre 0 y 1 que nos da la proporción de la variabilidad de Y que queda explicada por su regresión lineal en X ; cuanto más próximo a 1 esté r^2 , mejor predictor lineal de Y será la v.a. X .

En particular, si $b = 0$, la suma de cuadrados residual es la variación total y la regresión lineal en X no influye en la variación de Y . Si $b \neq 0$ los datos detectan la existencia de una cierta relación lineal entre ambas variables; esa relación lineal se dice directa si la pendiente muestral b es positiva (pues Y tiende a crecer cuando X crece) e inversa si b es negativa (pues Y tiende a decrecer cuando X crece). Nótese también que, si $\beta = 0$, la distribución de Y una vez conocido que $X = x$ es la distribución $N(\alpha, \sigma^2)$, que no depende del valor x de la v.a. X y, por tanto, Y es independiente de X .

Lo dicho en el párrafo anterior justifica que el problema de contrastar la hipótesis de independencia lineal de las variables X e Y consideradas contra la alternativa de que existe dependencia (o asociación) lineal entre ellas queda convertido en el problema de contrastar la hipótesis nula $H_0 : \beta = 0$ contra la alternativa $H_1 : \beta \neq 0$.

En general, para resolver el problema de contrastar la hipótesis nula $H_0 : \beta = \beta_0$, tendremos en cuenta que, dado $X = x$ y supuesta cierta la hipótesis nula, la pendiente de regresión muestral b tiene distribución $N(\beta_0, \sigma^2/S_{xx})$, por lo que la región crítica al nivel γ es

$$\frac{|b - \beta_0|}{S/\sqrt{S_{xx}}} > t_\gamma(n - 2)$$

En particular, la región crítica al nivel γ para el problema de contrastar la hipótesis de independencia lineal de las dos variables $H_0 : \beta = 0$ contra la alternativa de dependencia lineal $H_1 : \beta \neq 0$ es

$$\frac{|b|}{S/\sqrt{S_{xx}}} > t_\gamma(n - 2)$$

Para establecer un intervalo de confianza para la pendiente de regresión β al $100 \cdot (1 - \gamma) \%$ de confianza, tendremos en cuenta que, dado $X = x$, b tiene distribución normal $N(\beta; \sigma^2/S_{xx})$;

por tanto, el intervalo buscado es:

$$\beta \in b \pm t_\gamma(n-2) \frac{S}{\sqrt{S_{xx}}}$$

Ejemplo 6.4. (Ejemplo 6.3, continuación) Continuando con el ejemplo considerado anteriormente, con una confianza del 95 % podemos afirmar que

$$\beta \in 37.46 \pm 2.101 \sqrt{18.31/0.9} = [27.63, 47.29]$$

Por otra parte, para contrastar la hipótesis nula $H_0 : \beta = 0$ de independencia lineal entre las variables consumo de sodio y PSS calculamos el valor experimental

$$t_{\text{exp}} = \frac{|37.46|}{\sqrt{18.31/0.9}} = 8.31$$

que, comparado con el valor crítico $t_{0.05}(20-2) = 2.101$, resulta significativo (el valor P es ≤ 0.001 , pues $t_{0.001}(20-2) = 3.922$). Debemos concluir entonces que consumo de sal y presión sanguínea sistólica están relacionadas. \triangleleft

Observación 6.5. (★) Los programas estadísticos para ordenador suelen incorporar diversos métodos estadísticos que permiten verificar si los datos disponibles son compatibles con las suposiciones que se hacen en el modelo de regresión lineal; en particular, nos permitirán comprobar la normalidad de los residuos y la suposición de que la varianza no depende del valor de la variable independiente (que se suele conocer como homocedasticidad). \triangleleft

6.1.5. Inferencia sobre la media de Y dado un valor de X y sobre la ordenada en el origen

Ejemplo 6.6. (Ejemplo 6.3, continuación) Cabe preguntarse por la presión sanguínea sistólica (PSS) media de los adultos que consumen 7 gramos de sal por día. Si no tuviéramos en cuenta la relación lineal entre Y y X , tendríamos que tomar una muestra de individuos adultos que consuman 7 gramos de sal por día y obtener, a partir de ella, una estimación (puntual y/o por intervalos) de la PSS media. Haciendo uso de la relación lineal entre Y y X podemos dar esa estimación aunque en la muestra seleccionada no haya ningún individuo que consuma 7 gramos de sal por día. \triangleleft

De forma general, supuesto conocido que X toma el valor x , pretendemos estimar el valor $\alpha + \beta x$. Evidentemente, su estimación puntual es $\hat{y} = a + bx$. Para obtener un intervalo de confianza para esa media, utilizaremos que, dado $X = x$, $a + bx$ tiene distribución

$$N \left(\alpha + \beta x, \sigma^2 \left(\frac{1}{n} + \frac{(x - \bar{x})^2}{S_{xx}} \right) \right)$$

Por tanto, un intervalo de confianza para la media $\alpha + \beta x$ al nivel $1 - \gamma$ es

$$\alpha + \beta x \in a + bx \pm t_\gamma(n-2)S\sqrt{\frac{1}{n} + \frac{(x-\bar{x})^2}{S_{xx}}} \quad (2)$$

En particular, para $x = 0$ se obtiene un intervalo de confianza para la ordenada en el origen α :

$$\alpha \in a \pm t_\gamma(n-2)S\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}}$$

Observación 6.7. Variando x en (2), los extremos de los correspondientes intervalos determinan en el plano dos curvas que delimitan una banda alrededor de la recta de regresión. A medida que nos alejamos de la media \bar{x} en uno u otro sentido, la banda se abre más y más; por ello, en regresión, las inferencias sobre el valor medio de Y para $X = x$ son tanto más precisas cuanto más cerca de la media \bar{x} esté x . \triangleleft

Análogamente, para contrastar la hipótesis nula $H_0 : \alpha + \beta x = h$ contra $H_1 : \alpha + \beta x \neq h$ al nivel γ , utilizaremos la región crítica

$$\frac{|a + bx - h|}{S\sqrt{\frac{1}{n} + \frac{(x-\bar{x})^2}{S_{xx}}}} > t_\gamma(n-2)$$

En particular, tomando $x = 0$, la región crítica para contrastar $H_0 : \alpha = \alpha_0$ contra $H_1 : \alpha \neq \alpha_0$ al nivel γ es

$$\frac{|a - \alpha_0|}{S\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}}} > t_\gamma(n-2)$$

6.1.6. Intervalos de predicción

En muchas ocasiones el experimento de regresión se planifica para predecir la variable dependiente Y a partir de un cierto valor x de la variable X . Ya sabemos que la predicción sugerida para Y por la recta de regresión cuando $X = x$ es

$$\hat{y} = a + bx$$

Cabe preguntarse por la precisión con que podemos hacer esa afirmación. Aprovechando el ejemplo que venimos considerando, la respuesta a esta cuestión depende de si queremos predecir la PSS de un determinado individuo a partir de su consumo de sodio por día o

si estamos interesados en la PSS media de todos los individuos que consumen una cierta cantidad de sodio por día. Este último caso ya ha sido estudiado en la sección anterior; la desviación entre la estimación $\hat{y} = a + bx$ y la media $\alpha + \beta x$ de Y dado $X = x$ es $\hat{y} - \alpha - \beta x$. Para el primer caso, la desviación entre la estimación $\hat{y} = a + bx$ y el valor y de Y es $\hat{y} - y$; se prueba que, condicionalmente en $X = x$, la distribución de la variable $\hat{y} - Y = a + bx - \alpha - \beta x - \epsilon$ es

$$N\left(0; \sigma^2 \left(1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{S_{xx}}\right)\right)$$

Por tanto, un intervalo de predicción para y al $100 \cdot (1 - \gamma)\%$ de confianza es

$$y \in (a + bx) \pm t_{\gamma}(n - 2)S\sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{S_{xx}}}$$

Observación 6.8. Hablamos de intervalo de predicción y no de intervalo de confianza puesto que el valor y de la variable Y no es un parámetro desconocido de la distribución de Y , pero la interpretación que debemos dar es análoga a la de un intervalo de confianza. \triangleleft

Observación 6.9. Igual que ocurría con los intervalos de confianza para la media de Y dado un valor de X , la longitud del intervalo de predicción aumenta a medida que x se aleja de la media \bar{x} . Por tanto, hacer predicciones a partir de la recta de regresión cuando x está muy lejos de \bar{x} es peligroso, puesto que las predicciones pueden ser muy poco precisas. Como regla general, debería evitarse hacer predicciones para valores de X fuera del rango de sus valores observados. Así, en el ejemplo que venimos considerando sobre la relación entre el consumo de sal y la PSS, el modelo de regresión predice para un consumo nulo de sal una PSS igual a -178.69 , sin sentido que queda explicado por el hecho de que hemos realizado una predicción para un consumo nulo de sal utilizando un modelo de regresión que ha sido construido a partir de individuos que consumen entre 6.87 y 7.56 gr/día. \triangleleft

Ejemplo 6.10. (Ejemplo 6.3, continuación) En el ejemplo que consideramos en este tema, un intervalo de predicción al 95% de confianza para la PSS de un adulto que consume siete gr/día de sal es

$$-178.69 + 37.46 \cdot 7 \pm 2.101\sqrt{18.31 \left(1 + \frac{1}{20} + \frac{(7 - 7.15)^2}{0.9}\right)}$$

es decir, $[74.21, 92.85]$. Un adulto que consuma 7 gr/día de sal cuya PSS quede fuera de ese intervalo será declarado no normal (su PSS es anormalmente grande o anormalmente pequeña). La interpretación de ese intervalo de predicción es la misma que la que hemos dado en otras ocasiones para intervalos de confianza, es decir: supongamos que la experiencia de extraer una muestra de tamaño n de adultos con un consumo de sal entorno a los 7 gr/día

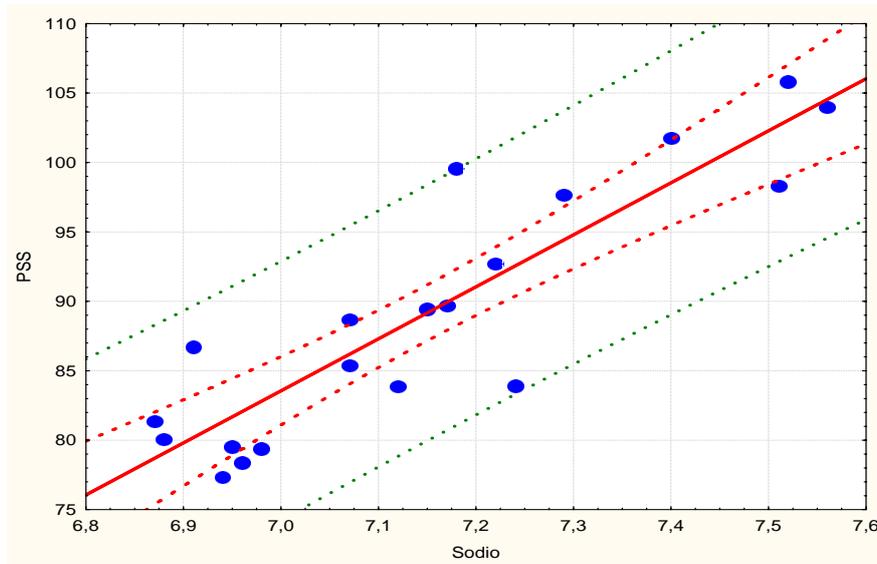


Figura 6.6: Intervalo de confianza (rayas) y e intervalo de predicción (puntos)

de sal (los valores mínimo y máximo en la muestra obtenida son 6,87 y 7,56) se realiza un número grande de veces; debemos pensar que para cada una de esas muestras (x_i, y_i) , $1 \leq i \leq n$, se obtiene un intervalo de confianza diferente y el 95% de esos intervalos contienen al valor y que mide la PSS del adulto con un consumo de 7 gr/día que nos interesa. La Figura 6.6 contiene el intervalo de confianza para la media de Y dado $X = x$ y el intervalo de predicción del valor de Y dado $X = x$. \triangleleft

6.2. Relación entre dos Variables Cuantitativas: Correlación

6.2.1. Coeficiente de correlación de Pearson y coeficiente de determinación

Ya hemos advertido que el estudio de la posible asociación lineal entre dos variables cuantitativas se puede llevar a cabo con la técnica de regresión lineal descrita anteriormente. Una medida del grado de asociación lineal entre ellas es el coeficiente de correlación (que también nos servirá para realizar un test de independencia, según veremos).

Buscando una medida del grado de asociación entre dos variables cuantitativas, podría pensarse en la pendiente b de la recta de regresión muestral, pues es ella quién determina la posible dependencia lineal entre X e Y . Pero b posee dos inconvenientes: 1) queda afectado por un cambio lineal de escala (si las unidades de medida de Y pasan de centímetros a metros, b queda dividido por 100) y 2) la pendiente de la recta de regresión de Y sobre X es en general distinta de la pendiente de la recta de regresión de X sobre Y , y deseamos una medida simétrica del grado de asociación entre ambas variables. Consideremos el llamado coeficiente de determinación muestral

$$r^2 = \frac{S_{xy}^2}{S_{xx}S_{yy}}$$

que es simétrico en ambas variables y no queda afectado por cambios lineales de escala. Nótese también que

$$r^2 = \frac{S_{xy}^2/S_{xx}}{S_{yy}} = \frac{bS_{xy}}{S_{yy}}$$

es decir, r^2 es el cociente entre la variabilidad explicada por la regresión de Y sobre X y la variación total de Y , o bien, r^2 es la fracción de la variabilidad de Y que queda explicada por su dependencia lineal de la variable X (en tanto que $1 - r^2$ es la fracción de variabilidad

de Y no explicada por su dependencia lineal con X , y puede deberse a la influencia de otras variables no consideradas en el problema o a una dependencia no lineal entre ambas). Como r^2 no conserva la información sobre el signo de S_{xy} (que era también el signo de la pendiente b), se llama coeficiente de correlación muestral o coeficiente de correlación de Pearson la cantidad

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$

cantidad que sí recoge el signo de S_{xy} , indicando dicho signo que el valor de Y crece (si r positivo) o decrece (si r negativo) cuando X crece. Se deduce de lo dicho en el tema anterior que el coeficiente de correlación muestral r toma valores entre -1 y 1 y que $(n - 2)S^2 = \sum(y_i - \hat{y}_i)^2 = S_{yy}(1 - r^2)$, con lo cual, si $r = \pm 1$, los valores y_i coinciden con sus predicciones y la nube de puntos está sobre la recta de regresión y la dependencia es lineal; en el caso opuesto, $r = 0$ y $(n - 2)S^2 = S_{yy}$ (que es lo máximo posible), la regresión de Y sobre X no rebaja en nada la variabilidad total de Y y las dos variables serán linealmente independientes.

Para el ejemplo del consumo de sal y la presión sanguínea sistólica considerado en el tema anterior, se verifica que $r = 33.82/\sqrt{0.9 \cdot 1596.56} = 0,8921$ y $r^2 = 0,796$, lo que se interpreta diciendo que el 79.6 % de la variabilidad de la PSS de los adultos de la población en cuestión queda justificada por su asociación lineal con el consumo de sal, y el 20.4 % restante depende de otros factores (ambiente, alimentación, etc) no contemplados en el estudio, del azar y de una posible relación no lineal entre las dos variables.

Ya habíamos hablado anteriormente del coeficiente de correlación de la distribución bidimensional de (X, Y) ; es el coeficiente de correlación poblacional ρ (o simplemente coeficiente de correlación) el cual es desconocido desde el punto de vista estadístico y del que el coeficiente de correlación muestral r es un estimador.

En un problema particular, el investigador estará interesado en determinar si $\rho = 0$ (variables linealmente independientes) o $\rho \neq 0$ (variables dependientes), pero el verdadero valor de ρ es desconocido. Para ello realizaremos un test de hipótesis para contrastar $H_0 : \rho = 0$ contra $H_1 : \rho \neq 0$. Recordemos que $b = S_{xy}/S_{xx}$ y $r = S_{xy}/\sqrt{S_{xx}S_{yy}}$; por tanto, $b = 0$ si y sólo si $r = 0$. Del mismo modo se verifica que $\beta = 0$ si y sólo si $\rho = 0$, de

ahí que el test para contrastar $H_0 : \rho = 0$ sea equivalente al test de independencia que ya habíamos estudiado para contrastar $\beta = 0$. Modificaremos el estadístico de contraste para que quede en términos de r . El cuadrado del estadístico de contraste considerado en el caso de regresión es $t^2 = b^2 S_{xx}/S^2$; sustituyendo b y S^2 por sus valores se tiene

$$t^2 = \frac{S_{xy}^2/S_{xx}^2}{S_{yy} - S_{xy}^2/S_{xx}} S_{xx}(n-2)$$

y dividiendo numerador y denominador por S_{yy} se obtiene que

$$t^2 = (n-2)r^2/(1-r^2)$$

La región crítica al nivel α para el problema de contrastar la hipótesis de independencia lineal $H_0 : \rho = 0$ contra $H_1 : \rho \neq 0$ es

$$|t| = |r| \sqrt{\frac{n-2}{1-r^2}} > t_\alpha(n-2)$$

Observación 6.11. (★) (Correlaciones espúreas) En estudios prospectivos en busca de posibles relaciones entre variables conviene tener presente algunos hechos. Una correlación alta entre dos variables no debe interpretarse sin más como una estrecha relación causa-efecto de una variable sobre otra, a menos que haya una teoría previamente establecida y avalada por estudios precedentes. Podría ocurrir que la presencia de una tercera variable no contemplada en el estudio estuviese distorsionando las conclusiones del mismo o que las suposiciones de independencia, normalidad, . . . que hemos hecho para el modelo de regresión lineal hayan sido seriamente violadas. Consideremos el siguiente ejemplo aclaratorio: en un estudio sobre alumnos de segundo de bachillerato se detectó una correlación negativa significativa entre la estatura y la longitud del cabello, de tal suerte que podemos terminar malinterpretando los resultados diciendo que el hecho de mantener el pelo corto aumenta la altura. Un análisis más detallado que tenga en cuenta el sexo del alumno puede aclarar las cosas, habida cuenta de que sí parece de acuerdo con nuestra experiencia que las chicas suelen ser más bajas y llevar el pelo largo. La primera figura del gráfico siguiente, realizada con datos ficticios, pretende ilustrar lo que acabamos de decir: en el grupo de los chicos (puntos redondos) una recta de regresión horizontal indica una correlación prácticamente nula en ese grupo; algo análogo ocurre en el grupo de las chicas (puntos cuadrados). Las conclusiones que obtenemos al incluir la variable sexo en el estudio está más acorde con nuestra experiencia que el caso en que el sexo no es tenido en cuenta, en el que aparece una recta de regresión con una marcada pendiente negativa que indica una correlación negativa significativa difícil de explicar. Se suele hablar de correlación espúrea entre dos variables para referirnos a una correlación que, lejos de estar justificada por una relación causa-efecto, es fundamentalmente debida a la dependencia común de una tercera variable (en nuestro caso, el sexo).

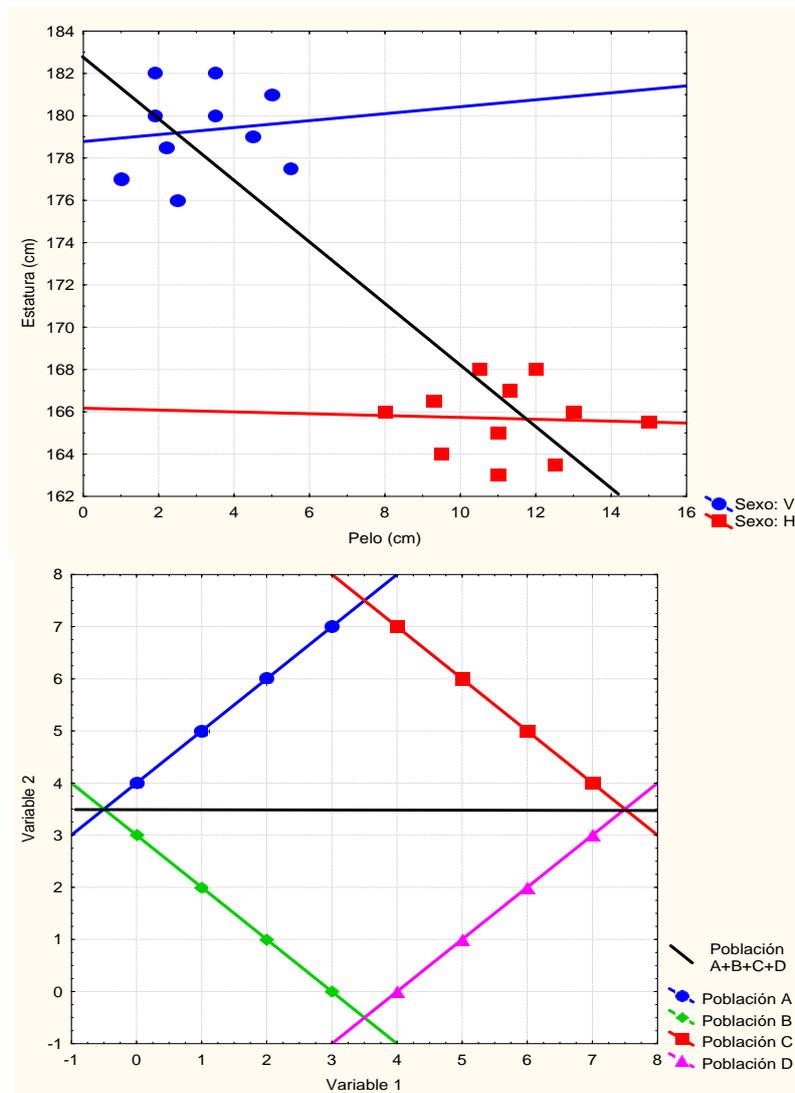


Figura 6.7: Correlaciones espúreas

La presencia inadvertida de dos o más poblaciones en la muestra puede tener el efecto de convertir en significativa una dependencia prácticamente inexistente (como acabamos de ver), o el efecto contrario de convertir en no significativas relaciones que pudieran ser de interés: mientras que en la muestra global la recta de regresión es prácticamente horizontal, que indicaría una ausencia de relación entre las dos variables consideradas, la dependencia entre ellas sería significativa si ajustamos rectas de regresión a las nubes de puntos correspondientes a cada una de las poblaciones presentes en la muestra, como se ve en la segunda figura del gráfico anterior, en el que la gruesa recta horizontal corresponde a la muestra global (donde se ha ignorado que hay cuatro poblaciones A, B, C y D en la muestra), mientras que las 4 rectas más finas, que señalan una dependencia significativa entre las dos variables, corresponden a cada una de las cuatro poblaciones involucradas.

Por otra parte, cuando aparecen involucradas un gran número de variables y estudiamos sistemáticamente la correlación entre cada dos de ellas, puesto que estamos dispuestos a asumir una probabilidad de error de tipo I del 5%, lo que viene a decir que una de cada 20 veces que realicemos el experimento de estudiar la relación entre dos variables incorreladas obtendremos por error una correlación significativamente no nula, puede ocurrir que se obtengan por azar algunas correlaciones significativamente no nulas; por ejemplo, si disponemos de un total de 20 variables y estudiamos la relación entre cada dos de ellas, tendremos que calcular un total de $20 \cdot 19/2 = 190$ coeficientes de correlación, entre los cuales, en el caso de que todos ellos fuesen nulos (cosa que, desde un punto de vista estadístico, desconocemos), pueden aparecer entre 9 y 10 correlaciones estadísticamente significativas por azar. \triangleleft

6.2.2. (★) Correlación no paramétrica: Coeficiente de correlación de Spearman

El coeficiente de correlación descrito anteriormente, y los métodos estadísticos presentados para él, funcionan bien bajo la hipótesis de normalidad; lejos de esa hipótesis puede ser preferible utilizar una alternativa no paramétrica. En la literatura se han propuesto diversas medidas no paramétricas de la correlación entre dos variables. La más usada es el coeficiente de correlación de Spearman. Dada una muestra (x_i, y_i) , $1 \leq i \leq n$, de tamaño n de la distribución conjunta bidimensional de dos variables X e Y , se asignan rangos R_1, \dots, R_n a las x 's y rangos R'_1, \dots, R'_n a las y 's como se explica en la sección 12.5. Se define el coeficiente de correlación de Spearman r_s como el coeficiente de correlación calculado para los n pares

de rangos (R_i, R'_i) , es decir,

$$r_s = \frac{\sum_{i=1}^n (R_i - \bar{R})(R'_i - \bar{R}')}{\sqrt{\sum_{i=1}^n (R_i - \bar{R})^2 \sum_{i=1}^n (R'_i - \bar{R}')^2}} = \frac{S_{RR'}}{\sqrt{S_{RR}S_{R'R'}}$$

Una forma más simple de calcular r_s , cuando no hay empates, viene dada por la expresión

$$r_s = 1 - \frac{6}{n(n^2 - 1)} \sum_{i=1}^n (R_i - R'_i)^2$$

Si hay empates (digamos, g grupos de empates en las x 's con t_1, \dots, t_g datos, respectivamente, y h grupos de empates en las y 's con s_1, \dots, s_h datos, respectivamente), entonces

$$r_s = \frac{(n^3 - n) - \frac{1}{2} \sum_{i=1}^g (t_i^3 - t_i) - \frac{1}{2} \sum_{i=1}^h (s_i^3 - s_i) - 6 \sum_{i=1}^n (R_i - R'_i)^2}{\sqrt{[n^3 - n - \sum_{i=1}^g (t_i^3 - t_i)][n^3 - n - \sum_{i=1}^h (s_i^3 - s_i)]}}$$

Para contrastar la hipótesis nula H_0 de independencia de X e Y , Spearman propone la región crítica

$$|r_s| > r_\alpha(n)$$

donde los valores $r_\alpha(n)$ se obtienen de una tabla especial (Tabla XIII) para tamaños de muestra n inferiores a 50 a los niveles de significación usuales. Para tamaños de muestra grandes (la aproximación se considera aceptable si $n \geq 10$) podemos utilizar el estadístico

$$t_s = r_s \sqrt{\frac{n-2}{1-r_s^2}}$$

obtenido del análogo para el caso normal, pero reemplazando r por r_s . Su distribución bajo la hipótesis nula de independencia es aproximadamente $t(n-2)$; rechazaremos la hipótesis nula cuando

$$|t_s| > z_\alpha$$

6.3. Relación entre dos Variables Cualitativas

6.3.1. Test de independencia de dos variables cualitativas

En esta sección abordaremos el problema de la relación entre dos variables cualitativas, también llamadas caracteres; por ejemplo, ¿el color de los ojos es independiente del grupo sanguíneo? Para ello los datos se suelen presentar en una tabla de contingencias. Un ejemplo nos ayudará a plantear el problema.

Ejemplo 6.12. Se ha llevado a cabo un estudio en una población con 7306 individuos elegidos al azar para ver si el grupo sanguíneo y el color de los ojos (se distinguen 3 categorías: azul, marrón y otros) están relacionados. La tabla siguiente recoge los resultados obtenidos:

		Grupo sanguíneo				Totales
		O	A	B	AB	
Color de ojos	Azul	$O_{11} = 920$	$O_{12} = 558$	$O_{13} = 84$	$O_{14} = 48$	$F_1 = 1610$
	Marrón	$O_{21} = 1817$	$O_{22} = 1711$	$O_{23} = 336$	$O_{24} = 223$	$F_2 = 4087$
	Otros	$O_{31} = 752$	$O_{32} = 678$	$O_{33} = 125$	$O_{34} = 54$	$F_3 = 1609$
Totales		$C_1 = 3489$	$C_2 = 2947$	$C_3 = 545$	$C_4 = 325$	$T = 7306$

Cantidades observadas

△

En general, una tabla de contingencias $r \times s$ clasifica los T individuos de una muestra con arreglo a los r niveles de una característica cualitativa A y a los s niveles de otra característica cualitativa B . Denotaremos por O_{ij} (O por valor observado) el número de individuos de la muestra que se encuentran en el nivel i del carácter A y en el nivel j del carácter B , F_i (F por fila) la suma de los O_{ij} , $1 \leq j \leq s$ y por C_j (C por columna) la suma de los O_{ij} , $1 \leq i \leq r$.

La cuestión que nos interesa es si el nivel en que se encuentre (o valor que tome) el carácter A condiciona a (o es condicionado por) el nivel en que se encuentre el carácter B o, por el contrario, ambos caracteres son independientes. La hipótesis nula H_0 puede enunciarse de diversos modos:

- Los caracteres A y B son independientes (o A y B no están relacionados, o también, A y B no están asociados).
- El nivel en que se encuentre el carácter A no condiciona el nivel en que se encuentra el carácter B .
- Las proporciones de individuos que caen en el nivel j del carácter B para los diferentes niveles del carácter A coinciden, y eso para cada $1 \leq j \leq s$.
- O, lo que equivalente, las proporciones de individuos que caen en el nivel i del carácter A para los diferentes niveles del carácter B coinciden, y eso para cada $1 \leq i \leq r$.

Matemáticamente, si denotamos por A_i el suceso “el carácter A está en el nivel i (o toma el valor i)” y B_j el suceso “el carácter B está en el nivel j ”, se trata de contrastar

$$H_0 : P(A_i \cap B_j) = P(A_i)P(B_j), \text{ para todo par } i, j$$

contra

$$H_1 : P(A_i \cap B_j) \neq P(A_i)P(B_j), \text{ para algún par } i, j$$

Si la hipótesis nula de independencia se supone cierta, el valor esperado E_{ij} para la casilla (i, j) de la tabla será aproximadamente igual a $T \cdot (F_i/T) \cdot (C_j/T) = F_i \cdot C_j/T$, pues F_i/T es una estimación de $P(A_i)$ y C_j/T es una estimación de $P(B_j)$. Así pues, si la hipótesis nula es cierta, los valores observados O_{ij} deberían estar próximos a los valores esperados E_{ij} . Usaremos el estadístico de contraste

$$\chi^2 = \sum_i \sum_j \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \sum_i \sum_j \frac{O_{ij}^2}{E_{ij}} - T$$

para medir la discrepancia entre los valores observados y los esperados, que tiene bajo la hipótesis nula, para T grande, distribución aproximada chi-cuadrado con $(r - 1) \cdot (s - 1)$ grados de libertad que venimos denotando por $\chi^2((r - 1) \cdot (s - 1))$.

Observación 6.13. (Condiciones de validez) Para que el test χ^2 descrito pueda aplicarse se exigen en la práctica las siguientes condiciones: ninguna de las cantidades esperadas E_{ij} debe ser menor que 1 (pues el sumando correspondiente daría una contribución excesiva al estadístico de contraste) y no más del 20% de ellas deben ser inferiores a 5. \triangleleft

La hipótesis nula de independencia de los dos caracteres se rechazará para valores grandes del estadístico de contraste. Concretamente, la región crítica del test es

$$\chi^2 = \sum_i \sum_j \frac{(O_{ij} - E_{ij})^2}{E_{ij}} > \chi_\alpha^2((r - 1) \cdot (s - 1))$$

donde el valor crítico $\chi_\alpha^2((r - 1) \cdot (s - 1))$ es el punto que deja a su derecha un área igual a α en la distribución χ^2 con $(r - 1) \cdot (s - 1)$ grados de libertad.

Ejemplo 6.14. (Ejemplo 6.12, continuación) La tabla siguiente, junto a las cantidades observadas O_{ij} , contiene entre paréntesis las cantidades esperadas correspondientes E_{ij} , con una precisión de 1 cifra decimal:

		Grupo sanguíneo				
		O	A	B	AB	Totales
Color de ojos	Azul	920(768,9)	558(649,4)	84(120,1)	48(71,6)	1610
	Marrón	1817(1951,8)	1711(1648,6)	336(304,9)	223(181,8)	4087
	Otros	752(768,4)	678(649)	125(120)	54(71,6)	1609
Totales		3489	2947	545	325	7306

Cantidades observadas y esperadas

Se deduce de ello que el valor experimental del test χ^2 es

$$\chi^2 = \sum_{i=1}^3 \sum_{j=1}^4 \frac{O_{ij}^2}{E_{ij}} - T = 91.6$$

Es claro que se verifican las condiciones de validez del test. Puesto que $\chi_{0.05}^2((3-1) \cdot (4-1)) = 12.59$, los datos son incompatibles con la hipótesis nula. Siendo $\chi_{0.001}^2((3-1) \cdot (3-1)) = 22.46$, debemos señalar diferencias altamente significativas entre los 4 grupos sanguíneos en lo que a la distribución del color de ojos se refiere, es decir, debemos rechazar la hipótesis de independencia entre el color de ojos y el grupo sanguíneo y aceptar que existe asociación entre ellos.

La tabla siguiente muestra los porcentajes por filas de individuos que caen en cada casilla, con el fin de visualizar las diferencias entre las distribuciones del grupo sanguíneo para los diferentes colores de ojos considerados.

		Grupo sanguíneo				
		O	A	B	AB	Totales
Color de ojos	Azul	57.1 %	34.7 %	5.2 %	3 %	100 %
	Marrón	44.5 %	41.9 %	8.2 %	5.5 %	100 %
	Otros	46.7 %	42.1 %	7.8 %	3.4 %	100 %

Porcentajes por filas

Acabamos de ver que las diferencias que se observan entre los porcentajes de una misma columna son altamente significativos. Análogamente se puede construir una tabla de porcentajes por columnas con el fin de visualizar las diferencias, altamente significativas, entre las distribuciones del color de ojos entre los individuos de un mismo grupo sanguíneo:

		Grupo sanguíneo			
		O	A	B	AB
Color de ojos	Azul	26.4 %	18.9 %	15.4 %	14.8 %
	Marrón	52.1 %	58.1 %	61.7 %	68.6 %
	Otros	21.6 %	23 %	22.9 %	16.6 %
Total		100 %	100 %	100 %	100 %

Porcentajes por columna

Finalmente, con fines descriptivos, tiene interés la observación de una tabla de porcentajes referidos al total de individuos que participan en el estudio, para visualizar la distribución conjunta de las dos variables cualitativas consideradas:

		Grupo sanguíneo				
		O	A	B	AB	Totales
Color de ojos	Azul	12.6 %	7.6 %	1.1 %	0.7 %	22 %
	Marrón	24.9 %	23.4 %	4.6 %	3.1 %	55.9 %
	Otros	10.3 %	9.3 %	1.7 %	0.7 %	22 %
		47.8 %	40.3 %	7.5 %	4.4 %	100 %

Distribución conjunta del grupo sanguíneo y el color de ojos

A partir de esta tabla se puede construir un diagrama de barras bidimensional en el que sobre una tabla 3×4 se levantan prismas sobre cada casilla de altura proporcional al porcentaje de la misma. \triangleleft

Observación 6.15. En el caso de una tabla de contingencias 2×2 , el estadístico de contraste toma la forma

$$\chi^2 = \frac{(O_{11}O_{22} - O_{12}O_{21})^2}{F_1 F_2 C_1 C_2} \cdot T$$

y debe compararse con el valor $\chi_a^2(1)$. Las condiciones de validez que suelen exigirse en el caso 2×2 son las siguientes: las F_i y las C_j deben ser mayores que $T/10$ y todas las E_{ij} deben ser mayores que 5. \triangleleft

Observación 6.16. (★) El test χ^2 de comparación de varias muestras cualitativas que presentamos en el epígrafe 5.2, en el que se pretendía comparar una variable cualitativa en varias poblaciones, es un caso particular de éste: la segunda variable cualitativa es la que asigna a cada individuo su población. Es conveniente, no obstante, tomar ciertas precauciones en este caso. No hay diferencia alguna entre estos dos tests si se seleccionan al azar T individuos de las poblaciones consideradas y se les clasifica en la tabla de acuerdo a los valores de la variable cualitativa de interés y la población a la que pertenece. Sin

embargo, si se seleccionan muestras de tamaños F_1, \dots, F_r elegidos por el investigador en las poblaciones $1, \dots, r$ que queremos comparar (así se ha hecho en el Tema), la hipótesis nula debe formularse del siguiente modo: “ H_0 : la distribución de la variable cualitativa B es la misma en las r poblaciones consideradas”; es decir, una vez construida la tabla de porcentajes por filas, la hipótesis nula viene a decir que no existen diferencias significativas entre los porcentajes que aparecen en una misma columna. No podríamos, en este caso, construir la tabla de porcentajes por columnas y formular la hipótesis nula diciendo que no existen diferencias significativas entre los porcentajes que figuran en una misma fila pues esos porcentajes por columnas no nos permiten estimar el porcentaje de individuos de cada población en cada nivel del carácter B , ya que los tamaños muestrales en cada población han sido elegidos arbitrariamente por el investigador. Por ejemplo, supongamos que, para estudiar una posible asociación entre el color de los ojos y el desprendimiento de retina, se ha llevado a cabo un estudio con 100 individuos con desprendimiento de retina y otros 100 sin esa afección, y que la tabla siguiente recoge los resultados de la experiencia:

	Azul	Marrón	Otros	Totales
Desprendimiento retina	32	51	17	100
No desprendimiento retina	37	40	23	100
Totales	69	91	40	200

Nótese que la tabla observada coincide en este caso con la tabla de porcentajes por fila, pues éstas suman 100. Si p_{ar}, p_{mr}, p_{or} denotan las proporciones de individuos con desprendimiento de retina que tienen los ojos de color azul, marrón u otros, resp., y p_{an}, p_{mn}, p_{on} denotan las correspondientes proporciones para los individuos sin desprendimiento de retina, la hipótesis nula debe formularse en la forma: $H_0 : p_{ar} = p_{an}, p_{mr} = p_{mn}, p_{or} = p_{on}$. La correspondiente tabla de porcentajes por columnas es:

	Azul	Marrón	Otros
Desprendimiento retina	46.4	56	42.5
No desprendimiento retina	53.6	44	57.5
Totales	100	100	100

Sin embargo, no debemos interpretar que un 46.4% de los individuos con ojos azules tiene desprendimiento de retina, pues el investigador ha decidido seleccionar 100 individuos con desprendimiento de retina y otros 100 sin desprendimiento de retina, sin que de ahí se pueda colegir que el 50% de los individuos de la población tiene desprendimiento de retina. Por tanto, si denotamos por p_a, p_m, p_o las proporciones de individuos con ojos de color azul, marrón y otros, resp., no es apropiado utilizar esa tabla para contrastar la hipótesis nula $H_0 : p_a = p_m = p_o$. Véase también la observación 6.28; una situación similar se encuentra en el ejemplo 1.39. Una precaución similar debemos tener cuando lo que selecciona el investigador son los totales marginales por columnas C_1, \dots, C_s . ◀

6.3.2. Medidas de asociación de dos variables cualitativas

Si el test χ^2 resulta significativo, admitiremos que existe una cierta asociación (dependencia) entre los caracteres considerados. En ese caso puede resultar de interés determinar la fuerza con que la asociación está presente. Con ese objetivo, diversas medidas de asociación han sido propuestas en la literatura; aquí recogemos brevemente algunas de ellas, especialmente en el caso 2×2 .

El coeficiente de contingencia de Pearson para una tabla de contingencias $r \times s$ se define a partir del estadístico de contraste χ^2 mediante la siguiente fórmula

$$C = \sqrt{\frac{\chi^2}{T + \chi^2}}$$

que toma valores entre 0 (asociación nula) y $\sqrt{\frac{q-1}{q}}$ (asociación máxima) donde $q = \min(r, s)$.

Ejemplo 6.17. (Ejemplo 6.12, continuación) En el ejemplo considerado se verifica que el coeficiente de contingencia de Pearson es igual a 0.111; el valor máximo que podría haber alcanzado es $\sqrt{2/3} = 0.816$, media que sugiere una débil asociación entre ambas variables cualitativas. \triangleleft

Observación 6.18. La cantidad χ^2 no constituye por sí misma una buena medida de la asociación entre los dos caracteres considerados debido a que queda afectado por cambios de escala; de hecho, al multiplicar los datos originales por una constante K , el estadístico de contraste queda también multiplicado por esa misma constante. Por tanto, la magnitud del estadístico de contraste χ^2 sólo indica evidencia de asociación, pero no es una buena medida de la misma. Tampoco lo pequeño que sea el valor P del test χ^2 mide el grado de asociación entre los caracteres. \triangleleft

Una medida de asociación muy popular en las ciencias del comportamiento para tablas de contingencias 2×2 es el coeficiente phi que se define por

$$\phi = \sqrt{\frac{\chi^2}{T}}$$

donde $\chi^2 = \frac{(O_{11}O_{22} - O_{12}O_{21})^2}{F_1 F_2 C_1 C_2} \cdot T$. Este coeficiente es un número comprendido entre 0 (asociación nula) y 1 (asociación máxima).

6.3.3. Otras medidas de asociación en el caso 2×2 :

Riesgo relativo y razón del producto cruzado

Definimos a continuación otras medidas de asociación para tablas 2×2 muy utilizadas en investigaciones epidemiológicas. Consideraremos, concretamente, el problema de medir el grado de asociación entre un factor de riesgo (descrito por una v.a. X) y una enfermedad (descrito por una v.a. Y) y supondremos que ambas variables son dicotómicas: X toma el valor 1 (respectivamente, el valor 2) si el factor de riesgo está presente (respectivamente, no está presente) en el individuo elegido; Y toma el valor 1 (respectivamente, el valor 2) si el individuo elegido ha desarrollado la enfermedad (respectivamente, no la ha desarrollado). Por ejemplo, la variable X puede describir si el individuo es o no fumador y la variable Y nos dirá si ha desarrollado o no un cáncer de pulmón. Denotaremos por p_{ij} la probabilidad de que $X = i$ e $Y = j$, $i, j = 1, 2$, en la población total. Denotaremos también $p_{i+} = p_{i1} + p_{i2}$ la probabilidad de que $X = i$, $i = 1, 2$, y $p_{+j} = p_{1j} + p_{2j}$ la probabilidad de que $Y = j$, $j = 1, 2$. Podemos distinguir, además de las consideradas anteriormente, las siguientes medidas de asociación entre el factor de riesgo y la enfermedad:

DEFINICIÓN 6.19. (Diferencia de riesgos) Esta medida de asociación se define como

$$D = P(Y = 1|X = 1) - P(Y = 1|X = 2) = \frac{p_{11}}{p_{1+}} - \frac{p_{21}}{p_{2+}}$$

Observación 6.20. La diferencia de riesgos mide el incremento de tasa de enfermedad atribuible al factor de riesgo, o la proporción esperada de individuos que no desarrollarán la enfermedad si el factor de riesgo fuese eliminado. Toma valores entre -1 y 1 , correspondiendo el valor cero a la independencia entre la enfermedad y el factor de riesgo. \triangleleft

Ejemplo 6.21. Si, entre los individuos con el factor de riesgo, el 20% desarrollan la enfermedad, mientras que entre los que no poseen el factor de riesgo la desarrollan el 5%, la diferencia de riesgos es igual a 0.15, lo que indica que la presencia del factor de riesgo produce un 15% de aumento en el desarrollo de la enfermedad. \triangleleft

DEFINICIÓN 6.22. (Riesgo relativo) El riesgo relativo, que se define como

$$R = \frac{P(Y = 1|X = 1)}{P(Y = 1|X = 2)}$$

Observación 6.23. Por tanto, R es el cociente entre la probabilidad de padecer la enfermedad cuando el factor de riesgo está presente y la probabilidad de padecerla cuando el

factor de riesgo no está presente. Toma valores entre 0 y $+\infty$, correspondiendo el valor 1 a la independencia entre la enfermedad y el factor de riesgo. \triangleleft

Ejemplo 6.24. (Continuación del ejemplo 6.21) En el ejemplo anterior, $R = 4$, lo que indica que entre los individuos que poseen el factor de riesgo la probabilidad de desarrollar la enfermedad es 4 veces superior que entre los que no lo poseen. \triangleleft

DEFINICIÓN 6.25. (Razón del producto cruzado -en inglés, odds ratio-) Esta medida de asociación se define por

$$OR = \frac{\frac{P(Y=1|X=1)}{P(Y=2|X=1)}}{\frac{P(Y=1|X=2)}{P(Y=2|X=2)}} = \frac{p_{22}p_{11}}{p_{12}p_{21}}$$

Observación 6.26. La razón del producto cruzado representa la razón entre las proporciones de individuos enfermos ($Y = 1$) y no enfermos ($Y = 2$) cuando el factor de riesgo está presente ($X = 1$) frente a la razón entre las proporciones de individuos enfermos y no enfermos cuando no lo está ($X = 2$). A diferencia del riesgo relativo y la diferencia de riesgos, ésta es una medida simétrica respecto a los papeles desempeñados por X e Y , pues también

$$OR = \frac{\frac{P(X=1|Y=1)}{P(X=2|Y=1)}}{\frac{P(X=1|Y=2)}{P(X=2|Y=2)}}$$

Toma valores entre 0 y $+\infty$, correspondiendo el valor 1 a la independencia entre la enfermedad y el factor de riesgo. \triangleleft

Ejemplo 6.27. (Continuación del ejemplo 6.21) En el ejemplo que venimos considerando, $OR = 4.75$, lo que indica que, entre los individuos que poseen el factor de riesgo, la razón entre la probabilidad de desarrollar la enfermedad y la de no desarrollarla es 4.75 veces superior a la que se obtiene entre los individuos que no poseen el factor de riesgo. \triangleleft

6.3.4. Inferencia sobre el riesgo relativo y la razón del producto cruzado

Para hacer inferencia estadística sobre las medidas de asociación introducidas en el epígrafe anterior, se lleva a cabo un estudio prospectivo, es decir, se seleccionan dos muestras independientes de tamaños F_1 y F_2 de individuos con el factor de riesgo y sin él, respectivamente, y se determina después cuántos de ellos desarrollan la enfermedad y cuántos no.

Los datos obtenidos se representan en una tabla 2×2 como la siguiente:

	$Y = 1$	$Y = 2$	Total
$X = 1$	O_{11}	O_{12}	F_1
$X = 2$	O_{21}	O_{22}	F_2
Total	C_1	C_2	T

donde T es el número total de individuos estudiados, O_{11} es el número de individuos que poseen el factor de riesgo y han desarrollado la enfermedad, etc.

Como estimador de la diferencia de riesgos poblacional utilizaremos la diferencia de riesgos muestral

$$\hat{D} = \frac{O_{11}}{F_1} - \frac{O_{21}}{F_2}$$

Más utilizadas como medidas de asociación en el caso 2×2 son el riesgo relativo y la razón del producto cruzado, cuyos estimadores respectivos son el riesgo relativo muestral, definido por

$$\hat{R} = \frac{\frac{O_{11}}{F_1}}{\frac{O_{21}}{F_2}} = \frac{O_{11}F_2}{O_{21}F_1},$$

y la razón del producto cruzado muestral, definida por

$$\widehat{OR} = \frac{\frac{O_{11}/F_1}{O_{12}/F_1}}{\frac{O_{21}/F_2}{O_{22}/F_2}} = \frac{O_{11}O_{22}}{O_{21}O_{12}}$$

Se verifica que la varianza del logaritmo (neperiano) del riesgo relativo muestral

$$\log \hat{R} = \log O_{11} + \log F_2 - \log O_{21} - \log F_1$$

se puede estimar por

$$S_{\log \hat{R}}^2 = \frac{O_{12}}{O_{11}F_1} + \frac{O_{22}}{O_{21}F_2}$$

y, bajo la hipótesis nula de no asociación (en cuyo caso el riesgo relativo es igual a 1 y su logaritmo igual a cero), el estadístico

$$\frac{(\log \hat{R})^2}{S_{\log \hat{R}}^2}$$

tiene, para muestras grandes, distribución aproximada $\chi^2(1)$. Por tanto, la región crítica al nivel α para el problema de contrastar la hipótesis nula $H_0 : R = 1$ de no asociación entre factor de riesgo y enfermedad contra la hipótesis alternativa $H_1 : R \neq 1$ es

$$\frac{(\log \hat{R})^2}{S_{\log \hat{R}}^2} > \chi_\alpha^2(1)$$

Del mismo modo, un intervalo al $100 \cdot (1 - \alpha) \%$ de confianza para el logaritmo del riesgo relativo es

$$\log R \in \log \hat{R} \pm z_\alpha S_{\log \hat{R}}$$

con lo que

$$e^{\log \hat{R} \pm z_\alpha S_{\log \hat{R}}}$$

se utilizará como intervalo de confianza para R .

Por otra parte, el logaritmo de la razón del producto cruzado se estima por

$$\log \widehat{OR} = \log \frac{O_{11}O_{22}}{O_{21}O_{12}}$$

y como estimador de su varianza utilizaremos

$$S_{\log \widehat{OR}}^2 = \frac{1}{O_{11}} + \frac{1}{O_{12}} + \frac{1}{O_{21}} + \frac{1}{O_{22}}$$

La región crítica al nivel α (aproximadamente para muestras grandes) para el problema de contrastar la hipótesis nula $H_0 : OR = 1$ de no asociación entre factor de riesgo y enfermedad contra la hipótesis alternativa $H_1 : OR \neq 1$ es

$$\frac{(\log \widehat{OR})^2}{S_{\log \widehat{OR}}^2} > \chi_\alpha^2(1)$$

y un intervalo al nivel de confianza $1 - \alpha$ para OR es

$$e^{\log \widehat{OR} \pm z_\alpha S_{\log \widehat{OR}}}$$

Observación 6.28. Además de los estudios prospectivos, se distinguen otros dos tipos de estudios: los retrospectivos y los cruzados. En los retrospectivos -o estudios caso control-, se extraen muestras independientes de tamaños C_1 y C_2 de individuos que han desarrollado la enfermedad y que no la han desarrollado, respectivamente, y el investigador mira en cada una de esas dos muestras cuántos individuos poseen el factor de riesgo y cuántos no. En

estudios cruzados se elige una muestra de tamaño T de la población general y sus individuos se clasifican en una de las 4 clases determinadas por los valores de las variables X e Y . Los argumentos de la observación 6.16.2 desaconsejan la estimación de la diferencia de riesgos y del riesgo relativo en estudios retrospectivos. Sin embargo, la razón del producto cruzado puede ser estimada en cualquiera de los 3 tipos de estudio citados.

¿Verdadero o Falso? Capítulo 6

Decidir si cada una de las proposiciones siguientes es verdadera (V) o falsa (F), justificando la respuesta.

- 6.1 El riesgo relativo es una medida simétrica respecto a los papeles desempeñados por el factor de riesgo y la enfermedad.
- 6.2 Una recta de regresión horizontal denota ausencia de dependencia lineal entre las dos variables consideradas.
- 6.3 La longitud del intervalo de predicción en un problema de regresión lineal aumenta a medida que el valor de la variable independiente se aleja de la media.
- 6.4 En el modelo de regresión lineal, la suma de cuadrados residual dividida por $n - 2$ es un estimador de la varianza de regresión.
- 6.5 La suma de cuadrados $\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$ en el modelo de regresión lineal se interpreta como la parte de la variación total de Y explicada por el azar.
- 6.6 En un estudio sobre la relación entre dos variables cuantitativas X e Y medidas en un mismo individuo, ¿es posible que el coeficiente de correlación muestral r y la pendiente de regresión muestral b verifiquen que $r \cdot b > 0$?
- 6.7 La suma de cuadrados $\sum_{i=1}^n (y_i - \hat{y}_i)^2$ en el modelo de regresión lineal se interpreta como la parte de la variación total de Y explicada por la regresión lineal en X .
- 6.8 La igualdad $\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2$ se conoce como análisis de la varianza para el modelo de regresión lineal.
- 6.9 Si la razón del producto cruzado (odds ratio) toma el valor 1, no existe asociación entre el factor de riesgo y la enfermedad.
- 6.10 Si la recta de regresión entre dos variables cuantitativas es decreciente, el coeficiente de correlación entre esas dos variables es negativo.

- 6.11** Si $y = a + bx$ es la ecuación de la recta de regresión muestral entre la variable dependiente Y =“talla” y la variable independiente X =“edad” calculada a partir de una muestra de niños entre 3 y 9 meses de edad, entonces a es la predicción que podemos hacer para la talla de un niño de tres meses de edad.
- 6.12** En el modelo de regresión lineal, la suma de cuadrados residual $(n - 2)S^2$ queda explicada exclusivamente por el azar.
- 6.13** (★) El coeficiente de correlación de Spearman para una muestra (x_i, y_i) , $1 \leq i \leq n$, de la distribución de dos variables cuantitativas X e Y es el coeficiente de correlación de Pearson calculado a partir de los pares (R_i, S_i) , $1 \leq i \leq n$, siendo R_i el rango de x_i y S_i el de y_i .
- 6.14** Cuando vale 1 el coeficiente de determinación r^2 para una muestra (x_i, y_i) , $1 \leq i \leq n$, de la distribución de dos variables cuantitativas X e Y , todos los puntos (x_i, y_i) están alineados.
- 6.15** Si la diferencia de riesgos toma el valor 1, no existe asociación entre el factor de riesgo y la enfermedad.
- 6.16** El riesgo relativo es el cociente entre las probabilidades de padecer la enfermedad cuando el factor de riesgo está presente y cuando no.
- 6.17** Dadas dos v.a.r. X e Y , la recta de regresión de Y sobre X es la función de X que mejor aproxima a Y en el sentido de los mínimos cuadrados.
- 6.18** Si en una tabla de contingencias $r \times s$ el test χ^2 para contrastar la independencia dos variables cualitativas resulta significativo, admitiremos que hay diferencias significativas entre las cantidades marginales por filas F_1, \dots, F_r .
- 6.19** Si el riesgo relativo para el estudio de la relación entre un factor de riesgo y una enfermedad es igual a 1.4, el porcentaje de individuos con el factor de riesgo que desarrollarán la enfermedad es un 40% superior al correspondiente a los individuos sin el factor de riesgo.

- 6.20** En un estudio en niños sobre la relación entre el hecho de haber recibido una determinada vacuna y padecer en los meses siguientes una invaginación intestinal, se ha facilitado una odds ratio igual a 1.1. Podemos concluir, entonces, que el cociente entre las probabilidades de sufrir una invaginación y no sufrirla es un 10% (es decir, 1.1 veces) superior en la población de niños vacunados que en la de no vacunados.

Problemas del Capítulo 6

- 6.1** Las presiones sanguíneas de 20 matrimonios cuyos miembros tienen edades entre 25 y 34 años elegidos al azar son recogidas en la tabla siguiente:

Matrimonio	1	2	3	4	5	6	7	8	9	10
X: Marido	136	121	128	100	110	116	127	150	180	172
Y: Mujer	110	112	128	106	127	100	98	142	143	150
Matrimonio	11	12	13	14	15	16	17	18	19	20
X: Marido	156	98	132	142	138	126	124	137	160	125
Y: Mujer	135	115	125	130	132	146	127	128	135	110

a) Contrasta la hipótesis de independencia lineal entre las presiones sanguíneas de las mujeres y la de sus maridos.

b) ¿Qué porcentaje de la variabilidad de las presiones sanguíneas de las mujeres queda explicada por su regresión lineal en X ?

c) Si un nuevo matrimonio entre esas edades se considera en el estudio y se sabe que el marido tiene presión sanguínea igual a 140, ¿qué predicción puede hacerse para la presión sanguínea de la mujer? Da un intervalo de predicción al 95% de confianza. Indicación: Utiliza que $\bar{x} = 133.9$, $\bar{y} = 124.95$, $\sum_{i=1}^{20} (x_i - \bar{x})^2 = 8923.8$, $\sum_{i=1}^{20} (x_i - \bar{x})(y_i - \bar{y}) = 4271.9$, $\sum_{i=1}^{20} (y_i - \bar{y})^2 = 4412.95$.

- 6.2** El volumen expiratorio forzado (VEF) es una medida estándar de la función pulmonar. A partir del VEF de 655 chicos de entre 10 y 15 años se obtuvieron los siguientes datos que relacionan la altura en cm con el VEF medio medido en litros:

Altura (X)	134	138	142	146	150	154
VEF Medio (Y)	1.7	1.9	2.0	2.1	2.2	2.5
Altura (X)	158	162	166	170	174	178
VEF Medio (Y)	2.7	3.0	3.1	3.4	3.8	3.9

(a) Determina la recta de regresión de Y sobre X .

(b) Contrasta la hipótesis de independencia lineal entre ambas variables.

(c) ¿Qué proporción de variabilidad del VEF medio queda explicado por su dependencia lineal de la altura? Nota: $\sum x_i = 1872$, $\sum x_i^2 = 294320$, $\sum y_i = 32.3$, $\sum y_i^2 = 93.11$, $\sum x_i y_i = 5156.2$.

- 6.3** En un estudio sobre una muestra de 264 personas con enfermedad depresiva unipolar se han registrado el sexo y si ha habido alguna hospitalización debida a la enfermedad, obteniéndose los siguientes resultados:

	Hosp. sí	Hosp. no	Total
Hombre	48	75	123
Mujer	32	109	141
Total	80	184	264

a) ¿Existe alguna asociación entre las dos variables? Ind.: Respóndase utilizando como medida de asociación la razón del producto cruzado.

b) Facilita una estimación y un intervalo de confianza para esa medida de asociación.

SOLUCIÓN: PÁG. 342

- 6.4** Para decidir si la concentración de amoníaco en sangre aumenta durante su almacenaje a cuatro grados centígrados, se han seleccionado diversas muestras de sangre y se han registrado para cada una de ellas el número de días (X) que llevan almacenadas y su concentración de amoníaco (Y) en $\mu\text{g}/\text{dL}$. La tabla siguiente contiene los resultados obtenidos, aunque debe tenerse en cuenta que a la muestra de sangre con 0 días de almacenamiento se le determinó su concentración de amoníaco 2 horas después de haber sido almacenada.

Días (X)	0	1	3	4	5	5	20	21	21	21
Amoníaco (Y)	128	167	244	185	194	214	685	770	930	620

a) Dibuja la nube de puntos y, por inspección, decidir si la relación entre las variables X e Y puede ser aproximada por una línea recta.

- b) Ejecuta un test apropiado para responder estadísticamente a la cuestión planteada en a), y cuantifica de algún modo la relación entre esas dos variables.
- c) Estima la tasa de incremento de concentración de amoníaco y proporciona un intervalo de confianza al 95 % para ella.
- d) Se ha determinado que el nivel de amoníaco en la sangre recién extraída es prácticamente nula. ¿Los datos anteriores son compatibles con esta afirmación? Explica el resultado.
- e) Utiliza los datos precedentes para predecir el contenido de amoníaco a los 10 días de almacenamiento, y facilita un intervalo al 95 % para esa predicción.
- f) Estima el nivel medio de amoníaco en muestras de sangre que llevan 10 días almacenadas y facilita un intervalo al 95 % de confianza para él.

6.5 La tabla siguiente contiene el porcentaje de reticulocitos (variable X) y el número de linfocitos por mm^2 (variable Y) de 9 pacientes con aplasia medular.

X	3,6	2	0,3	0,3	0,2	3	0	1	2,2
Y	1700	3078	1820	2706	2086	2299	676	2088	2013

- (a) Calcula el coeficiente de determinación entre Y y X e interprétalo. (b) A partir de ese valor, decide si la relación lineal entre Y y X es estadísticamente significativa; calcula y comenta el valor P .

6.6 Se ha llevado a cabo un estudio sobre 1000 niños para estudiar la posible relación entre el grado de cumplimiento de su calendario de vacunaciones y el nivel socio-cultural de sus padres. La tabla de contingencias 3×3 siguiente recoge los datos obtenidos:

		Cumplimiento			
		Alto	Medio	Bajo	Totales
Nivel	Alto	65	23	15	103
	Medio	318	147	10	475
	Bajo	170	171	81	422
Totales		558	346	96	1000

¿Qué conclusiones se pueden obtener de esos datos? Proporciona una medida de asociación entre ambas variables. **SOLUCIÓN: PÁG. 342**

¿Verdadero o Falso? SEGUNDA PARTE

Decidir si cada una de las proposiciones siguientes es verdadera (V) o falsa (F), justificando la respuesta.

- II.1** El nivel mínimo de significación de un test no depende de los datos obtenidos.
- II.2** Si X es una v.a. con distribución binomial $B(n, p)$ de parámetro desconocido p , entonces X/n es un estimador insesgado de mínima varianza de p .
- II.3** Si en el problema de comparación de dos medias de dos distribuciones normales la hipótesis nula es $H_0 : \mu_1 \leq \mu_2$, la hipótesis nula se rechazará siempre que la media muestral de la primera muestra sea estrictamente mayor que la de la segunda.
- II.4** Si el test de comparación de dos medias de dos distribuciones normales resulta no significativo a un nivel de significación $\alpha \in]0, 1[$, entonces sigue siendo no significativo para cualquier nivel de significación inferior a α .
- II.5** Se dispone de una muestra de tamaño n de una distribución normal de varianza conocida σ_0^2 y media desconocida μ . Si $[a, b]$ es un intervalo de confianza para μ al nivel de confianza $1 - \alpha$ y $] - \infty, c]$ es un intervalo de confianza unilateral para μ al mismo nivel de confianza $1 - \alpha$ entonces $b < c$.
- II.6** La longitud del intervalo de confianza para la media de una distribución normal aumenta con el nivel de confianza que deseamos para el mismo.
- II.7** El nivel mínimo de significación de un test es la probabilidad de obtener un valor del estadístico de contraste tan extremo o menos que el realmente obtenido supuesta cierta la hipótesis nula.
- II.8** La longitud del intervalo de confianza para la media de una distribución normal de varianza conocida aumenta con la varianza.
- II.9** El nivel mínimo de significación de un test es la probabilidad de obtener bajo la hipótesis nula un valor del estadístico de contraste tan o más extremo que el realmente obtenido.

- II.10** Un intervalo de confianza para un parámetro desconocido es una estimación conjuntista del mismo y, como tal, sus extremos, además de depender de los datos, pueden depender del parámetro.
- II.11** Un test de hipótesis es un método estadístico que nos permite decidir si se acepta o no la hipótesis nula formulada a partir de los datos obtenidos de la experimentación.
- II.12** Un estadístico es una función de los datos o, lo que es lo mismo, una aplicación definida en una estructura estadística.
- II.13** Dado un intervalo de confianza al nivel de confianza $1 - \alpha$ para la media desconocida μ , entenderemos que a lo sumo un $100 \cdot \alpha \%$ de las veces que realicemos el experimento obtendremos un intervalo que deja fuera a la media μ .
- II.14** En un problema de contraste de hipótesis la decisión final depende de las suposiciones que se hacen en la estructura estadística de partida.
- II.15** Un intervalo al 95 % de confianza para un parámetro es un intervalo que contiene el 95 % de los valores posibles del parámetro.
- II.16** (★) Se usa el test de rangos con signo de Wilcoxon para muestras relacionadas para una suma de rangos igual a 121 basado en 17 diferencias no nulas. El valor P del test verifica $0.01 < P < 0.05$.
- II.17** Consideremos el problema de contrastar la hipótesis $\mu \leq 0$ contra $\mu > 0$ a partir de una muestra de una distribución normal de media y varianza desconocidas. Siempre que la media muestral sea negativa aceptaremos la hipótesis nula.
- II.18** En un problema de comparación de dos muestras independientes de tamaños 9 y 14 en ausencia de normalidad, la suma de los rangos de los datos de la primera muestra toma el valor 69. El valor P del test está entre 0.01 y 0.05.
- II.19** Supongamos que, en el problema bilateral de comparación de dos medias a partir de dos muestras independientes de tamaños 17 y 15 de dos distribuciones normales

con varianzas desconocidas pero iguales, el valor experimental del test es igual a 3.16. Entonces el test es muy significativo.

- II.20** El test de D'Agostino pretende comprobar la hipótesis de independencia de los datos obtenidos de una muestra de la población.
- II.21** Se quiere contrastar $\sigma \leq 3$ contra $\sigma > 3$ a partir de una muestra de tamaño 10 de una distribución normal con parámetros μ y σ^2 desconocidos. Si la varianza muestral toma el valor 18, la decisión es aceptar la hipótesis nula.
- II.22** Para calcular el estadístico de contraste del test de comparación de dos proporciones en el caso de muestras independientes no necesitamos conocer exactamente las proporciones muestrales \hat{p}_1 y \hat{p}_2 de ambas muestras: basta conocer su diferencia $\hat{p}_1 - \hat{p}_2$.
- II.23** Con el objetivo de comparar dos medicamentos A y B para el tratamiento de una enfermedad, se piensa en seleccionar 100 enfermos al azar y basar la comparación en el porcentaje de enfermos que sanan con cada uno de ellos. El test de McNemar es un instrumento razonable para tomar una decisión de ese tipo.
- II.24** En el problema de comparación de dos medias para poblaciones normales en el caso de muestras relacionadas, la media muestral \bar{D} de las diferencias $X_i - Y_i$, $1 \leq i \leq n$, coincide con la diferencia $\bar{X} - \bar{Y}$ de las medias muestrales.
- II.25** En la determinación del intervalo de confianza para la diferencia de dos proporciones relacionadas sólo se tienen en cuenta los individuos con respuesta dispar.
- II.26** Admitamos que el peso (en Kg) de los individuos varones de 18 años de una cierta población sigue una distribución normal de media desconocida y varianza conocida 400. Si la media muestral de una muestra de tamaño 100 de varones de 18 años es igual a 75, el peso del 95 % de los varones de 18 años de la población estará entre 71.08 y 78.92 Kg.

- II.27** En el modelo de clasificación simple en análisis de la varianza, si la variación “dentro” aumenta y la variación “entre” permanece constante, el nivel mínimo de significación del test F disminuye.
- II.28** El modelo de clasificación simple en análisis de la varianza nos permite estudiar la relación entre una variable cuantitativa y otra cualitativa.
- II.29** (★) El test de Kruskal-Wallis está inspirado en la idea de que debe rechazarse la hipótesis de homogeneidad de las muestras comparadas si son pequeñas las diferencias observadas entre los rangos medios de las mismas.
- II.30** Si las varianzas de las diferentes muestras en un análisis de la varianza son muy diferentes, es cuestionable la decisión proporcionada por el test F .
- II.31** Si en el modelo de clasificación simple en análisis de la varianza la variación “entre” es nula, entonces las medias muestrales de las distintas muestras coinciden entre sí y con la media general de todos los datos obtenidos.
- II.32** Si en una tabla de contingencias $r \times s$ para la comparación de varias muestras en el caso discreto se supone cierta la hipótesis nula de que todas las muestras han sido extraídas de la misma población, la cantidad esperada de individuos en la casilla (i, j) es igual a $F_i C_j T$, siendo F_i el número de individuos en la fila i , C_j el número de individuos en la columna j y T el número total de individuos.
- II.33** En el modelo de clasificación simple en análisis de la varianza, la parte de la variación total explicada por el azar dividida por sus grados de libertad constituye un estimador de la varianza común de las variables que observamos.
- II.34** La suma de cuadrados $\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$ en el modelo de regresión lineal se interpreta como la parte de la variación total de Y explicada por la regresión lineal en X .
- II.35** La diferencia de riesgos se interpreta como la proporción esperada de individuos que no desarrollarán la enfermedad si el factor de riesgo es eliminado.

- II.36** El modelo de regresión lineal se puede utilizar para estimar la media de la variable dependiente Y a un valor dado de la variable independiente X aunque entre los datos disponibles no figure ningún valor de Y para ese valor de X .
- II.37** El modelo de regresión lineal presupone que, dado un valor x de la variable independiente X , la variable dependiente Y tiene distribución normal de media $\alpha + \beta x$ y varianza $\sigma^2 > 0$, donde α , β y σ^2 son parámetros desconocidos.
- II.38** Las pendientes de las rectas de regresión muestrales de Y sobre X y de X sobre Y coinciden.
- II.39** Si en una tabla de contingencias $r \times s$ para el estudio de la relación entre dos variables cualitativas todos los valores observados son iguales a 10, el test χ^2 rechaza la hipótesis de independencia entre esas dos variables.
- II.40** En un estudio sobre la relación entre dos variables cuantitativas X e Y medidas en un mismo individuo, si el coeficiente de determinación r^2 está próximo a 1, entonces Y tiende a crecer cuando X crece.
- II.41** La pendiente de la recta de regresión muestral de X sobre Y es igual a S_{xy}/S_{yy} .
- II.42** Supongamos conocido que el coeficiente de determinación muestral para una muestra de tamaño 22 de la distribución conjunta de dos variables cuantitativas es igual a 0.4. El valor P del test de independencia entre ambas variables verifica que $0.001 < P < 0.01$.
- II.43** Si tras contrastar la hipótesis de independencia lineal $\beta = 0$ entre X e Y en el modelo de regresión lineal, la decisión es rechazar la hipótesis nula entonces debemos aceptar que existen constantes a y b tales que $Y = a + bX$.
- II.44** Si $y = a + bx$ es la ecuación de la recta de regresión muestral entre la variable dependiente Y =“talla” y la variable independiente X =“edad” calculada a partir de una muestra de niños entre 3 y 9 meses de edad, entonces $2b$ es la estimación que ofrece el modelo de regresión lineal del aumento de talla de un niño que aumenta su edad en dos meses.

- II.45** Si el coeficiente de correlación de Pearson entre dos variables cuantitativas Y y X es igual a 0.5, entenderemos que el 50% de la variabilidad de Y queda explicada por la regresión lineal en X .
- II.46** Si en una tabla de contingencias $r \times s$ para estudiar la relación entre dos variables cualitativas X e Y el coeficiente de contingencia de Pearson es igual a 0.5, entenderemos que el 50% de la variabilidad de Y queda explicada por su asociación con X .
- II.47** En el estudio de la asociación entre dos variables dicotómicas se obtiene una tabla de contingencias 2×2 en la que la segunda fila es exactamente igual a la primera. Entonces se acepta la hipótesis de independencia de esas variables.
- II.48** Si el coeficiente de correlación de Pearson de un conjunto de pares de datos numéricos es muy alto entenderemos que existe una fuerte dependencia lineal entre las dos variables que generaron esos datos.
- II.49** En una tabla de contingencias 5×4 para contrastar la hipótesis de independencia de las dos variables cualitativas consideradas, el valor del estadístico de contraste es igual a 32.76. Podemos concluir, por tanto, que el test es altamente significativo.
- II.50** La distribución normal bivariante es la distribución de probabilidad de un par de variables normales X e Y que depende solamente de 4 parámetros: las medias y las varianzas de X e Y .
- II.51** Si en una tabla de contingencias las cantidades observadas coinciden con las esperadas, existe máxima asociación entre las variables cualitativas consideradas.
- II.52** Si la diferencia de riesgos para el estudio de la relación entre un factor de riesgo y una enfermedad es igual a 0.3, el porcentaje de individuos con el factor de riesgo que desarrollarán la enfermedad es un 3% superior al correspondiente a los individuos sin el factor de riesgo.

PROBLEMAS DE LA SEGUNDA PARTE

- II.1** En un estudio realizado en 1985 sobre mujeres entre 30 y 39 años de edad se obtuvo que el 20% de las mujeres que utilizaban anticonceptivos se habían decidido por un método anticonceptivo irreversible. Se sospechaba que eso podría haber cambiado en la década siguiente y en 1995 se eligieron al azar 100 mujeres entre 30 y 39 años que usan anticonceptivos y se observó que 25 de ellas utilizan un método irreversible. Con ese planteamiento, decide cuáles son las hipótesis nula H_0 y alternativa H_1 y determina el valor P del test correspondiente.
- II.2** Los niveles de testosterona en sangre de diez personas elegidas al azar de una población son 15.49, 15.20, 20.98, 10.35, 28.30, 10.66, 19.78, 13.57, 14.38 y 20.51.
- (a) Decide si es asumible la hipótesis de normalidad para la distribución de la que se han extraído esos datos.
- (b) Asumiendo normalidad, confirma si esos datos soportan la hipótesis de que el nivel medio de testosterona en la población en cuestión es igual a 21. Calcula y comenta el valor P .
- II.3** Se desea comparar las precisiones de dos métodos A y B de medida de la tiroxina libre, para lo cual se toma una muestra de tiroxina y se le mide 10 veces por el método A y 9 por el método B. Los datos se resumen en las medias muestrales $\bar{x} = 0.97$ e $\bar{y} = 1.20$ y en las desviaciones típicas muestrales $S_1 = 0.067$ y $S_2 = 0.118$. ¿Qué conclusión podemos obtener de esos datos?
- II.4** Con la intención de obtener los límites de normalidad o tolerancia para la presión sanguínea sistólica en una población se ha extraído de la misma una muestra de tamaño 25 de la misma y se han calculado la media y varianzas muestrales a partir de los datos obtenidos: $\bar{x} = 130$ y $S^2 = 729$. Calcular esos límites de tolerancia si pretendemos que entre ellos se encuentre el 90% de la población con una confianza del 95%.
- II.5** Una muestra de 13 personas ha sido dividida al azar en dos grupos de 6 y 7 personas,

respectivamente. Se ha administrado un fármaco A al primer grupo y un fármaco B al segundo. Después se ha extraído sangre a cada una de esas personas y se ha determinado el tiempo en minutos que tarda en coagular la muestra de sangre extraída. Esos tiempos son los siguientes: Fármaco A: 8.8, 8.4, 7.9, 8.7, 9.1, 9.6; fármaco B: 9.9, 9.0, 11.1, 9.6, 8.7, 10.4, 9.5.

(a) Asumiendo normalidad e igualdad de varianzas, ¿podemos concluir que el tiempo medio de coagulación para el fármaco A es superior al del fármaco B?

(b) ¿Es significativo el test de comparación de las varianzas de ambas muestras?

II.6 En un estudio llevado a cabo en varones mayores de 14 años en la ciudad de Badajoz con el objetivo de estimar el porcentaje de bebedores de riesgo (BR) y su distribución por grupos de edad y por centro de salud (CS), se ha extraído una muestra de 856 individuos y se ha aplicado a cada uno de ellos una encuesta a partir de la cual se toma la decisión de declarar o no bebedor de riesgo al individuo haciendo uso de un criterio elaborado por Altisent y otros. Los datos obtenidos se resumen en las tablas siguientes categorizados por grupos de edad y por centro de salud:

Edad	14 a 29	30 a 64	> 64	Total
BR	62	85	11	158
Total	268	440	148	856

CS	La Paz	San Fernando	San Roque	Anexo	Centro	Total
BR	47	50	15	36	10	158
Total	198	173	133	266	86	856

a) Estima la proporción de bebedores de riesgo en la población considerada y facilitar un intervalo de confianza al 95 % para ella.

b) Contrasta la hipótesis nula bilateral de que el porcentaje de bebedores de riesgo es del 20 %.

c) ¿Existen diferencias significativas entre las proporciones de bebedores de riesgo entre los grupos de edad 14–29 y ≥ 30 ? Facilitar un intervalo de confianza para la

diferencia de proporciones.

II.7 Uno de los modos de valorar un cemento dental consiste en comprobar si la pieza dental a la que fue aplicado ha sufrido algún deterioro a los 5 años de su aplicación. Con objeto de comparar dos cementos dentales A y B, fueron aplicados a 90 y 104 individuos, respectivamente, y al cabo de 5 años se encontró que 21 de los que recibieron el cemento A y 32 de los que recibieron el B sufrieron algún deterioro en las piezas cementadas. A la luz de estos datos, ¿podemos concluir que ambos cementos son igualmente efectivos?

II.8 Un programa de desintoxicación ética para adultos posee un 70 por ciento de éxito. En un grupo de 100 individuos que siguieron el programa, 79 dejaron definitivamente el alcohol.

¿Representa este dato evidencia suficiente de que el porcentaje de éxito ha aumentado?

II.9 Se ha llevado a cabo un estudio para estudiar el efecto del etanol en la vaso-dilatación retinal tras la administración de metanol (1 g/Kg) mediante un tubo nasogástrico. Se ha contado el número de venas dilatadas en la retina en las dos situaciones siguientes: la primera corresponde a la administración de metanol tal y como se ha indicado, y la segunda experiencia se repitió dos semanas después pero, en esta ocasión, tras la dosis de metanol se añadió media hora más tarde una dosis oral de etanol (1.4 g/Kg). Los datos obtenidos son los siguientes:

Metanol solo	10	9	7	10	7	4	8	4	10
Metano+Etanol	6	5	5	1	6	7	5	5	7

Contrasta la hipótesis de igualdad de medias.

II.10 Con el objetivo de decidir si el uso de medicación tiroidea puede influir en el desarrollo de un cáncer de mama, se ha realizado un estudio sobre 1851 mujeres, de las cuales 974 fueron elegidas al azar entre mujeres que han desarrollado un cáncer de mama y 877 entre mujeres que no han desarrollado cáncer de mama. Se ha observado que 90 entre las primeras y 66 entre las segundas habían tomado tal medicación. Analiza

esos datos y comenta los resultados, realizando un test apropiado y presentando un intervalo al 95 % de confianza.

II.11 La tabla siguiente contiene las temperaturas de 12 niñas de 4 años afectadas de gripe antes y después de haber ingerido aspirina.

Antes	39.1	39.6	38.8	39.4	38.4	38.2
Después	37.6	37.8	37.9	38.4	37.7	37.9
Antes	39.2	39.5	39.3	39.1	38.8	38.6
Después	38.3	37.8	38.2	38.4	38.5	37.9

a) ¿Podemos concluir que la aspirina es efectiva en la reducción de la temperatura de niñas de 4 años con gripe? Ind.: Antes de seleccionar un test para responder a esa cuestión, verifica si es asumible la hipótesis de normalidad y actúa en consecuencia.

b) Responde a la misma cuestión del apartado a) haciendo uso de un método estadístico alternativo al utilizado allí.

c) Asumiendo normalidad, estima la reducción media de temperatura tras la administración de aspirina y facilita un intervalo al 95 % de confianza para la diferencia de medias.

II.12 Con el objetivo de comparar las longitudes de los brazos derecho e izquierdo de una cierta población, se seleccionan al azar 10 individuos de la misma, observándose las siguientes diferencias en mm entre la longitud del brazo derecho y la del brazo izquierdo: 5.5, -3.2, 3.9, -1.5, 0.8, 4.7, 2.8, -1.9, 4.4, 2.5. Puede sostenerse, a la luz de estos datos y asumiendo normalidad, que la longitud del brazo derecho es mayor que la del brazo izquierdo. Calcula y comenta el valor P .

II.13 En un estudio realizado sobre recién nacidos se pretende decidir si el número de latidos de corazón por minuto depende de la raza del niño. Los datos obtenidos se resumen en la siguiente tabla:

Raza	Media muestral	Desviación típica muestral	n
Blanca	125	11	218
Negra	133	12	156

Decide si existen o no diferencias entre el número medio de latidos en ambas razas, comprobando previamente la hipótesis de igualdad de varianzas. Proporciona cotas para el valor P en ambos casos.

II.14 Se determinaron los niveles de testosterona en sangre a 10 personas supervivientes de infarto de miocardio y a 10 personas sanas, obteniéndose las medias y desviaciones típicas (en formato $\bar{x} \pm s$) siguientes:

Grupo infarto: 15.84 ± 4

Grupo control: 19.27 ± 7.8

¿A la luz de esos datos, puede afirmarse que los individuos supervivientes de un infarto de miocardio tienen un déficit en sangre de testosterona?

II.15 En un estudio realizado en EE.UU. sobre 300 mujeres entre 50 y 55 años cuyas madres tuvieron cáncer de mama, se observó que 12 habían padecido un cáncer de mama en algún momento de sus vidas. Supuesto conocido que la prevalencia de cáncer de mama en mujeres estadounidenses entre 50 y 55 años es del 2%, ¿podemos afirmar que el hecho de que la madre haya sufrido cáncer de mama aumenta la probabilidad de tener esa enfermedad? Realiza un test apropiado para responder a esa cuestión y construye un intervalo al 95% de confianza adecuado. Comenta todos los resultados obtenidos.

II.16 Para evaluar la influencia del tipo de acidosis del recién nacido en los niveles de glucemia medidos en el cordón umbilical del mismo, se obtuvieron los siguientes datos:

Controles	51	56	58	60	62	63	65	68	73	73
Acidosis respiratoria	60	65	66	68	68	69	73	75	78	80
Acidosis metabólica	69	73	74	78	79	79	82	85	87	88
Acidosis mixta	70	75	76	77	79	80	82	86	88	89

Asumiendo normalidad, analiza esos datos y decide si el tipo de acidosis influye realmente en el nivel de glucemia, comentando el nivel mínimo de significación P . Nota: la suma de los datos de las filas 1, 2, 3 y 4 son, respectivamente, 628, 702, 794, 802. La suma de todos los datos es igual a 2926.

- II.17** Se pretende contrastar una teoría sobre el cáncer de mama que afirma que el tiempo que transcurre entre la primera menstruación y el primer parto influye en la probabilidad de desarrollar un cáncer de mama. Para ello se lleva a cabo un estudio caso/control en el que interviene un total de 13465 mujeres que han tenido al menos un hijo; 3220 de ellas habían desarrollado un cáncer de mama (caso) y 10245 no lo habían desarrollado (control). A cada una de ellas se le preguntó por la edad a la que tuvieron su primer hijo. Los datos obtenidos se recogen en la tabla siguiente:

	< 20	20-24	25-29	30-34	≥ 35	Total
Caso	320	1206	1011	463	220	3220
Control	1422	4432	2893	1092	406	10245
Total	1742	5638	3904	1555	626	13465

¿Podemos concluir a partir de esos datos si hay relación entre la edad del primer parto y el desarrollo de cáncer de mama?

- II.18** Un cirujano sospecha que determinados tipos de dolores se alivian sólo con dar un medicamento, sea éste potente o no. Para contrastar esa hipótesis selecciona 15 pacientes a los que se acaba de extirpar las amígdalas y los asigna aleatoriamente a uno de 3 grupos; al primer grupo le administra un placebo, y al segundo y al tercero dos analgésicos A y B. Los tiempos (en horas) transcurridos hasta que los pacientes se quejaron de dolor son recogidos en la siguiente tabla:

Placebo	2.2	0.3	1.1	2.0	3.4
A	2.8	1.4	1.7	.3	4.3
B	1.1	4.2	3.8	2.6	0.5

¿Confirman esos datos las sospechas del cirujano?

II.19 Un grupo de 28 pacientes que han sufrido una hemorragia a causa de un aneurisma intracraneal han sido asignados aleatoriamente a uno de cuatro posibles tratamientos: reposo en cama (RC), hipotensión inducida por medicación (HIM), ligamento de carótida (LC) y cirugía intracraneal (CI). La tabla siguiente muestra el número de semanas hasta la muerte para esos pacientes.

RC	55	32	13	34	6	48	11
HIM	37	49	66	21	101	38	51
LC	20	16	58	34	85	67	50
CI	12	19	38	25	18	54	46

- Realiza un test paramétrico para evaluar las posibles diferencias en los tiempos de supervivencia entre los 4 grupos.
- En caso de que las diferencias sean significativas, compara esos grupos dos a dos.
- Facilita intervalos al 95 % de confianza para las diferencias de medias.
- Repite los apartados a) y b) utilizando un test no paramétrico.

II.20 Estamos interesados en la posible relación entre el infarto de miocardio y el consumo de café en varones de edad comprendida entre 60 y 64 años. Se eligen 200 individuos al azar de esa población y se obtienen los siguientes resultados:

Tazas de café/día	0	0	1	1	2	2	≥ 3	≥ 3
Infarto en los 5 últimos años	Si	No	Si	No	Si	No	Si	No
Núm. personas	3	57	7	43	8	42	12	28

Realiza un test para investigar si existe relación entre el número de tazas de café consumidas por día y el hecho de haber sufrido o no un infarto de miocardio en los últimos cinco años.

II.21 La tabla siguiente contiene determinaciones en útero de catalasa (CAT) y malondialdehído (MDA) de 14 ratas.

Rata	1	2	3	4	5	6	7
CAT	51.6	48.7	42.6	38	50.1	60.2	41.3
MDA	24.8	25	23.8	23.5	27.2	21.3	24.2
Rata	8	9	10	11	12	13	14
CAT	47.6	40.6	49.5	42	55.1	46.2	52.1
MDA	22.8	22.9	20.6	24	24.7	23.2	25.1

a) Determina la recta de regresión de CAT sobre MDA y decidir si existe dependencia lineal entre esas variables.

b) Calcula el coeficiente de correlación entre ambas y cuantificar el porcentaje de variabilidad de CAT explicada por la regresión lineal en la variable CAT.

II.22 En un hospital materno se han registrado los nacimientos prematuros y el consumo de tabaco de la madre durante un año, habiéndose obtenido los resultados siguientes: 10 nacimientos prematuros se detectaron en madres fumadoras y 80 nacimientos prematuros en las no fumadoras; entre los nacimientos normales, se observaron 30 casos en fumadoras y 270 en las no fumadoras.

a) ¿Existe alguna asociación entre las dos variables? Ind.: Respóndase utilizando como medida de asociación el riesgo relativo .

b) Facilita una estimación y un intervalo de confianza para esa medida de asociación.

II.23 Para un grupo de 27 individuos entre 19 y 65 años de edad con cardiomiopatía dilatada aguda se ha registrado su edad X y la fracción de eyección del ventrículo izquierdo Y . Los datos obtenidos se resumen como sigue:

$$\sum x_i = 1137, \quad \sum x_i^2 = 54749, \quad \sum y_i = 6.05$$

$$\sum y_i^2 = 1.522, \quad \sum x_i y_i = 262.93$$

a) Estima los coeficientes de la recta de regresión de Y sobre X , y proporciona intervalos al 95 % de confianza para ellos.

b) Contrasta la significación del modelo de regresión.

- c) Estima el valor medio de Y para un individuo de esas características con 45 años de edad.
- d) Ofrece una medida de asociación entre X e Y y cuantifica la relación lineal entre X e Y .
- e) Determina además una medida de asociación no paramétrica entre X e Y y contrasta su significación.

II.24 En un estudio sobre tumores cerebrales se desea averiguar si existe alguna relación (asociación) entre la localización del tumor y la naturaleza del mismo. Para ello se eligieron al azar 141 pacientes afectados de tumor cerebral y los datos se clasificaron en la tabla de contingencias 3×3 siguiente, en la que la localización (que llamaremos carácter A) aparece en el margen izquierdo y la naturaleza del tumor (que llamaremos carácter B) en la parte superior:

	Benigno	Maligno	Otros	Totales
Lóbulo frontal	$O_{11} = 23$	$O_{12} = 9$	$O_{13} = 6$	$F_1 = 38$
Lóbulo temporal	$O_{21} = 21$	$O_{22} = 4$	$O_{23} = 3$	$F_2 = 28$
Otras áreas	$O_{31} = 34$	$O_{32} = 24$	$O_{33} = 17$	$F_3 = 75$
Totales	$C_1 = 78$	$C_2 = 37$	$C_3 = 26$	$T = 141$

Decide si existe algún tipo de relación entre la localización del tumor y la naturaleza del mismo y cuantifica esa relación con alguna medida de asociación.

II.25 Se ha llevado a cabo un estudio en 3 países (1, 2 y 3) de la Unión Europea con la intención de comparar la distribución del color del cabello de sus individuos. Para ello, se han seleccionado en esos países muestras de tamaños respectivos 282, 315 y 87. Los datos obtenidos son los de la siguiente tabla de contingencias, y nos preguntamos si, a la luz de esos datos, podemos aceptar la hipótesis nula de que la distribución de los individuos por su color de pelo es la misma en esos 3 países o si, por el contrario, existen diferencias significativas en al menos dos de ellos.

	Rubio	Castaño	Moreno	Rojo	Totales
País 1	177	81	19	5	282
País 2	95	139	75	6	315
País 3	12	44	29	2	87
Totales	284	264	123	13	684

¿Qué conclusiones podemos extraer de esos datos?

- II.26** Se ha llevado a cabo un estudio con 300 individuos para ver si la capacidad para las matemáticas (se distinguen 3 niveles, baja, media y alta, y se seleccionan 100 individuos en cada una de esas categorías) y el interés por la estadística (3 niveles: bajo, medio, alto) están relacionados.

		Capacidad para matemáticas			
		Baja	Media	Alta	Totales
Interés por estadística	Bajo	50	25	15	90
	Medio	25	60	5	90
	Alto	25	15	80	120
Totales		100	100	100	300

Plantea el problema de test de hipótesis apropiado para, a la luz de esos datos, decidir si existe asociación entre ambas variables; facilita una medida de asociación apropiada.

Epílogo: ¿Qué hacer con los datos?

Primera etapa: entrada de datos

Los datos son los ingredientes básicos de cualquier problema de inferencia estadística y, para resaltar su importancia, hemos querido comenzar este libro con un conjunto real de datos. Queremos también finalizarlo hablando de ellos, pues un investigador en ciencias de la salud se encuentra, tras el trabajo de campo, con un conjunto de datos al que quiere extraer el mayor provecho posible.

Para el análisis de datos es conveniente utilizar alguno de los muchos programas estadísticos para ordenador existentes en el mercado, como pueden ser STATISTICA, SPSS, SAS, MINITAB, etc. Esos programas exigen que los datos sean presentados en una tabla. Cada columna recogerá los valores de una de las variables consideradas en el estudio. Cada fila recogerá los valores de esas variables para uno de los individuos (casos) que intervienen en el mismo. Ya hemos dicho (pág. 14) que las variables pueden ser cuantitativas (valores numéricos) o cualitativas (valores de naturaleza no numérica), y que las primeras pueden ser discretas o continuas. Podemos utilizar una variable cualitativa para clasificar a un individuo en una de varias clases posibles; por ejemplo, la variable sexo es una variable cualitativa que toma los valores Hombre y Mujer (que pueden ser representados por 0 y 1, respectivamente, aunque la naturaleza de estos datos no es numérica), o si queremos comparar los valores de una variable cuantitativa (estatura, por ejemplo) en distintas clases sociales, deberemos incluir una variable cualitativa en la que se indique la clase social a que pertenece el individuo (con valores Alta, Media o Baja, por ejemplo). Igualmente, si se ha determinado una variable X para individuos de varios grupos, conviene escribir todos los valores de X en una sola columna e introducir una variable cualitativa que nos indique a qué grupo pertenece cada uno de esos valores. Las variables cuantitativas discretas se utilizan usualmente para contar (por ejemplo, el número de piezas dentales con caries, número de copas de cerveza consumidas por día, etc), mientras que las variables cuantitativas continuas se suelen utilizar para medir (nivel de colesterol, estatura, presión sanguínea sistólica, etc). Una tabla de datos contendrá al menos una variable y , en general, varias variables de alguno de los

tres tipos mencionados.

Segunda etapa: estadística descriptiva

Una vez introducidos los datos, la primera fase en el análisis de los mismos corresponde a la estadística descriptiva, cuya misión queda descrita con dos verbos: resumir y visualizar (véanse los temas 0 y 1). Los datos son, habitualmente, muy numerosos y es conveniente sintetizarlos en ciertas cantidades que nos proporcionen información comprensible sobre los mismos.

Para cada variable cuantitativa conviene determinar estadísticos descriptivos que nos permitan localizar el conjunto de datos (media, mediana, valores mínimo y máximo y cuartiles inferior y superior son obligados) y hacernos una idea sobre su dispersión (desviación típica, rango intercuartílico y rango). También es conveniente estudiar la presencia de datos anómalos para detectar posibles errores de transcripción o experimentación que debieran ser subsanados; es cuestionable la eliminación de datos por el simple hecho de que parecen discrepar excesivamente del resto. Todo ello es tratado en el tema 1.

Para una variable cualitativa suele ser suficiente, en principio, determinar las frecuencias absoluta y relativa de cada clase.

Pero interesa también obtener estadísticos descriptivos de la relación entre variables. Así, deberíamos obtener alguna medida de correlación (coeficiente de correlación de Pearson o de Spearman) entre variables cuantitativas; los coeficientes de la recta de regresión y el coeficiente de determinación admiten una sencilla interpretación; consúltense los temas 16 y 17. Entre una variable cuantitativa y otra cualitativa podría resultar de interés determinar diferentes estadísticos descriptivos (media, mediana, varianza, cuartiles inferior y superior y valores mínimo y máximo) de la primera en cada una de las clases determinadas por la segunda. Entre dos variables cualitativas podemos construir una tabla de contingencias que recoge el número de datos para cada combinación posible de valores de las dos variables y facilitar alguna medida de asociación entre ellas (coeficiente de contingencia de Pearson o, en el caso de tablas 2×2 , el riesgo relativo y la razón del producto cruzado); véase el tema

17. El problema se complica -y, habitualmente, se sale de los límites de este libro- cuando, en lugar de dos, se pretende estudiar la relación entre 3 o más variables; podría resultar de interés, por ejemplo, estudiar la correlación entre dos variables cuantitativas en cada una de las clases en que el conjunto de individuos que forman parte del estudio ha quedado dividido de acuerdo con los valores de una cierta variable cualitativa; o estudiar la posible dependencia lineal entre tres o más variables cuantitativas.

Además de resumir (o sintetizar) los datos mediante el uso de apropiados estadísticos descriptivos, la estadística descriptiva ofrece una amplia gama de gráficos cuyo objetivo es proporcionar una rápida visualización de los mismos. Entre ellos queremos destacar los siguientes:

1) Para una primera impresión de la distribución de una variable cuantitativa, conviene presentar un diagrama de caja (pág. 29) que recoge las posiciones de la mediana, los cuartiles inferior y superior y los valores mínimo y máximo; estos valores dividen el conjunto de datos ordenados de menor a mayor en cuatro partes, cada una con un 25 % de los datos, y, mediante un simple vistazo, conseguimos información sobre la localización y dispersión de cada una de esas cuatro partes, la simetría del conjunto de datos, y, si se incluyen en el gráfico (los programas estadísticos suelen hacerlo), la posible existencia de datos anómalos (outliers, en inglés) que conviene revisar cuidadosamente uno a uno (ver pág. 31).

2) Con el mismo objetivo del punto anterior, una información más detallada (y, por tanto, menos fácil de interpretar) sobre la distribución de una variable aleatoria cuantitativa se puede obtener mediante un histograma (pág. 28) con un adecuado número de clases. Además de los objetivos del punto anterior, podemos ahora obtener una primera impresión sobre la posible normalidad de los datos; muchos programas estadísticos superponen al histograma la función de densidad de la distribución normal que mejor aproxima a esos datos (i.e., la distribución normal que tiene la misma media y varianza que el conjunto de datos). La suposición de normalidad está en la base de muchos de los métodos estadísticos que podrían utilizarse en una etapa posterior del análisis de los datos; nos referimos a la etapa de inferencia estadística. Un histograma excesivamente sesgado (a derecha o izquierda) sugeriría rechazar esa hipótesis de normalidad; un histograma con dos “jorobas” podría ser

un indicio de la presencia de dos grupos claramente distintos en el conjunto de datos que no han sido detectados a priori (como cuando mezclamos las alturas de un grupo de chicos de seis años con otro grupo de chicos de 20 años, en cuyo caso sería inapropiado en la etapa de inferencia suponer que esos son datos de una población normal).

3) Para obtener una rápida impresión de la distribución de una variable cualitativa (o, incluso, de una variable cuantitativa discreta con pocos valores posibles), la mejor opción puede ser el diagrama de barras (ver pág. 28). Los valores de una variable cualitativa en un grupo de individuos suelen llamarse datos categóricos.

4) Para visualizar la relación entre dos variables cuantitativas, nada mejor que dibujar en el plano la nube de puntos (véase el tema 15) situando una variable en el eje de abscisas y la otra en el eje de ordenadas. Más completo queda el gráfico si incluimos la recta de regresión, que nos marca la posible dependencia lineal entre ambas; cuando se utilizan escalas equivalentes en ambos ejes, una recta de regresión que forme un ángulo próximo a 45 o a 135 grados con la parte positiva del eje X será indicio de una fuerte relación lineal entre ambas variables, mientras que rectas de regresión casi horizontales vendrían a indicar una dependencia lineal prácticamente nula, aunque en esta interpretación hay que tener también en cuenta la distribución de la nube de puntos en el plano. Una imagen de la distribución conjunta de ambas variables puede obtenerse mediante un histograma tridimensional que divide en clases los valores de las dos variables y sitúa dichas clases en los ejes X e Y de un plano horizontal, y levanta sobre cada cuadrado de ese plano determinado por una combinación de clases un prisma de volumen proporcional a la frecuencia absoluta de esa combinación; hay programas estadísticos que incluyen la posibilidad de obtener diferentes perspectivas de este histograma tridimensional para facilitar su visualización.

5) Si una de las variables es cuantitativa y la otra cualitativa, una forma sencilla de visualizar la relación entre ambas se obtiene mediante un gráfico que contenga un diagrama de caja de la variable cuantitativa para cada uno de los valores posibles de la variable cualitativa; pueden situarse estos últimos en el eje de abscisas y dibujar sobre cada valor el diagrama de caja de los datos de la variable cuantitativa correspondientes al mismo; de ese modo podemos valorar, por ejemplo, cómo varía la mediana de la variable cuantitativa al

cambiar el valor de la cualitativa (por ejemplo, interesados en la estatura de un grupo de personas, convendría obtener dos diagramas de caja para las estaturas, uno para varones y otro para mujeres). También podemos presentar un gráfico con múltiples histogramas (uno por cada valor de la variable cualitativa) para visualizar, por ejemplo, la posible normalidad de la variable cuantitativa en cada una de las clases definidas por la cualitativa.

6) Para el caso de dos variables cualitativas no hemos descrito ningún gráfico concreto. La tabla de contingencias puede considerarse también como un pseudográfico, que suele ser suficiente si no es excesivamente grande. Existen métodos gráficos (por ejemplo, un diagrama de barras tridimensional que levantaría sobre cada combinación posible de valores de las dos variables un prisma de altura proporcional a la frecuencia absoluta de esa combinación, o un gráfico superpuesto a la tabla de contingencias en la que las frecuencias absolutas que aparecen son reemplazadas por rectángulos de áreas proporcionales a ellas), pero son poco utilizados.

7) Un mayor número de variables entraña desde luego una mayor complejidad gráfica. Consideraremos simplemente dos casos: la relación lineal entre tres variables cuantitativas puede visualizarse dibujando en el espacio la nube de puntos (se asigna una variable a cada eje coordenado) y ajustando a ella el plano de regresión (disponible habitualmente en los programas estadísticos en el módulo de regresión múltiple); si tenemos dos variables cuantitativas y una cualitativa, podemos dibujar la nube de puntos de las dos primeras variables para cada uno de los valores de la variable cualitativa, y ajustar una recta de regresión a cada una de las nubes de puntos obtenidas, para comprobar si la relación lineal entre las dos variables cuantitativas queda afectada por el valor de la cualitativa.

Tercera etapa: inferencia estadística

Observación 6.29. (Aclaraciones previas sobre diseño de experimentos) Supongamos que los datos de que disponemos se refieren a 100 individuos para los que hemos determinado los valores de una o más variables; los métodos descriptivos que acabamos de mencionar nos ayudan a sintetizar ese conjunto de datos y a obtener una primera impresión gráfica de los mismos. Pero surge una cuestión inevitable: ¿por qué tomarse tantas molestias en escudriñar datos obtenidos de un conjunto de 100 individuos? A menos que esos individuos tengan algo que ver personalmente con el autor del estudio, es muy probable que la justificación última

de su afán se encuentre en la posibilidad de que las conclusiones extraídas de esos 100 individuos puedan extrapolarse a una población más grande, y, para ello, necesitamos ir a una nueva etapa: la de inferencia estadística. Los datos pasan a ser considerados realizaciones concretas de variables definidas en la población objetivo del estudio y tendremos que hacer ciertas suposiciones (independencia, normalidad,...) sobre la distribución teórica de las variables observadas para construir un modelo estadístico en el que nos apoyaremos para realizar esa extrapolación, y, en la fase de diseño del experimento, previa a la selección de la muestra y obtención de los datos, se deberá hacer un análisis riguroso de las condiciones de experimentación que justifique el cumplimiento de esas suposiciones.

Para que la extrapolación sea posible, los individuos que forman parte del estudio deben constituir una muestra representativa de la población, y la representatividad se conseguirá con un procedimiento que garantice que todos los individuos de la población objetivo tengan la misma probabilidad de resultar elegidos, es decir, mediante un muestreo aleatorio simple (m.a.s.); la puesta en práctica de un m.a.s. puede entrañar serias dificultades, y se procurará garantizar, al menos, que ciertas características que puedan parecer relevantes en el estudio se distribuyan en la muestra de forma similar a como lo hacen en la población objetivo. Esta fase de diseño del experimento, previa a la obtención de los datos, resulta fundamental a la hora de garantizar la representatividad de la muestra.

En esta fase de diseño del experimento se suele abordar también la cuestión del tamaño de la muestra. Conviene dejar claro que la estadística nos proporciona instrumentos útiles para tomar decisiones sobre grandes poblaciones a partir de un pequeño conjunto de datos y ciertas suposiciones sobre la distribución de las variables consideradas en la población. Por tanto, sea cual sea el tamaño de muestra disponible, el método estadístico que utilicemos nos va a ofrecer una decisión (un mínimo número de individuos es requerido por muchos métodos estadísticos para que sean posibles los cálculos). Pero es igualmente claro que: 1) los métodos estadísticos se van haciendo más y más precisos (y más costosos) a medida que aumenta el tamaño de muestra; 2) suposiciones cuestionables para tamaños de muestra pequeños pueden dejar de serlo para muestras grandes; 3) los problemas de determinación del tamaño muestral se plantean en términos análogos a los que se hicieron en la observación 2 de la página 156, es decir, se trata de determinar el tamaño de muestra necesario para conseguir una precisión dada, pero si nos conformamos con la precisión que obtengamos, podemos obviar el problema de determinación del tamaño muestral planteado en estos términos. <

1) En la etapa de inferencia estamos interesados en utilizar los datos para estimar (tema 9), puntualmente o por intervalos, o contrastar hipótesis (tema 10) sobre uno o más parámetros -desconocidos, desde el punto de vista estadístico- de las variables consideradas en la población objetivo: estos parámetros pueden ser medias, proporciones, varianzas, coeficientes de correlación, etc. La estimación puntual de un parámetro poblacional nos proporcionará a partir de los datos un valor estimado del mismo, mientras que la estimación por intervalos de ese parámetro nos proporcionará un intervalo de confianza alrededor del mismo que nos

permitirá hacernos una idea del error cometido en la estimación puntual. Los problemas de contraste de hipótesis sobre uno o más parámetros poblacionales tienen como objetivo hacer uso de los datos para tomar una de dos decisiones posibles: aceptar como verdadera o rechazar una afirmación (hipótesis nula, denotada por H_0) que se ha formulado sobre los mismos.

2) Además de los datos, para poder abordar un problema de inferencia estadística, tenemos que fijar un modelo teórico -una estructura estadística- para las variables consideradas en el estudio, es decir, tendremos que hacer algunas suposiciones sobre las distribuciones de esas variables:

- 2.1) (Independencia) Una suposición habitual en muchos métodos estadísticos es, por ejemplo, la de considerar los datos x_1, \dots, x_n de una variable X como valores concretos tomados por variables independientes X_1, \dots, X_n cada una de ellas con la misma distribución que X ; la independencia es automática cuando se determina una misma variable en individuos que han sido seleccionados al azar de una población, pero no es razonable suponer independientes observaciones en un mismo individuo de una variable o de diferentes variables (como pueden ser las presiones sanguíneas sistólicas tomadas mensualmente a una persona, o la estatura y el peso de un mismo individuo).
- 2.2) (Normalidad) Otra suposición habitual para variables cuantitativas es la de que tiene distribución normal. Muchos de los métodos estadísticos más utilizados en la práctica están basados en la suposición de normalidad. La distribución normal es la más importante de la estadística pues, además de que estudios epidemiológicos a gran escala han permitido confirmar como razonable el uso de la distribución normal para muchas variables biológicas (como puede ser el nivel de colesterol) en grandes poblaciones, los métodos estadísticos basados en la suposición de normalidad pueden ser aproximadamente válidos en ausencia de esta si los tamaños muestrales son grandes. Es precisamente esta cualidad de la distribución normal de aproximar en muestras grandes a otras muchas distribuciones (discretas, incluso) lo que la hace especialmente importante. Pero, además de esto, podemos hacer uso de los datos para decidir si es

razonable o no asumir normalidad: un histograma excesivamente sesgado a izquierda o derecha, un coeficiente de asimetría muy distinto de cero o un coeficiente de curtosis muy diferente de 0 serían indicios importantes de ausencia de normalidad. Los programas estadísticos suelen incluir otros métodos descriptivos (como el gráfico de probabilidad normal, que representa en un plano los cuantiles muestrales frente a los cuantiles correspondientes a la distribución normal estándar, siendo cuestionable la normalidad si los puntos así obtenidos están lejos de estar alineados) y diversos tests de normalidad, como son los de Kolmogorov-Smirnov, Lilliefors, Shapiro-Wilks o D'Agostino, que permiten contrastar si los datos obtenidos son compatibles con la suposición de normalidad.

- 2.3) (Métodos no paramétricos) En ausencia de normalidad y en el caso de pequeñas muestras puede ser preferible utilizar métodos estadísticos no paramétricos; son métodos que no asumen ninguna distribución especial para las variables observadas y que compiten con éxito con los métodos tradicionales basados en la suposición de normalidad.

3) Por tanto, puesto que para un mismo problema tenemos a nuestra disposición diversos métodos estadísticos, antes de decidirnos por uno de ellos conviene hacer un estudio detallado de las condiciones de experimentación y de los datos obtenidos para garantizar que se cumplen las suposiciones en las que se basa el método que finalmente vayamos a usar.

4) Supongamos que tenemos una variable cualitativa dicotómica, es decir, que solo toma dos valores que denotaremos por 1 y 0 (que pueden representar varón y hembra, o fracaso y éxito en una prueba, o que ha ocurrido un cierto suceso y que no ha ocurrido, respectivamente). En este caso podemos estar interesados en estimar la probabilidad de éxito (o proporción de éxitos) en futuras realizaciones del experimento, y en proporcionar un intervalo de confianza alrededor de esa estimación. O, sabiendo que en otros estudios similares se ha obtenido una proporción p_0 de éxitos, podríamos estar interesados en contrastar la hipótesis $p = p_0$ con nuestros datos (ver tema 11).

5) Para una variable cualitativa que tome más de dos valores también podemos estimar la probabilidad de que la variable en la población objetivo tome uno de los valores de la

variable, además de dar un intervalo de confianza para esa probabilidad. También podría interesarnos contrastar la hipótesis de que las proporciones de los n valores de la variable cualitativa son, respectivamente, p_1, \dots, p_n (pág. 170).

6) Para una variable cuantitativa normalmente distribuida en la población objetivo, pero con media y varianza desconocidas, podríamos estar interesados en estimar puntualmente y por intervalos la media y/o la varianza, o, sabiendo que en otra población se ha estimado que la media de la variable que nos interesa es un valor μ_0 , podríamos estar interesados en contrastar la hipótesis nula bilateral $H_0 : \mu = \mu_0$ contra $H_1 : \mu \neq \mu_0$ (ver tema 11). Hay casos en los que en vez de realizar un test bilateral conviene realizar un test unilateral (ver pág. 154).

7) Si más que hacer inferencia sobre una sola variable estamos interesados en estudiar conjuntamente dos o más variables, hay dos problemas básicos que pueden interesarnos y que tienen que ver con los términos “comparación” y “relación”.

8) Supongamos que tenemos una variable X dicotómica (que toma los valores 0 y 1, que pueden corresponder a fracaso y éxito, respectivamente, en un experimento aleatorio) cuyos valores en dos grupos (o poblaciones) A y B queremos comparar. Introduciremos en la tabla de datos una nueva variable cualitativa Y que asigna a cada individuo el grupo (A o B) al que pertenece, y cualquier programa estadístico nos permitirá comparar las proporciones de éxito en ambos grupos tal y como queda explicado en la página 186. Lo mismo podemos hacer si queremos comparar las proporciones de éxito en más de dos grupos (ver pág. 223). El test de McNemar (pág. 188) nos permite comparar las proporciones de éxito para dos variables dicotómicas determinadas en un mismo individuo.

9) Si es una variable cuantitativa X la que queremos comparar en dos grupos (o poblaciones) A y B, introducimos una variable dicotómica Y en la tabla de datos como en el punto anterior; interesa habitualmente comparar las medias de la variable X en los dos grupos, pero, si vamos a suponer normalidad, convendría previamente comparar las varianzas para ver si aplicamos el test t de Student (en el caso de varianzas iguales) o el de Welch (en el caso de varianzas desiguales); si no suponemos normalidad, podemos utilizar el test no paramétrico de Mann-Whitney-Wilcoxon (ver tema 12, comparación de dos muestras

independientes).

10) Si una misma variable cuantitativa X se mide dos veces (antes y después de aplicar un tratamiento, por ejemplo) en cada individuo de una población, podemos estar interesados en comparar los valores medios de X “antes” con los valores “después”. Para ello, bajo la suposición de normalidad, podemos utilizar el test t de Student para muestras relacionadas o, sin esa suposición, el test de Wilcoxon para muestras relacionadas (ver tema 12, comparación de dos muestras relacionadas).

11) Si tenemos una variable cuantitativa X cuyos valores en tres o más grupos queremos comparar, introduciremos en la tabla de datos una variable cualitativa Y que asigna a cada individuo de la muestra su grupo, y, bajo la suposición de normalidad e igualdad de varianza de X en todos los grupos, el tema 13 nos enseña a realizar un análisis de la varianza (ANOVA) seguido de un estudio de comparaciones múltiples en el caso de que el test resulte significativo (y tengamos que declarar no homogéneos esos grupos) para decidir qué subgrupos se pueden considerar homogéneos. Podemos usar el test de Kruskal-Wallis (tema 13) si no se verifican las condiciones para hacer un análisis de la varianza. En realidad, estos son métodos que permiten estudiar la relación entre la variable cuantitativa X y la cualitativa Y .

12) La relación entre dos variables cualitativas (que toman r y s valores, respectivamente) se puede estudiar mediante el test χ^2 en una tabla de contingencias $r \times s$ (tema 17) que recoge el número de individuos de la muestra que hay en cada combinación de valores de las dos variables. En este caso, se puede facilitar una medida del grado de asociación (o de dependencia) que hay entre ambas variables (coeficiente de contingencia de Pearson). El caso $r \times 2$ corresponde al problema de comparación de r proporciones (la variable que toma dos valores es dicotómica y realmente estaríamos comparando las proporciones de éxito en cada una de las clases en que la población queda dividida de acuerdo con la otra variable). El caso 2×2 es especialmente interesante en ciencias de la salud (por ejemplo, para estudiar la dependencia entre el desarrollo de una enfermedad y la presencia de un cierto factor de riesgo); unas medidas de asociación interesantes en este caso son el riesgo relativo y la razón del producto cruzado, sobre las que también podría interesarnos hacer algún tipo de

inferencia (contrastar, por ejemplo, si son iguales a 1, lo que supondría que las dos variables son independientes).

13) La relación entre dos variables cuantitativas X (llamada variable independiente) e Y (llamada variable dependiente) puede hacerse a través del modelo de regresión lineal (tema 15) estimando puntualmente y por intervalos la pendiente de la recta de regresión, la ordenada en el origen y la varianza de regresión, y cuantificando el grado de relación lineal entre ambas mediante el coeficiente de correlación de Pearson r (sobre el que también podemos plantear problemas de estimación y contraste de hipótesis); el coeficiente de determinación r^2 admite una sencilla interpretación como fracción de variabilidad de la variable Y explicada por su dependencia lineal de X (tema 16). El modelo de regresión lineal se puede aprovechar también para estimar el valor medio de la variable Y para un valor concreto de la variable X , y para predecir el valor de la variable Y en un nuevo individuo si conocemos su valor de X .

14) En todos los contrastes de hipótesis conviene facilitar su valor P (pág. 153), también llamado probabilidad de significación o nivel mínimo de significación. El test se dice significativo si $P \leq 0.05$ y no significativo en otro caso. Por ejemplo, para comparar dos medias poblacionales (desconocidas desde el punto de vista estadístico pues no tenemos acceso a toda la población), podemos fijarnos en las medias muestrales (que son estimadores de aquellas); el hecho de que estas sean distintas no significa que aquellas lo sean, pues las diferencias observadas entre las medias muestrales podrían ser debidas al azar (siempre presente en experimentos aleatorios); para evaluar en qué medida puedan ser debidas al azar, el estadístico de contraste compara esa diferencia con una medida de la variabilidad de los datos, y solo cuando la discrepancia entre las medias muestrales es grande en comparación con la variabilidad de los datos decimos que las diferencias entre las medias poblacionales son (estadísticamente) significativas. El término “significativo” tiene otra connotación que queremos mencionar: si el test es significativo, la decisión es rechazar la hipótesis nula y esta decisión, aunque pudiera ser errónea, es fiable pues nos hemos asegurado de que la probabilidad de cometer un error sea pequeña (5%); pero si el test es no significativo, aceptaremos la hipótesis nula con cautela, pues la probabilidad de haber cometido un error al tomar esa

decisión no ha sido previamente controlada (aunque se hace más y más pequeña a medida que aumenta el tamaño de la muestra).

Observación 6.30. (¿Y qué más?) Lo dicho en las secciones precedentes resume los métodos estadísticos estudiados en este manual que, como su título indica, solo incluye técnicas estadísticas elementales. La misma línea argumental desarrollada sugiere inmediatamente nuevas situaciones (comparación y relación entre tres o más variables de diferentes tipos) para las que la estadística nos propone nuevos métodos (estadísticos descriptivos, gráficos y soluciones a problemas de estimación y contraste de hipótesis). Muchos de ellos aparecen implementados en los programas estadísticos habituales bajo el epígrafe genérico de métodos multivariantes, como puedan ser la regresión múltiple multivariante o el análisis multivariante de la varianza. Otras técnicas multivariantes han sido diseñadas para estudiar y simplificar la estructura de correlación en uno o dos grupos de variables (análisis de componentes principales, análisis factorial y análisis de correlaciones canónicas), o para clasificar individuos de acuerdo con los valores de varias variables (análisis discriminante y análisis cluster). El modelo log-lineal ha sido diseñado para el análisis de tablas de contingencias más sofisticadas que las que hemos considerado en este manual. El análisis de series temporales tiene que ver con el análisis de datos que se van generando periódicamente a lo largo del tiempo y que, por tanto, no pueden suponerse realizaciones concretas de variables independientes. ◁

Bibliografía

- ALTMAN (1991), *Practical Statistics for Medical Research*, Chapman and Hall.
- ARMITAGE, BERRY (1992), *Estadística para la Investigación Biomédica*, Doyma.
- DANIEL (1983), *Biostatistics: a Foundation for Analysis in the Health Sciences*, Wiley.
- FISHER, VAN BELLE (1993), *Biostatistics, a Methodology for the Health Sciences*, Wiley.
- MARTÍN ANDRÉS, LUNA DEL CASTILLO (2004), *Bioestadística para las Ciencias de la Salud*, Norma.
- REMINGTON, SCHORK (1974), *Estadística Biométrica y Sanitaria*, Prentice-Hall International.
- ROSNER (2006), *Fundamentals of Biostatistics*, PWS Kent.
- SNEDECOR, COCHRAN (1973), *Statistical Methods*, The Iowa State University Press.
- WOOLSON (2002) *Statistical Methods for the Analysis of Biomedical Data*, Wiley.
- ZAR (1999), *Biostatistical Analysis*, Prentice-Hall.

Apéndice: Tablas estadísticas

- Tabla I: 3000 dígitos aleatorios
- Tabla II: Distribución binomial $B(n, p)$
- Tabla III.1: Distribución $N(0, 1)$. Cuantiles
- Tabla III.2: Distribución $N(0, 1)$. Función de distribución
- Tabla IV: Distribución t de Student
- Tabla V: Distribución chi-cuadrado $\chi^2(n)$
- Tabla VI: Tablas de la distribución $F(m, n)$ de Fisher
- Tabla VII: Tabla para el test de normalidad de D'Agostino
- Tabla VIII: Región de aceptación para el test de Mann-Whitney-Wilcoxon: muestras independientes (★)
- Tabla IX: Distribución T de Wilcoxon para muestras relacionadas (★)
- Tabla X: Distribución q de Tukey para comparaciones múltiples
- Tabla XI: Distribución Q para comparaciones múltiples no paramétricas (★)
- Tabla XII: Distribución de Bonferroni
- Tabla XIII: Valores críticos para el coeficiente de correlación de Spearman (★)
- Tabla XIV: Factores K para límites de tolerancia bilaterales para distribuciones normales

Tabla II: Distribución binomial $B(n, p)$ (p. 73)

n	k	p									
		0.050	0.100	0.150	0.200	0.250	0.300	0.350	0.400	0.450	0.500
2	0	0.903	0.810	0.723	0.640	0.563	0.490	0.423	0.360	0.303	0.250
2	1	0.095	0.180	0.255	0.320	0.375	0.420	0.455	0.480	0.495	0.500
2	2	0.003	0.010	0.023	0.040	0.063	0.090	0.123	0.160	0.203	0.250
3	0	0.857	0.729	0.614	0.512	0.422	0.343	0.275	0.216	0.166	0.125
3	1	0.135	0.243	0.325	0.384	0.422	0.441	0.444	0.432	0.408	0.375
3	2	0.007	0.027	0.057	0.096	0.141	0.189	0.239	0.288	0.334	0.375
3	3	0.000	0.001	0.003	0.008	0.016	0.027	0.043	0.064	0.091	0.125
4	0	0.815	0.656	0.522	0.410	0.316	0.240	0.179	0.130	0.092	0.062
4	1	0.171	0.292	0.368	0.410	0.422	0.412	0.384	0.346	0.299	0.250
4	2	0.014	0.049	0.098	0.154	0.211	0.265	0.311	0.346	0.368	0.375
4	3	0.000	0.004	0.011	0.026	0.047	0.076	0.111	0.154	0.200	0.250
4	4	0.000	0.000	0.001	0.002	0.004	0.008	0.015	0.026	0.041	0.063
5	0	0.774	0.590	0.444	0.328	0.237	0.168	0.116	0.078	0.050	0.031
5	1	0.204	0.328	0.392	0.410	0.396	0.360	0.312	0.259	0.206	0.156
5	2	0.021	0.073	0.138	0.205	0.264	0.309	0.336	0.346	0.337	0.313
5	3	0.001	0.008	0.024	0.051	0.088	0.132	0.181	0.230	0.276	0.313
5	4	0.000	0.000	0.002	0.006	0.015	0.028	0.049	0.077	0.113	0.156
5	5	0.000	0.000	0.000	0.000	0.001	0.002	0.005	0.010	0.018	0.031
6	0	0.735	0.531	0.377	0.262	0.178	0.118	0.075	0.047	0.028	0.016
6	1	0.232	0.354	0.399	0.393	0.356	0.303	0.244	0.187	0.136	0.094
6	2	0.031	0.098	0.176	0.246	0.297	0.324	0.328	0.311	0.278	0.234
6	3	0.002	0.015	0.041	0.082	0.132	0.185	0.235	0.276	0.303	0.313
6	4	0.000	0.001	0.005	0.015	0.033	0.060	0.095	0.138	0.186	0.234
6	5	0.000	0.000	0.000	0.002	0.004	0.010	0.020	0.037	0.061	0.094
6	6	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.004	0.008	0.016
7	0	0.698	0.478	0.321	0.210	0.133	0.082	0.049	0.028	0.015	0.008
7	1	0.257	0.372	0.396	0.367	0.311	0.247	0.185	0.131	0.087	0.055
7	2	0.041	0.124	0.210	0.275	0.311	0.318	0.298	0.261	0.214	0.164
7	3	0.004	0.023	0.062	0.115	0.173	0.227	0.268	0.290	0.292	0.273
7	4	0.000	0.003	0.011	0.029	0.058	0.097	0.144	0.194	0.239	0.273
7	5	0.000	0.000	0.001	0.004	0.012	0.025	0.047	0.077	0.117	0.164
7	6	0.000	0.000	0.000	0.000	0.001	0.004	0.008	0.017	0.032	0.055
7	7	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.004	0.008

Tabla II (continuación): Distribución binomial $B(n, p)$

n	k	p									
		0.050	0.100	0.150	0.200	0.250	0.300	0.350	0.400	0.450	0.500
8	0	0.663	0.430	0.272	0.168	0.100	0.058	0.032	0.017	0.008	0.004
8	1	0.279	0.383	0.385	0.336	0.267	0.198	0.137	0.090	0.055	0.031
8	2	0.051	0.149	0.238	0.294	0.311	0.296	0.259	0.209	0.157	0.109
8	3	0.005	0.033	0.084	0.147	0.208	0.254	0.279	0.279	0.257	0.219
8	4	0.000	0.005	0.018	0.046	0.087	0.136	0.188	0.232	0.263	0.273
8	5	0.000	0.000	0.003	0.009	0.023	0.047	0.081	0.124	0.172	0.219
8	6	0.000	0.000	0.000	0.001	0.004	0.010	0.022	0.041	0.070	0.109
8	7	0.000	0.000	0.000	0.000	0.000	0.001	0.003	0.008	0.016	0.031
8	8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.004
9	0	0.630	0.387	0.232	0.134	0.075	0.040	0.021	0.010	0.005	0.002
9	1	0.299	0.387	0.368	0.302	0.225	0.156	0.100	0.060	0.034	0.018
9	2	0.063	0.172	0.260	0.302	0.300	0.267	0.216	0.161	0.111	0.070
9	3	0.008	0.045	0.107	0.176	0.234	0.267	0.272	0.251	0.212	0.164
9	4	0.001	0.007	0.028	0.066	0.117	0.172	0.219	0.251	0.260	0.246
9	5	0.000	0.001	0.005	0.017	0.039	0.074	0.118	0.167	0.213	0.246
9	6	0.000	0.000	0.001	0.003	0.009	0.021	0.042	0.074	0.116	0.164
9	7	0.000	0.000	0.000	0.000	0.001	0.004	0.010	0.021	0.041	0.070
9	8	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.004	0.008	0.018
9	9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.002
10	0	0.599	0.349	0.197	0.107	0.056	0.028	0.013	0.006	0.003	0.001
10	1	0.315	0.387	0.347	0.268	0.188	0.121	0.072	0.040	0.021	0.010
10	2	0.075	0.194	0.276	0.302	0.282	0.233	0.176	0.121	0.076	0.044
10	3	0.010	0.057	0.130	0.201	0.250	0.267	0.252	0.215	0.166	0.117
10	4	0.001	0.011	0.040	0.088	0.146	0.200	0.238	0.251	0.238	0.205
10	5	0.000	0.001	0.008	0.026	0.058	0.103	0.154	0.201	0.234	0.246
10	6	0.000	0.000	0.001	0.006	0.016	0.037	0.069	0.111	0.160	0.205
10	7	0.000	0.000	0.000	0.001	0.003	0.009	0.021	0.042	0.075	0.117
10	8	0.000	0.000	0.000	0.000	0.000	0.001	0.004	0.011	0.023	0.044
10	9	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.004	0.010
10	10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
11	0	0.569	0.314	0.167	0.086	0.042	0.020	0.009	0.004	0.001	0.000
11	1	0.329	0.384	0.325	0.236	0.155	0.093	0.052	0.027	0.013	0.005
11	2	0.087	0.213	0.287	0.295	0.258	0.200	0.140	0.089	0.051	0.027

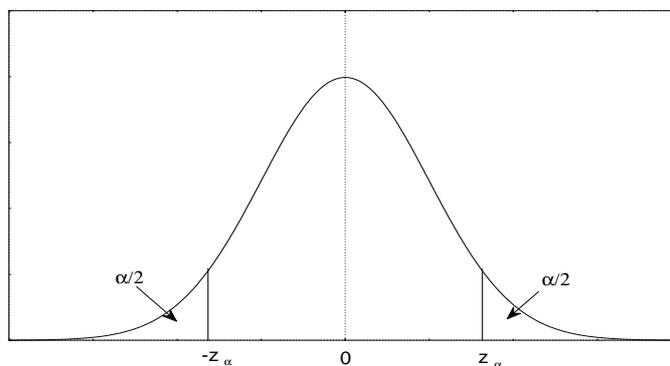
Tabla II (continuación): Distribución binomial $B(n, p)$

n	k	p									
		0.050	0.100	0.150	0.200	0.250	0.300	0.350	0.400	0.450	0.500
11	3	0.014	0.071	0.152	0.221	0.258	0.257	0.225	0.177	0.126	0.081
11	4	0.001	0.016	0.054	0.111	0.172	0.220	0.243	0.236	0.206	0.161
11	5	0.000	0.002	0.013	0.039	0.080	0.132	0.183	0.221	0.236	0.226
11	6	0.000	0.000	0.002	0.010	0.027	0.057	0.099	0.147	0.193	0.226
11	7	0.000	0.000	0.000	0.002	0.006	0.017	0.038	0.070	0.113	0.161
11	8	0.000	0.000	0.000	0.000	0.001	0.004	0.010	0.023	0.046	0.081
11	9	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.005	0.013	0.027
11	10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.005
11	11	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
12	0	0.540	0.282	0.142	0.069	0.032	0.014	0.006	0.002	0.001	0.000
12	1	0.341	0.377	0.301	0.206	0.127	0.071	0.037	0.017	0.008	0.003
12	2	0.099	0.230	0.292	0.283	0.232	0.168	0.109	0.064	0.034	0.016
12	3	0.017	0.085	0.172	0.236	0.258	0.240	0.195	0.142	0.092	0.054
12	4	0.002	0.021	0.068	0.133	0.194	0.231	0.237	0.213	0.170	0.121
12	5	0.000	0.004	0.019	0.053	0.103	0.158	0.204	0.227	0.222	0.193
12	6	0.000	0.000	0.004	0.016	0.040	0.079	0.128	0.177	0.212	0.226
12	7	0.000	0.000	0.001	0.003	0.011	0.029	0.059	0.101	0.149	0.193
12	8	0.000	0.000	0.000	0.001	0.002	0.008	0.020	0.042	0.076	0.121
12	9	0.000	0.000	0.000	0.000	0.000	0.001	0.005	0.012	0.028	0.054
12	10	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.007	0.016
12	11	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.003
12	12	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
13	0	0.513	0.254	0.121	0.055	0.024	0.010	0.004	0.001	0.000	0.000
13	1	0.351	0.367	0.277	0.179	0.103	0.054	0.026	0.011	0.004	0.002
13	2	0.111	0.245	0.294	0.268	0.206	0.139	0.084	0.045	0.022	0.010
13	3	0.021	0.100	0.190	0.246	0.252	0.218	0.165	0.111	0.066	0.035
13	4	0.003	0.028	0.084	0.154	0.210	0.234	0.222	0.184	0.135	0.087
13	5	0.000	0.006	0.027	0.069	0.126	0.180	0.215	0.221	0.199	0.157
13	6	0.000	0.001	0.006	0.023	0.056	0.103	0.155	0.197	0.217	0.209
13	7	0.000	0.000	0.001	0.006	0.019	0.044	0.083	0.131	0.177	0.209
13	8	0.000	0.000	0.000	0.001	0.005	0.014	0.034	0.066	0.109	0.157
13	9	0.000	0.000	0.000	0.000	0.001	0.003	0.010	0.024	0.050	0.087
13	10	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.006	0.016	0.035

Tabla II (continuación): Distribución binomial $B(n, p)$

		<i>p</i>									
<i>n</i>	<i>k</i>	0.050	0.100	0.150	0.200	0.250	0.300	0.350	0.400	0.450	0.500
17	15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
17	16	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
17	17	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
18	0	0.397	0.150	0.054	0.018	0.006	0.002	0.000	0.000	0.000	0.000
18	1	0.376	0.300	0.170	0.081	0.034	0.013	0.004	0.001	0.000	0.000
18	2	0.168	0.284	0.256	0.172	0.096	0.046	0.019	0.007	0.002	0.001
18	3	0.047	0.168	0.241	0.230	0.170	0.105	0.055	0.025	0.009	0.003
18	4	0.009	0.070	0.159	0.215	0.213	0.168	0.110	0.061	0.029	0.012
18	5	0.001	0.022	0.079	0.151	0.199	0.202	0.166	0.115	0.067	0.033
18	6	0.000	0.005	0.030	0.082	0.144	0.187	0.194	0.166	0.118	0.071
18	7	0.000	0.001	0.009	0.035	0.082	0.138	0.179	0.189	0.166	0.121
18	8	0.000	0.000	0.002	0.012	0.038	0.081	0.133	0.173	0.186	0.167
18	9	0.000	0.000	0.000	0.003	0.014	0.039	0.079	0.128	0.169	0.185
18	10	0.000	0.000	0.000	0.001	0.004	0.015	0.038	0.077	0.125	0.167
18	11	0.000	0.000	0.000	0.000	0.001	0.005	0.015	0.037	0.074	0.121
18	12	0.000	0.000	0.000	0.000	0.000	0.001	0.005	0.015	0.035	0.071
18	13	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.004	0.013	0.033
18	14	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.004	0.012
18	15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.003
18	16	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
18	17	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
18	18	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
19	0	0.377	0.135	0.046	0.014	0.004	0.001	0.000	0.000	0.000	0.000
19	1	0.377	0.285	0.153	0.068	0.027	0.009	0.003	0.001	0.000	0.000
19	2	0.179	0.285	0.243	0.154	0.080	0.036	0.014	0.005	0.001	0.000
19	3	0.053	0.180	0.243	0.218	0.152	0.087	0.042	0.017	0.006	0.002
19	4	0.011	0.080	0.171	0.218	0.202	0.149	0.091	0.047	0.020	0.007
19	5	0.002	0.027	0.091	0.164	0.202	0.192	0.147	0.093	0.050	0.022
19	6	0.000	0.007	0.037	0.095	0.157	0.192	0.184	0.145	0.095	0.052
19	7	0.000	0.001	0.012	0.044	0.097	0.153	0.184	0.180	0.144	0.096
19	8	0.000	0.000	0.003	0.017	0.049	0.098	0.149	0.180	0.177	0.144
19	9	0.000	0.000	0.001	0.005	0.020	0.051	0.098	0.146	0.177	0.176
19	10	0.000	0.000	0.000	0.001	0.007	0.022	0.053	0.098	0.145	0.176

Tabla III.1: Distribución $N(0, 1)$. Cuantiles (p. 85)



α	z_α								
0.01	2.576	0.21	1.254	0.41	0.824	0.61	0.510	0.81	0.240
0.02	2.326	0.22	1.227	0.42	0.806	0.62	0.496	0.82	0.228
0.03	2.170	0.23	1.200	0.43	0.789	0.63	0.482	0.83	0.215
0.04	2.054	0.24	1.175	0.44	0.772	0.64	0.468	0.84	0.202
0.05	1.960	0.25	1.150	0.45	0.755	0.65	0.454	0.85	0.189
0.06	1.881	0.26	1.126	0.46	0.739	0.66	0.440	0.86	0.176
0.07	1.812	0.27	1.103	0.47	0.722	0.67	0.426	0.87	0.164
0.08	1.751	0.28	1.080	0.48	0.706	0.68	0.412	0.88	0.151
0.09	1.695	0.29	1.058	0.49	0.690	0.69	0.399	0.89	0.138
0.10	1.645	0.30	1.036	0.50	0.674	0.70	0.385	0.90	0.126
0.11	1.598	0.31	1.015	0.51	0.659	0.71	0.372	0.91	0.113
0.12	1.555	0.32	0.994	0.52	0.643	0.72	0.358	0.92	0.100
0.13	1.514	0.33	0.974	0.53	0.628	0.73	0.345	0.93	0.088
0.14	1.476	0.34	0.954	0.54	0.613	0.74	0.332	0.94	0.075
0.15	1.440	0.35	0.935	0.55	0.598	0.75	0.319	0.95	0.063
0.16	1.405	0.36	0.915	0.56	0.583	0.76	0.305	0.96	0.050
0.17	1.372	0.37	0.896	0.57	0.568	0.77	0.292	0.97	0.038
0.18	1.341	0.38	0.878	0.58	0.553	0.78	0.279	0.98	0.025
0.19	1.311	0.39	0.860	0.59	0.539	0.79	0.266	0.99	0.013
0.20	1.282	0.40	0.842	0.60	0.524	0.80	0.253	1.00	0.000

Tabla III.2: Función de Distribución $N(0, 1)$ (p. 84)

$(F(x) = P(Z \leq x)$ si Z es una variable con distribución $N(0, 1)$)

x	$F(x)$	x	$F(x)$	x	$F(x)$	x	$F(x)$
0.01	0.5040	0.31	0.6217	0.61	0.7291	0.91	0.8186
0.02	0.5080	0.32	0.6255	0.62	0.7324	0.92	0.8212
0.03	0.5120	0.33	0.6293	0.63	0.7357	0.93	0.8238
0.04	0.5160	0.34	0.6331	0.64	0.7389	0.94	0.8264
0.05	0.5199	0.35	0.6368	0.65	0.7422	0.95	0.8289
0.06	0.5239	0.36	0.6406	0.66	0.7454	0.96	0.8315
0.07	0.5279	0.37	0.6443	0.67	0.7486	0.97	0.8340
0.08	0.5319	0.38	0.6480	0.68	0.7517	0.98	0.8365
0.09	0.5359	0.39	0.6517	0.69	0.7549	0.99	0.8389
0.10	0.5398	0.40	0.6554	0.70	0.7580	1.00	0.8413
0.11	0.5438	0.41	0.6591	0.71	0.7611	1.01	0.8438
0.12	0.5478	0.42	0.6628	0.72	0.7642	1.02	0.8461
0.13	0.5517	0.43	0.6664	0.73	0.7673	1.03	0.8485
0.14	0.5557	0.44	0.6700	0.74	0.7704	1.04	0.8508
0.15	0.5596	0.45	0.6736	0.75	0.7734	1.05	0.8531
0.16	0.5636	0.46	0.6772	0.76	0.7764	1.06	0.8554
0.17	0.5675	0.47	0.6808	0.77	0.7794	1.07	0.8577
0.18	0.5714	0.48	0.6844	0.78	0.7823	1.08	0.8599
0.19	0.5753	0.49	0.6879	0.79	0.7852	1.09	0.8621
0.20	0.5793	0.50	0.6915	0.80	0.7881	1.10	0.8643
0.21	0.5832	0.51	0.6950	0.81	0.7910	1.11	0.8665
0.22	0.5871	0.52	0.6985	0.82	0.7939	1.12	0.8686
0.23	0.5910	0.53	0.7019	0.83	0.7967	1.13	0.8708
0.24	0.5948	0.54	0.7054	0.84	0.7995	1.14	0.8729
0.25	0.5987	0.55	0.7088	0.85	0.8023	1.15	0.8749
0.26	0.6026	0.56	0.7123	0.86	0.8051	1.16	0.8770
0.27	0.6064	0.57	0.7157	0.87	0.8078	1.17	0.8790
0.28	0.6103	0.58	0.7190	0.88	0.8106	1.18	0.8810
0.29	0.6141	0.59	0.7224	0.89	0.8133	1.19	0.8830
0.30	0.6179	0.60	0.7257	0.90	0.8159	1.20	0.8849

Tabla III.2: Función de Distribución $N(0, 1)$ (cont.)

$(F(x) = P(Z \leq x)$ si Z es una variable con distribución $N(0, 1)$)

x	$F(x)$	x	$F(x)$	x	$F(x)$	x	$F(x)$
1.21	0.8869	1.51	0.9345	1.81	0.9649	2.11	0.9826
1.22	0.8888	1.52	0.9357	1.82	0.9656	2.12	0.9830
1.23	0.8907	1.53	0.9370	1.83	0.9664	2.13	0.9834
1.24	0.8925	1.54	0.9382	1.84	0.9671	2.14	0.9838
1.25	0.8944	1.55	0.9394	1.85	0.9678	2.15	0.9842
1.26	0.8962	1.56	0.9406	1.86	0.9686	2.16	0.9846
1.27	0.8980	1.57	0.9418	1.87	0.9693	2.17	0.9850
1.28	0.8997	1.58	0.9429	1.88	0.9699	2.18	0.9854
1.29	0.9015	1.59	0.9441	1.89	0.9706	2.19	0.9857
1.30	0.9032	1.60	0.9452	1.90	0.9713	2.20	0.9861
1.31	0.9049	1.61	0.9463	1.91	0.9719	2.21	0.9864
1.32	0.9066	1.62	0.9474	1.92	0.9726	2.22	0.9868
1.33	0.9082	1.63	0.9484	1.93	0.9732	2.23	0.9871
1.34	0.9099	1.64	0.9495	1.94	0.9738	2.24	0.9875
1.35	0.9115	1.65	0.9505	1.95	0.9744	2.25	0.9878
1.36	0.9131	1.66	0.9515	1.96	0.9750	2.26	0.9881
1.37	0.9147	1.67	0.9525	1.97	0.9756	2.27	0.9884
1.38	0.9162	1.68	0.9535	1.98	0.9761	2.28	0.9887
1.39	0.9177	1.69	0.9545	1.99	0.9767	2.29	0.9890
1.40	0.9192	1.70	0.9554	2.00	0.9772	2.30	0.9893
1.41	0.9207	1.71	0.9564	2.01	0.9778	2.31	0.9896
1.42	0.9222	1.72	0.9573	2.02	0.9783	2.32	0.9898
1.43	0.9236	1.73	0.9582	2.03	0.9788	2.33	0.9901
1.44	0.9251	1.74	0.9591	2.04	0.9793	2.34	0.9904
1.45	0.9265	1.75	0.9599	2.05	0.9798	2.35	0.9906
1.46	0.9279	1.76	0.9608	2.06	0.9803	2.36	0.9909
1.47	0.9292	1.77	0.9616	2.07	0.9808	2.37	0.9911
1.48	0.9306	1.78	0.9625	2.08	0.9812	2.38	0.9913
1.49	0.9319	1.79	0.9633	2.09	0.9817	2.39	0.9916
1.50	0.9332	1.80	0.9641	2.10	0.9821	2.40	0.9918

Tabla III.2: Función de Distribución $N(0, 1)$ (cont.)

$(F(x) = P(Z \leq x))$ si Z es una variable con distribución $N(0, 1)$

x	$F(x)$	x	$F(x)$	x	$F(x)$
2.41	0.9920	2.61	0.9955	2.81	0.9975
2.42	0.9922	2.62	0.9956	2.82	0.9976
2.43	0.9925	2.63	0.9957	2.83	0.9977
2.44	0.9927	2.64	0.9959	2.84	0.9977
2.45	0.9929	2.65	0.9960	2.85	0.9978
2.46	0.9931	2.66	0.9961	2.86	0.9979
2.47	0.9932	2.67	0.9962	2.87	0.9979
2.48	0.9934	2.68	0.9963	2.88	0.9980
2.49	0.9936	2.69	0.9964	2.89	0.9981
2.50	0.9938	2.70	0.9965	2.90	0.9981
2.51	0.9940	2.71	0.9966	2.91	0.9982
2.52	0.9941	2.72	0.9967	2.92	0.9982
2.53	0.9943	2.73	0.9968	2.93	0.9983
2.54	0.9945	2.74	0.9969	2.94	0.9984
2.55	0.9946	2.75	0.9970	2.95	0.9984
2.56	0.9948	2.76	0.9971	2.96	0.9985
2.57	0.9949	2.77	0.9972	2.97	0.9985
2.58	0.9951	2.78	0.9973	2.98	0.9986
2.59	0.9952	2.79	0.9974	2.99	0.9986
2.60	0.9953	2.80	0.9974	3.00	0.9987

Tabla IV: Distribución t de Student (p. 97)

$t_\alpha(n)$	α							
n	0.5	0.2	0.1	0.05	0.025	0.02	0.01	0.001
1	1.000	3.078	6.314	12.706	25.452	31.821	63.656	636.611
2	0.816	1.886	2.920	4.303	6.205	6.965	9.925	31.602
3	0.765	1.638	2.353	3.182	4.177	4.541	5.841	12.923
4	0.741	1.533	2.132	2.776	3.495	3.747	4.604	8.610
5	0.727	1.476	2.015	2.571	3.163	3.365	4.032	6.869
6	0.718	1.440	1.943	2.447	2.969	3.143	3.707	5.959
7	0.711	1.415	1.895	2.365	2.841	2.998	3.499	5.408
8	0.706	1.397	1.860	2.306	2.752	2.896	3.355	5.041
9	0.703	1.383	1.833	2.262	2.685	2.821	3.250	4.781
10	0.700	1.372	1.812	2.228	2.634	2.764	3.169	4.587
11	0.697	1.363	1.796	2.201	2.593	2.718	3.106	4.437
12	0.695	1.356	1.782	2.179	2.560	2.681	3.055	4.318
13	0.694	1.350	1.771	2.160	2.533	2.650	3.012	4.221
14	0.692	1.345	1.761	2.145	2.510	2.624	2.977	4.140
15	0.691	1.341	1.753	2.131	2.490	2.602	2.947	4.073
16	0.690	1.337	1.746	2.120	2.473	2.583	2.921	4.015
17	0.689	1.333	1.740	2.110	2.458	2.567	2.898	3.965
18	0.688	1.330	1.734	2.101	2.445	2.552	2.878	3.922
19	0.688	1.328	1.729	2.093	2.433	2.539	2.861	3.883
20	0.687	1.325	1.725	2.086	2.423	2.528	2.845	3.850
21	0.686	1.323	1.721	2.080	2.414	2.518	2.831	3.819
22	0.686	1.321	1.717	2.074	2.405	2.508	2.819	3.792

Tabla IV: Distribución t de Student (cont.)

$t_\alpha(n)$	α							
n	0.5	0.2	0.1	0.05	0.025	0.02	0.01	0.001
23	0.685	1.319	1.714	2.069	2.398	2.500	2.807	3.768
24	0.685	1.318	1.711	2.064	2.391	2.492	2.797	3.745
25	0.684	1.316	1.708	2.060	2.385	2.485	2.787	3.725
26	0.684	1.315	1.706	2.056	2.379	2.479	2.779	3.707
27	0.684	1.314	1.703	2.052	2.373	2.473	2.771	3.690
28	0.683	1.313	1.701	2.048	2.368	2.467	2.763	3.674
29	0.683	1.311	1.699	2.045	2.364	2.462	2.756	3.659
30	0.683	1.310	1.697	2.042	2.360	2.457	2.750	3.646
35	0.682	1.306	1.690	2.030	2.342	2.438	2.724	3.591
40	0.681	1.303	1.684	2.021	2.329	2.423	2.704	3.551
45	0.680	1.301	1.679	2.014	2.319	2.412	2.690	3.520
50	0.679	1.299	1.676	2.009	2.311	2.403	2.678	3.496
55	0.679	1.297	1.673	2.004	2.304	2.396	2.668	3.476
60	0.679	1.296	1.671	2.000	2.299	2.390	2.660	3.460
65	0.678	1.295	1.669	1.997	2.295	2.385	2.654	3.447
70	0.678	1.294	1.667	1.994	2.291	2.381	2.648	3.435
75	0.678	1.293	1.665	1.992	2.287	2.377	2.643	3.425
80	0.678	1.292	1.664	1.990	2.284	2.374	2.639	3.416
85	0.677	1.292	1.663	1.988	2.282	2.371	2.635	3.409
90	0.677	1.291	1.662	1.987	2.280	2.368	2.632	3.402
95	0.677	1.291	1.661	1.985	2.277	2.366	2.629	3.396
100	0.677	1.290	1.660	1.984	2.276	2.364	2.626	3.390
∞	0.674	1.282	1.645	1.960	2.241	2.326	2.576	3.291

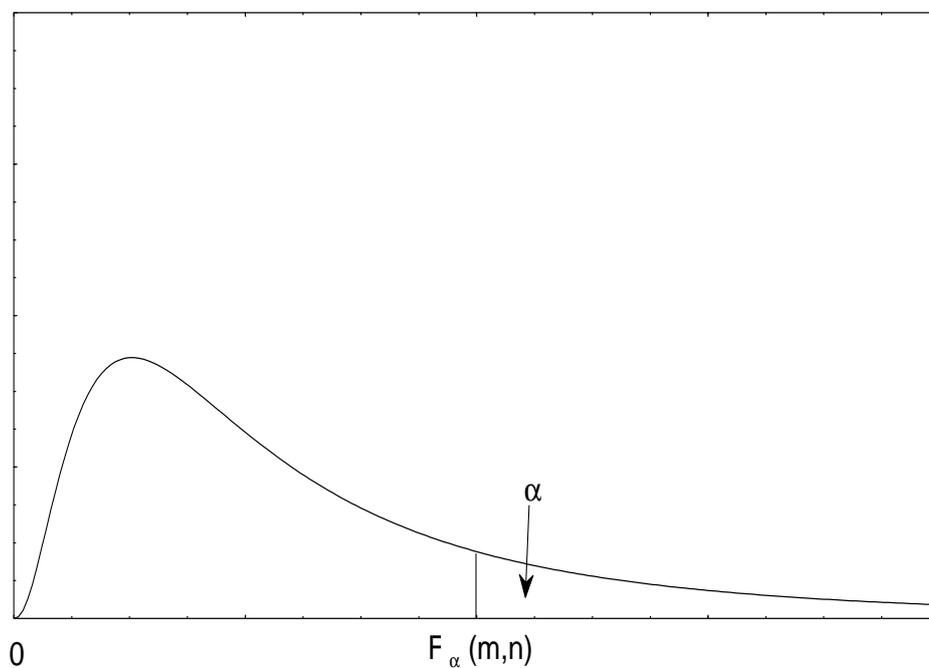
Tabla V: Distribución chi-cuadrado $\chi^2(n)$ (p. 166)

$\chi^2_\alpha(n)$	α													
n	0.001	0.005	0.01	0.025	0.05	0.1	0.2	0.8	0.9	0.95	0.975	0.99	0.995	
1	10.83	7.88	6.64	5.02	3.84	2.71	1.64	0.06	0.02	0.00	0.00	0.00	0.00	
2	13.82	10.60	9.21	7.38	5.99	4.61	3.22	0.45	0.21	0.10	0.05	0.02	0.01	
3	16.27	12.84	11.35	9.35	7.82	6.25	4.64	1.01	0.58	0.35	0.22	0.11	0.07	
4	18.47	14.86	13.28	11.14	9.49	7.78	5.99	1.65	1.06	0.71	0.48	0.30	0.21	
5	20.52	16.75	15.09	12.83	11.07	9.24	7.29	2.34	1.61	1.15	0.83	0.55	0.41	
6	22.46	18.55	16.81	14.45	12.59	10.65	8.56	3.07	2.20	1.64	1.24	0.87	0.68	
7	24.33	20.28	18.48	16.01	14.07	12.02	9.80	3.82	2.83	2.17	1.69	1.24	0.99	
8	26.13	21.96	20.09	17.54	15.51	13.36	11.03	4.59	3.49	2.73	2.18	1.65	1.34	
9	27.88	23.59	21.67	19.03	16.92	14.68	12.24	5.38	4.17	3.32	2.70	2.09	1.73	
10	29.59	25.19	23.21	20.49	18.31	15.99	13.44	6.18	4.87	3.94	3.25	2.56	2.16	
11	31.27	26.76	24.73	21.92	19.68	17.28	14.63	6.99	5.58	4.57	3.82	3.05	2.60	
12	32.92	28.30	26.22	23.34	21.03	18.55	15.81	7.81	6.30	5.23	4.40	3.57	3.07	
13	34.54	29.83	27.69	24.74	22.36	19.81	16.99	8.63	7.04	5.89	5.01	4.11	3.56	
14	36.13	31.33	29.15	26.12	23.69	21.07	18.15	9.47	7.79	6.57	5.63	4.66	4.07	
15	37.71	32.81	30.58	27.49	25.00	22.31	19.31	10.31	8.55	7.26	6.26	5.23	4.60	
16	39.26	34.27	32.01	28.85	26.30	23.54	20.47	11.15	9.31	7.96	6.91	5.81	5.14	
17	40.80	35.73	33.41	30.20	27.59	24.77	21.62	12.00	10.08	8.67	7.56	6.41	5.70	
18	42.32	37.16	34.81	31.53	28.87	25.99	22.76	12.86	10.86	9.39	8.23	7.01	6.26	

Tabla V: Distribución chi-cuadrado $\chi^2(n)$ (cont.)

$\chi^2_\alpha(n)$	α												
n	0.001	0.005	0.01	0.025	0.05	0.1	0.2	0.8	0.9	0.95	0.975	0.99	0.995
19	43.83	38.59	36.20	32.86	30.15	27.21	23.90	13.72	11.65	10.12	8.91	7.63	6.84
20	45.33	40.01	37.57	34.18	31.41	28.41	25.04	14.58	12.44	10.85	9.59	8.26	7.43
21	46.81	41.41	38.94	35.48	32.68	29.62	26.17	15.44	13.24	11.59	10.28	8.90	8.03
22	48.28	42.81	40.30	36.79	33.93	30.82	27.30	16.31	14.04	12.34	10.98	9.54	8.64
23	49.74	44.19	41.65	38.08	35.18	32.01	28.43	17.19	14.85	13.09	11.69	10.19	9.26
24	51.19	45.57	42.99	39.37	36.42	33.20	29.56	18.06	15.66	13.85	12.40	10.86	9.89
25	52.64	46.94	44.32	40.65	37.66	34.39	30.68	18.94	16.47	14.61	13.12	11.52	10.52
26	54.07	48.30	45.65	41.93	38.89	35.57	31.80	19.82	17.29	15.38	13.84	12.20	11.16
27	55.49	49.66	46.97	43.20	40.12	36.75	32.91	20.70	18.11	16.15	14.57	12.88	11.81
28	56.91	51.01	48.29	44.47	41.34	37.92	34.03	21.59	18.94	16.93	15.31	13.56	12.46
29	58.32	52.35	49.60	45.73	42.56	39.09	35.14	22.47	19.77	17.71	16.05	14.25	13.12
30	59.72	53.69	50.90	46.99	43.78	40.26	36.25	23.36	20.60	18.49	16.79	14.95	13.78
40	73.43	66.79	63.71	59.35	55.77	51.81	47.27	32.34	29.05	26.51	24.43	22.16	20.70
50	86.69	79.52	76.18	71.44	67.52	63.18	58.17	41.45	37.68	34.76	32.35	29.70	27.99
60	99.65	91.98	88.41	83.32	79.10	74.41	68.98	50.64	46.45	43.18	40.47	37.48	35.53
70	112.37	104.25	100.46	95.05	90.55	85.54	79.72	59.89	55.32	51.73	48.75	45.43	43.26
80	124.90	116.37	112.37	106.66	101.90	96.60	90.42	69.20	64.27	60.38	57.14	53.53	51.16
90	137.28	128.35	124.16	118.17	113.17	107.59	101.07	78.55	73.28	69.11	65.63	61.74	59.18
100	149.53	140.23	135.86	129.60	124.37	118.52	111.68	87.94	82.35	77.91	74.20	70.05	67.31

Tabla VI: Distribución $F(m, n)$ de Fisher (p. 95)



Nota: En cada una de las tablas de la distribución $F(m, n)$ siguientes, los grados de libertad m del numerador aparecen en la primera columna, mientras que los grados de libertad n del denominador aparecen en la primera fila. Las tablas proporcionan los valores $F_\alpha(m, n)$.

Tabla VI.1: Distribución $F(m, n)$. $\alpha = 0.1$

1^a fila= g.l. denominador. 1^a columna= g.l. numerador

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	39.86	8.53	5.54	4.54	4.06	3.78	3.59	3.46	3.36	3.29	3.23	3.18	3.14	3.10	3.07	3.05	3.03	3.01	2.99	2.97
2	49.50	9.00	5.46	4.32	3.78	3.46	3.26	3.11	3.01	2.92	2.86	2.81	2.76	2.73	2.70	2.67	2.64	2.62	2.61	2.59
3	53.59	9.16	5.39	4.19	3.62	3.29	3.07	2.92	2.81	2.73	2.66	2.61	2.56	2.52	2.49	2.46	2.44	2.42	2.40	2.38
4	55.83	9.24	5.34	4.11	3.52	3.18	2.96	2.81	2.69	2.61	2.54	2.48	2.43	2.39	2.36	2.33	2.31	2.29	2.27	2.25
5	57.24	9.29	5.31	4.05	3.45	3.11	2.88	2.73	2.61	2.52	2.45	2.39	2.35	2.31	2.27	2.24	2.22	2.20	2.18	2.16
6	58.20	9.33	5.28	4.01	3.40	3.05	2.83	2.67	2.55	2.46	2.39	2.33	2.28	2.24	2.21	2.18	2.15	2.13	2.11	2.09
7	58.91	9.35	5.27	3.98	3.37	3.01	2.78	2.62	2.51	2.41	2.34	2.28	2.23	2.19	2.16	2.13	2.10	2.08	2.06	2.04
8	59.44	9.37	5.25	3.95	3.34	2.98	2.75	2.59	2.47	2.38	2.30	2.24	2.20	2.15	2.12	2.09	2.06	2.04	2.02	2.00
9	59.86	9.38	5.24	3.94	3.32	2.96	2.72	2.56	2.44	2.35	2.27	2.21	2.16	2.12	2.09	2.06	2.03	2.00	1.98	1.96
10	60.20	9.39	5.23	3.92	3.30	2.94	2.70	2.54	2.42	2.32	2.25	2.19	2.14	2.10	2.06	2.03	2.00	1.98	1.96	1.94
11	60.47	9.40	5.22	3.91	3.28	2.92	2.68	2.52	2.40	2.30	2.23	2.17	2.12	2.07	2.04	2.01	1.98	1.95	1.93	1.91
12	60.71	9.41	5.22	3.90	3.27	2.90	2.67	2.50	2.38	2.28	2.21	2.15	2.10	2.05	2.02	1.99	1.96	1.93	1.91	1.89
13	60.90	9.41	5.21	3.89	3.26	2.89	2.65	2.49	2.36	2.27	2.19	2.13	2.08	2.04	2.00	1.97	1.94	1.92	1.89	1.87
14	61.07	9.42	5.20	3.88	3.25	2.88	2.64	2.48	2.35	2.26	2.18	2.12	2.07	2.02	1.99	1.95	1.93	1.90	1.88	1.86
15	61.22	9.42	5.20	3.87	3.24	2.87	2.63	2.46	2.34	2.24	2.17	2.10	2.05	2.01	1.97	1.94	1.91	1.89	1.86	1.84
16	61.35	9.43	5.20	3.86	3.23	2.86	2.62	2.45	2.33	2.23	2.16	2.09	2.04	2.00	1.96	1.93	1.90	1.87	1.85	1.83
17	61.46	9.43	5.19	3.86	3.22	2.85	2.61	2.45	2.32	2.22	2.15	2.08	2.03	1.99	1.95	1.92	1.89	1.86	1.84	1.82
18	61.57	9.44	5.19	3.85	3.22	2.85	2.61	2.44	2.31	2.22	2.14	2.08	2.02	1.98	1.94	1.91	1.88	1.85	1.83	1.81
19	61.66	9.44	5.19	3.85	3.21	2.84	2.60	2.43	2.30	2.21	2.13	2.07	2.01	1.97	1.93	1.90	1.87	1.84	1.82	1.80
20	61.74	9.44	5.18	3.84	3.21	2.84	2.59	2.42	2.30	2.20	2.12	2.06	2.01	1.96	1.92	1.89	1.86	1.84	1.81	1.79
21	61.81	9.44	5.18	3.84	3.20	2.83	2.59	2.42	2.29	2.19	2.12	2.05	2.00	1.96	1.92	1.88	1.86	1.83	1.81	1.79
22	61.88	9.45	5.18	3.84	3.20	2.83	2.58	2.41	2.29	2.19	2.11	2.05	1.99	1.95	1.91	1.88	1.85	1.82	1.80	1.78
23	61.95	9.45	5.18	3.83	3.19	2.82	2.58	2.41	2.28	2.18	2.11	2.04	1.99	1.94	1.90	1.87	1.84	1.82	1.79	1.77
24	62.00	9.45	5.18	3.83	3.19	2.82	2.58	2.40	2.28	2.18	2.10	2.04	1.98	1.94	1.90	1.87	1.84	1.81	1.79	1.77
25	62.05	9.45	5.17	3.83	3.19	2.81	2.57	2.40	2.27	2.17	2.10	2.03	1.98	1.93	1.89	1.86	1.83	1.80	1.78	1.76
26	62.10	9.45	5.17	3.83	3.18	2.81	2.57	2.40	2.27	2.17	2.09	2.03	1.97	1.93	1.89	1.86	1.83	1.80	1.78	1.76
27	62.15	9.45	5.17	3.82	3.18	2.81	2.56	2.39	2.26	2.17	2.09	2.02	1.97	1.92	1.88	1.85	1.82	1.80	1.77	1.75
28	62.19	9.46	5.17	3.82	3.18	2.81	2.56	2.39	2.26	2.16	2.08	2.02	1.96	1.92	1.88	1.85	1.82	1.79	1.77	1.75
29	62.23	9.46	5.17	3.82	3.18	2.80	2.56	2.39	2.26	2.16	2.08	2.01	1.96	1.92	1.88	1.84	1.81	1.79	1.76	1.74
30	62.26	9.46	5.17	3.82	3.17	2.80	2.56	2.38	2.25	2.16	2.08	2.01	1.96	1.91	1.87	1.84	1.81	1.78	1.76	1.74
40	62.53	9.47	5.16	3.80	3.16	2.78	2.54	2.36	2.23	2.13	2.05	1.99	1.93	1.89	1.85	1.81	1.78	1.75	1.73	1.71
50	62.69	9.47	5.15	3.80	3.15	2.77	2.52	2.35	2.22	2.12	2.04	1.97	1.92	1.87	1.83	1.79	1.76	1.74	1.71	1.69
60	62.79	9.47	5.15	3.79	3.14	2.76	2.51	2.34	2.21	2.11	2.03	1.96	1.90	1.86	1.82	1.78	1.75	1.72	1.70	1.68
70	62.87	9.48	5.15	3.79	3.14	2.76	2.51	2.33	2.20	2.10	2.02	1.95	1.90	1.85	1.81	1.77	1.74	1.71	1.69	1.67
80	62.93	9.48	5.15	3.78	3.13	2.75	2.50	2.33	2.20	2.09	2.01	1.95	1.89	1.84	1.80	1.77	1.74	1.71	1.68	1.66
90	62.97	9.48	5.15	3.78	3.13	2.75	2.50	2.32	2.19	2.09	2.01	1.94	1.89	1.84	1.80	1.76	1.73	1.70	1.68	1.65
100	63.01	9.48	5.14	3.78	3.13	2.75	2.50	2.32	2.19	2.09	2.01	1.94	1.88	1.83	1.79	1.76	1.73	1.70	1.67	1.65
120	63.06	9.48	5.14	3.78	3.12	2.74	2.49	2.32	2.18	2.08	2.00	1.93	1.88	1.83	1.79	1.75	1.72	1.69	1.67	1.64
∞	63.32	9.49	5.13	3.76	3.11	2.72	2.47	2.29	2.16	2.06	1.97	1.90	1.85	1.80	1.76	1.72	1.69	1.66	1.63	1.61

Tabla VI.1 (cont.): Distribución $F(m, n)$. $\alpha = 0.1$

1^a fila= g.l. denominador. 1^a columna= g.l. numerador

	21	22	23	24	25	26	27	28	29	30	40	50	60	70	80	90	100	110	120	∞
1	2.96	2.95	2.94	2.93	2.92	2.91	2.90	2.89	2.89	2.88	2.84	2.81	2.79	2.78	2.77	2.76	2.76	2.75	2.75	2.71
2	2.57	2.56	2.55	2.54	2.53	2.52	2.51	2.50	2.50	2.49	2.44	2.41	2.39	2.38	2.37	2.36	2.36	2.35	2.35	2.30
3	2.36	2.35	2.34	2.33	2.32	2.31	2.30	2.29	2.28	2.28	2.23	2.20	2.18	2.16	2.15	2.15	2.14	2.13	2.13	2.08
4	2.23	2.22	2.21	2.19	2.18	2.17	2.17	2.16	2.15	2.14	2.09	2.06	2.04	2.03	2.02	2.01	2.00	2.00	1.99	1.95
5	2.14	2.13	2.11	2.10	2.09	2.08	2.07	2.06	2.06	2.05	2.00	1.97	1.95	1.93	1.92	1.91	1.91	1.90	1.90	1.85
6	2.08	2.06	2.05	2.04	2.02	2.01	2.00	2.00	1.99	1.98	1.93	1.90	1.87	1.86	1.85	1.84	1.83	1.83	1.82	1.77
7	2.02	2.01	1.99	1.98	1.97	1.96	1.95	1.94	1.93	1.93	1.87	1.84	1.82	1.80	1.79	1.78	1.78	1.77	1.77	1.72
8	1.98	1.97	1.95	1.94	1.93	1.92	1.91	1.90	1.89	1.88	1.83	1.80	1.77	1.76	1.75	1.74	1.73	1.73	1.72	1.67
9	1.95	1.93	1.92	1.91	1.89	1.88	1.87	1.87	1.86	1.85	1.79	1.76	1.74	1.72	1.71	1.70	1.69	1.69	1.68	1.63
10	1.92	1.90	1.89	1.88	1.87	1.86	1.85	1.84	1.83	1.82	1.76	1.73	1.71	1.69	1.68	1.67	1.66	1.66	1.65	1.60
11	1.90	1.88	1.87	1.85	1.84	1.83	1.82	1.81	1.80	1.79	1.74	1.70	1.68	1.66	1.65	1.64	1.64	1.63	1.63	1.57
12	1.87	1.86	1.84	1.83	1.82	1.81	1.80	1.79	1.78	1.77	1.71	1.68	1.66	1.64	1.63	1.62	1.61	1.61	1.60	1.55
13	1.86	1.84	1.83	1.81	1.80	1.79	1.78	1.77	1.76	1.75	1.70	1.66	1.64	1.62	1.61	1.60	1.59	1.59	1.58	1.52
14	1.84	1.83	1.81	1.80	1.79	1.77	1.76	1.75	1.75	1.74	1.68	1.64	1.62	1.60	1.59	1.58	1.57	1.57	1.56	1.51
15	1.83	1.81	1.80	1.78	1.77	1.76	1.75	1.74	1.73	1.72	1.66	1.63	1.60	1.59	1.57	1.56	1.56	1.55	1.55	1.49
16	1.81	1.80	1.78	1.77	1.76	1.75	1.74	1.73	1.72	1.71	1.65	1.61	1.59	1.57	1.56	1.55	1.54	1.54	1.53	1.47
17	1.80	1.79	1.77	1.76	1.75	1.73	1.72	1.71	1.71	1.70	1.64	1.60	1.58	1.56	1.55	1.54	1.53	1.52	1.52	1.46
18	1.79	1.78	1.76	1.75	1.74	1.72	1.71	1.70	1.69	1.69	1.62	1.59	1.56	1.55	1.53	1.52	1.52	1.51	1.50	1.44
19	1.78	1.77	1.75	1.74	1.73	1.71	1.70	1.69	1.68	1.68	1.61	1.58	1.55	1.54	1.52	1.51	1.50	1.50	1.49	1.43
20	1.78	1.76	1.74	1.73	1.72	1.71	1.70	1.69	1.68	1.67	1.61	1.57	1.54	1.53	1.51	1.50	1.49	1.49	1.48	1.42
21	1.77	1.75	1.74	1.72	1.71	1.70	1.69	1.68	1.67	1.66	1.60	1.56	1.53	1.52	1.50	1.49	1.48	1.48	1.47	1.41
22	1.76	1.74	1.73	1.71	1.70	1.69	1.68	1.67	1.66	1.65	1.59	1.55	1.53	1.51	1.49	1.48	1.48	1.47	1.46	1.40
23	1.75	1.74	1.72	1.71	1.70	1.68	1.67	1.66	1.65	1.64	1.58	1.54	1.52	1.50	1.49	1.48	1.47	1.46	1.46	1.39
24	1.75	1.73	1.72	1.70	1.69	1.68	1.67	1.66	1.65	1.64	1.57	1.54	1.51	1.49	1.48	1.47	1.46	1.45	1.45	1.38
25	1.74	1.73	1.71	1.70	1.68	1.67	1.66	1.65	1.64	1.63	1.57	1.53	1.50	1.49	1.47	1.46	1.45	1.45	1.44	1.38
26	1.74	1.72	1.70	1.69	1.68	1.67	1.65	1.64	1.63	1.63	1.56	1.52	1.50	1.48	1.47	1.45	1.45	1.44	1.43	1.37
27	1.73	1.72	1.70	1.69	1.67	1.66	1.65	1.64	1.63	1.62	1.56	1.52	1.49	1.47	1.46	1.45	1.44	1.43	1.43	1.36
28	1.73	1.71	1.69	1.68	1.67	1.66	1.64	1.63	1.62	1.62	1.55	1.51	1.49	1.47	1.45	1.44	1.43	1.43	1.42	1.35
29	1.72	1.71	1.69	1.68	1.66	1.65	1.64	1.63	1.62	1.61	1.55	1.51	1.48	1.46	1.45	1.44	1.43	1.42	1.41	1.35
30	1.72	1.70	1.69	1.67	1.66	1.65	1.64	1.63	1.62	1.61	1.54	1.50	1.48	1.46	1.44	1.43	1.42	1.42	1.41	1.34
40	1.69	1.67	1.66	1.64	1.63	1.61	1.60	1.59	1.58	1.57	1.51	1.46	1.44	1.42	1.40	1.39	1.38	1.37	1.37	1.34
50	1.67	1.65	1.64	1.62	1.61	1.59	1.58	1.57	1.56	1.55	1.48	1.44	1.41	1.39	1.38	1.36	1.35	1.35	1.34	1.33
60	1.66	1.64	1.62	1.61	1.59	1.58	1.57	1.56	1.55	1.54	1.47	1.42	1.40	1.37	1.36	1.35	1.34	1.33	1.32	1.33
70	1.65	1.63	1.61	1.60	1.58	1.57	1.56	1.55	1.54	1.53	1.46	1.41	1.38	1.36	1.34	1.33	1.32	1.31	1.31	1.32
80	1.64	1.62	1.61	1.59	1.58	1.56	1.55	1.54	1.53	1.52	1.45	1.40	1.37	1.35	1.33	1.32	1.31	1.30	1.29	1.32
90	1.63	1.62	1.60	1.58	1.57	1.56	1.54	1.53	1.52	1.51	1.44	1.39	1.36	1.34	1.33	1.31	1.30	1.29	1.28	1.31
100	1.63	1.61	1.59	1.58	1.56	1.55	1.54	1.53	1.52	1.51	1.43	1.39	1.36	1.34	1.32	1.30	1.29	1.28	1.28	1.31
120	1.62	1.60	1.59	1.57	1.56	1.54	1.53	1.52	1.51	1.50	1.42	1.38	1.35	1.32	1.31	1.29	1.28	1.27	1.26	1.30
∞	1.59	1.57	1.55	1.53	1.52	1.50	1.49	1.48	1.47	1.46	1.45	1.44	1.43	1.42	1.41	1.40	1.40	1.39	1.38	1.03

Tabla VI.2: Distribución $F(m, n)$. $\alpha = 0.05$

1ª fila= g.l. denominador. 1ª columna= g.l. numerador.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	161.45	18.51	10.13	7.71	6.61	5.99	5.59	5.32	5.12	4.96	4.84	4.75	4.67	4.60	4.54	4.49	4.45	4.41	4.38	4.35
2	199.50	19.00	9.55	6.94	5.79	5.14	4.74	4.46	4.26	4.10	3.98	3.89	3.81	3.74	3.68	3.63	3.59	3.55	3.52	3.49
3	215.71	19.16	9.28	6.59	5.41	4.76	4.35	4.07	3.86	3.71	3.59	3.49	3.41	3.34	3.29	3.24	3.20	3.16	3.13	3.10
4	224.58	19.25	9.12	6.39	5.19	4.53	4.12	3.84	3.63	3.48	3.36	3.26	3.18	3.11	3.06	3.01	2.96	2.93	2.90	2.87
5	230.16	19.30	9.01	6.26	5.05	4.39	3.97	3.69	3.48	3.33	3.20	3.11	3.03	2.96	2.90	2.85	2.81	2.77	2.74	2.71
6	233.99	19.33	8.94	6.16	4.95	4.28	3.87	3.58	3.37	3.22	3.09	3.00	2.92	2.85	2.79	2.74	2.70	2.66	2.63	2.60
7	236.77	19.35	8.89	6.09	4.88	4.21	3.79	3.50	3.29	3.14	3.01	2.91	2.83	2.76	2.71	2.66	2.61	2.58	2.54	2.51
8	238.88	19.37	8.85	6.04	4.82	4.15	3.73	3.44	3.23	3.07	2.95	2.85	2.77	2.70	2.64	2.59	2.55	2.51	2.48	2.45
9	240.54	19.38	8.81	6.00	4.77	4.10	3.68	3.39	3.18	3.02	2.90	2.80	2.71	2.65	2.59	2.54	2.49	2.46	2.42	2.39
10	241.88	19.40	8.79	5.96	4.74	4.06	3.64	3.35	3.14	2.98	2.85	2.75	2.67	2.60	2.54	2.49	2.45	2.41	2.38	2.35
11	242.98	19.40	8.76	5.94	4.70	4.03	3.60	3.31	3.10	2.94	2.82	2.72	2.63	2.57	2.51	2.46	2.41	2.37	2.34	2.31
12	243.91	19.41	8.74	5.91	4.68	4.00	3.57	3.28	3.07	2.91	2.79	2.69	2.60	2.53	2.48	2.42	2.38	2.34	2.31	2.28
13	244.69	19.42	8.73	5.89	4.66	3.98	3.55	3.26	3.05	2.89	2.76	2.66	2.58	2.51	2.45	2.40	2.35	2.31	2.28	2.25
14	245.36	19.42	8.71	5.87	4.64	3.96	3.53	3.24	3.03	2.86	2.74	2.64	2.55	2.48	2.42	2.37	2.33	2.29	2.26	2.22
15	245.95	19.43	8.70	5.86	4.62	3.94	3.51	3.22	3.01	2.85	2.72	2.62	2.53	2.46	2.40	2.35	2.31	2.27	2.23	2.20
16	246.46	19.43	8.69	5.84	4.60	3.92	3.49	3.20	2.99	2.83	2.70	2.60	2.51	2.44	2.38	2.33	2.29	2.25	2.21	2.18
17	246.92	19.44	8.68	5.83	4.59	3.91	3.48	3.19	2.97	2.81	2.69	2.58	2.50	2.43	2.37	2.32	2.27	2.23	2.20	2.17
18	247.32	19.44	8.67	5.82	4.58	3.90	3.47	3.17	2.96	2.80	2.67	2.57	2.48	2.41	2.35	2.30	2.26	2.22	2.18	2.15
19	247.69	19.44	8.67	5.81	4.57	3.88	3.46	3.16	2.95	2.79	2.66	2.56	2.47	2.40	2.34	2.29	2.24	2.20	2.17	2.14
20	248.01	19.45	8.66	5.80	4.56	3.87	3.44	3.15	2.94	2.77	2.65	2.54	2.46	2.39	2.33	2.28	2.23	2.19	2.16	2.12
21	248.31	19.45	8.65	5.79	4.55	3.86	3.43	3.14	2.93	2.76	2.64	2.53	2.45	2.38	2.32	2.26	2.22	2.18	2.14	2.11
22	248.58	19.45	8.65	5.79	4.54	3.86	3.43	3.13	2.92	2.75	2.63	2.52	2.44	2.37	2.31	2.25	2.21	2.17	2.13	2.10
23	248.83	19.45	8.64	5.78	4.53	3.85	3.42	3.12	2.91	2.75	2.62	2.51	2.43	2.36	2.30	2.24	2.20	2.16	2.12	2.09
24	249.05	19.45	8.64	5.77	4.53	3.84	3.41	3.12	2.90	2.74	2.61	2.51	2.42	2.35	2.29	2.24	2.19	2.15	2.11	2.08
25	249.26	19.46	8.63	5.77	4.52	3.83	3.40	3.11	2.89	2.73	2.60	2.50	2.41	2.34	2.28	2.23	2.18	2.14	2.11	2.07
26	249.45	19.46	8.63	5.76	4.52	3.83	3.40	3.10	2.89	2.72	2.59	2.49	2.41	2.33	2.27	2.22	2.17	2.13	2.10	2.07
27	249.63	19.46	8.63	5.76	4.51	3.82	3.39	3.10	2.88	2.72	2.59	2.48	2.40	2.33	2.27	2.21	2.17	2.13	2.09	2.06
28	249.80	19.46	8.62	5.75	4.50	3.82	3.39	3.09	2.87	2.71	2.58	2.48	2.39	2.32	2.26	2.21	2.16	2.12	2.08	2.05
29	249.95	19.46	8.62	5.75	4.50	3.81	3.38	3.08	2.87	2.70	2.58	2.47	2.39	2.31	2.25	2.20	2.15	2.11	2.08	2.05
30	250.09	19.46	8.62	5.75	4.50	3.81	3.38	3.08	2.86	2.70	2.57	2.47	2.38	2.31	2.25	2.19	2.15	2.11	2.07	2.04
40	251.14	19.47	8.59	5.72	4.46	3.77	3.34	3.04	2.83	2.66	2.53	2.43	2.34	2.27	2.20	2.15	2.10	2.06	2.03	1.99
50	251.77	19.48	8.58	5.70	4.44	3.75	3.32	3.02	2.80	2.64	2.51	2.40	2.31	2.24	2.18	2.12	2.08	2.04	2.00	1.97
60	252.20	19.48	8.57	5.69	4.43	3.74	3.30	3.01	2.79	2.62	2.49	2.38	2.30	2.22	2.16	2.11	2.06	2.02	1.98	1.95
70	252.50	19.48	8.57	5.68	4.42	3.73	3.29	2.99	2.78	2.61	2.48	2.37	2.28	2.21	2.15	2.09	2.05	2.00	1.97	1.93
80	252.72	19.48	8.56	5.67	4.41	3.72	3.29	2.99	2.77	2.60	2.47	2.36	2.27	2.20	2.14	2.08	2.03	1.99	1.96	1.92
90	252.90	19.48	8.56	5.67	4.41	3.72	3.28	2.98	2.76	2.59	2.46	2.36	2.27	2.19	2.13	2.07	2.03	1.98	1.95	1.91
100	253.04	19.49	8.55	5.66	4.41	3.71	3.27	2.97	2.76	2.59	2.46	2.35	2.26	2.19	2.12	2.07	2.02	1.98	1.94	1.91
120	253.25	19.49	8.55	5.66	4.40	3.70	3.27	2.97	2.75	2.58	2.45	2.34	2.25	2.18	2.11	2.06	2.01	1.97	1.93	1.90
∞	254.30	19.50	8.53	5.63	4.37	3.67	3.23	2.93	2.71	2.54	2.41	2.30	2.21	2.13	2.07	2.01	1.96	1.92	1.88	1.84

Tabla VI.2 (cont.): Distribución $F(m, n)$. $\alpha = 0.05$

1^a fila= g.l. denominador. 1^a columna= g.l. numerador.

	21	22	23	24	25	26	27	28	29	30	40	50	60	70	80	90	100	110	120	∞
1	4.32	4.30	4.28	4.26	4.24	4.23	4.21	4.20	4.18	4.17	4.08	4.03	4.00	3.98	3.96	3.95	3.94	3.93	3.92	3.84
2	3.47	3.44	3.42	3.40	3.39	3.37	3.35	3.34	3.33	3.32	3.23	3.18	3.15	3.13	3.11	3.10	3.09	3.08	3.07	3.00
3	3.07	3.05	3.03	3.01	2.99	2.98	2.96	2.95	2.93	2.92	2.84	2.79	2.76	2.74	2.72	2.71	2.70	2.69	2.68	2.61
4	2.84	2.82	2.80	2.78	2.76	2.74	2.73	2.71	2.70	2.69	2.61	2.56	2.53	2.50	2.49	2.47	2.46	2.45	2.45	2.37
5	2.68	2.66	2.64	2.62	2.60	2.59	2.57	2.56	2.55	2.53	2.45	2.40	2.37	2.35	2.33	2.32	2.31	2.30	2.29	2.21
6	2.57	2.55	2.53	2.51	2.49	2.47	2.46	2.45	2.43	2.42	2.34	2.29	2.25	2.23	2.21	2.20	2.19	2.18	2.18	2.10
7	2.49	2.46	2.44	2.42	2.40	2.39	2.37	2.36	2.35	2.33	2.25	2.20	2.17	2.14	2.13	2.11	2.10	2.09	2.09	2.01
8	2.42	2.40	2.37	2.36	2.34	2.32	2.31	2.29	2.28	2.27	2.18	2.13	2.10	2.07	2.06	2.04	2.03	2.02	2.02	1.94
9	2.37	2.34	2.32	2.30	2.28	2.27	2.25	2.24	2.22	2.21	2.12	2.07	2.04	2.02	2.00	1.99	1.97	1.97	1.96	1.88
10	2.32	2.30	2.27	2.25	2.24	2.22	2.20	2.19	2.18	2.16	2.08	2.03	1.99	1.97	1.95	1.94	1.93	1.92	1.91	1.83
11	2.28	2.26	2.24	2.22	2.20	2.18	2.17	2.15	2.14	2.13	2.04	1.99	1.95	1.93	1.91	1.90	1.89	1.88	1.87	1.79
12	2.25	2.23	2.20	2.18	2.16	2.15	2.13	2.12	2.10	2.09	2.00	1.95	1.92	1.89	1.88	1.86	1.85	1.84	1.83	1.75
13	2.22	2.20	2.18	2.15	2.14	2.12	2.10	2.09	2.08	2.06	1.97	1.92	1.89	1.86	1.84	1.83	1.82	1.81	1.80	1.72
14	2.20	2.17	2.15	2.13	2.11	2.09	2.08	2.06	2.05	2.04	1.95	1.89	1.86	1.84	1.82	1.80	1.79	1.78	1.78	1.69
15	2.18	2.15	2.13	2.11	2.09	2.07	2.06	2.04	2.03	2.01	1.92	1.87	1.84	1.81	1.79	1.78	1.77	1.76	1.75	1.67
16	2.16	2.13	2.11	2.09	2.07	2.05	2.04	2.02	2.01	1.99	1.90	1.85	1.82	1.79	1.77	1.76	1.75	1.74	1.73	1.64
17	2.14	2.11	2.09	2.07	2.05	2.03	2.02	2.00	1.99	1.98	1.89	1.83	1.80	1.77	1.75	1.74	1.73	1.72	1.71	1.62
18	2.12	2.10	2.08	2.05	2.04	2.02	2.00	1.99	1.97	1.96	1.87	1.81	1.78	1.75	1.73	1.72	1.71	1.70	1.69	1.60
19	2.11	2.08	2.06	2.04	2.02	2.00	1.99	1.97	1.96	1.95	1.85	1.80	1.76	1.74	1.72	1.70	1.69	1.68	1.67	1.59
20	2.10	2.07	2.05	2.03	2.01	1.99	1.97	1.96	1.94	1.93	1.84	1.78	1.75	1.72	1.70	1.69	1.68	1.67	1.66	1.57
21	2.08	2.06	2.04	2.01	2.00	1.98	1.96	1.95	1.93	1.92	1.83	1.77	1.73	1.71	1.69	1.67	1.66	1.65	1.64	1.56
22	2.07	2.05	2.02	2.00	1.98	1.97	1.95	1.93	1.92	1.91	1.81	1.76	1.72	1.70	1.68	1.66	1.65	1.64	1.63	1.54
23	2.06	2.04	2.01	1.99	1.97	1.96	1.94	1.92	1.91	1.90	1.80	1.75	1.71	1.68	1.67	1.65	1.64	1.63	1.62	1.53
24	2.05	2.03	2.01	1.98	1.96	1.95	1.93	1.91	1.90	1.89	1.79	1.74	1.70	1.67	1.65	1.64	1.63	1.62	1.61	1.52
25	2.05	2.02	2.00	1.97	1.96	1.94	1.92	1.91	1.89	1.88	1.78	1.73	1.69	1.66	1.64	1.63	1.62	1.61	1.60	1.51
26	2.04	2.01	1.99	1.97	1.95	1.93	1.91	1.90	1.88	1.87	1.77	1.72	1.68	1.65	1.63	1.62	1.61	1.60	1.59	1.50
27	2.03	2.00	1.98	1.96	1.94	1.92	1.90	1.89	1.88	1.86	1.77	1.71	1.67	1.65	1.63	1.61	1.60	1.59	1.58	1.49
28	2.02	2.00	1.97	1.95	1.93	1.91	1.90	1.88	1.87	1.85	1.76	1.70	1.66	1.64	1.62	1.60	1.59	1.58	1.57	1.48
29	2.02	1.99	1.97	1.95	1.93	1.91	1.89	1.88	1.86	1.85	1.75	1.69	1.66	1.63	1.61	1.59	1.58	1.57	1.56	1.47
30	2.01	1.98	1.96	1.94	1.92	1.90	1.88	1.87	1.85	1.84	1.74	1.69	1.65	1.62	1.60	1.59	1.57	1.56	1.55	1.46
40	1.96	1.94	1.91	1.89	1.87	1.85	1.84	1.82	1.81	1.79	1.69	1.63	1.59	1.57	1.54	1.53	1.52	1.50	1.50	1.45
50	1.94	1.91	1.88	1.86	1.84	1.82	1.81	1.79	1.77	1.76	1.66	1.60	1.56	1.53	1.51	1.49	1.48	1.47	1.46	1.44
60	1.92	1.89	1.86	1.84	1.82	1.80	1.79	1.77	1.75	1.74	1.64	1.58	1.53	1.50	1.48	1.46	1.45	1.44	1.43	1.44
70	1.90	1.88	1.85	1.83	1.81	1.79	1.77	1.75	1.74	1.72	1.62	1.56	1.52	1.49	1.46	1.44	1.43	1.42	1.41	1.43
80	1.89	1.86	1.84	1.82	1.80	1.78	1.76	1.74	1.73	1.71	1.61	1.54	1.50	1.47	1.45	1.43	1.41	1.40	1.39	1.42
90	1.88	1.86	1.83	1.81	1.79	1.77	1.75	1.73	1.72	1.70	1.60	1.53	1.49	1.46	1.44	1.42	1.40	1.39	1.38	1.42
100	1.88	1.85	1.82	1.80	1.78	1.76	1.74	1.73	1.71	1.70	1.59	1.52	1.48	1.45	1.43	1.41	1.39	1.38	1.37	1.41
120	1.87	1.84	1.81	1.79	1.77	1.75	1.73	1.71	1.70	1.68	1.58	1.51	1.47	1.44	1.41	1.39	1.38	1.36	1.35	1.40
∞	1.81	1.78	1.76	1.73	1.71	1.69	1.67	1.65	1.64	1.62	1.61	1.60	1.58	1.57	1.56	1.55	1.54	1.53	1.52	1.03

Tabla VI.3: Distribución $F(m, n)$. $\alpha = 0.025$ 1^a fila= g.l. denominador. 1^a columna= g.l. numerador.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	647.79	38.51	17.44	12.22	10.01	8.81	8.07	7.57	7.21	6.94	6.72	6.55	6.41	6.30	6.20	6.12	6.04	5.98	5.92	5.87
2	799.50	39.00	16.04	10.65	8.43	7.26	6.54	6.06	5.71	5.46	5.26	5.10	4.97	4.86	4.77	4.69	4.62	4.56	4.51	4.46
3	864.16	39.17	15.44	9.98	7.76	6.60	5.89	5.42	5.08	4.83	4.63	4.47	4.35	4.24	4.15	4.08	4.01	3.95	3.90	3.86
4	899.59	39.25	15.10	9.60	7.39	6.23	5.52	5.05	4.72	4.47	4.28	4.12	4.00	3.89	3.80	3.73	3.66	3.61	3.56	3.51
5	921.85	39.30	14.88	9.36	7.15	5.99	5.29	4.82	4.48	4.24	4.04	3.89	3.77	3.66	3.58	3.50	3.44	3.38	3.33	3.29
6	937.11	39.33	14.73	9.20	6.98	5.82	5.12	4.65	4.32	4.07	3.88	3.73	3.60	3.50	3.41	3.34	3.28	3.22	3.17	3.13
7	948.22	39.36	14.62	9.07	6.85	5.70	4.99	4.53	4.20	3.95	3.76	3.61	3.48	3.38	3.29	3.22	3.16	3.10	3.05	3.01
8	956.66	39.37	14.54	8.98	6.76	5.60	4.90	4.43	4.10	3.85	3.66	3.51	3.39	3.29	3.20	3.12	3.06	3.01	2.96	2.91
9	963.29	39.39	14.47	8.90	6.68	5.52	4.82	4.36	4.03	3.78	3.59	3.44	3.31	3.21	3.12	3.05	2.98	2.93	2.88	2.84
10	968.62	39.40	14.42	8.84	6.62	5.46	4.76	4.30	3.96	3.72	3.53	3.37	3.25	3.15	3.06	2.99	2.92	2.87	2.82	2.77
11	973.02	39.41	14.37	8.79	6.57	5.41	4.71	4.24	3.91	3.66	3.47	3.32	3.20	3.09	3.01	2.93	2.87	2.81	2.76	2.72
12	976.70	39.41	14.34	8.75	6.52	5.37	4.67	4.20	3.87	3.62	3.43	3.28	3.15	3.05	2.96	2.89	2.82	2.77	2.72	2.68
13	979.83	39.42	14.30	8.71	6.49	5.33	4.63	4.16	3.83	3.58	3.39	3.24	3.12	3.01	2.92	2.85	2.79	2.73	2.68	2.64
14	982.54	39.43	14.28	8.68	6.46	5.30	4.60	4.13	3.80	3.55	3.36	3.21	3.08	2.98	2.89	2.82	2.75	2.70	2.65	2.60
15	984.86	39.43	14.25	8.66	6.43	5.27	4.57	4.10	3.77	3.52	3.33	3.18	3.05	2.95	2.86	2.79	2.72	2.67	2.62	2.57
16	986.93	39.44	14.23	8.63	6.40	5.24	4.54	4.08	3.74	3.50	3.30	3.15	3.03	2.92	2.84	2.76	2.70	2.64	2.59	2.55
17	988.73	39.44	14.21	8.61	6.38	5.22	4.52	4.05	3.72	3.47	3.28	3.13	3.00	2.90	2.81	2.74	2.67	2.62	2.57	2.52
18	990.35	39.44	14.20	8.59	6.36	5.20	4.50	4.03	3.70	3.45	3.26	3.11	2.98	2.88	2.79	2.72	2.65	2.60	2.55	2.50
19	991.80	39.45	14.18	8.58	6.34	5.18	4.48	4.02	3.68	3.44	3.24	3.09	2.96	2.86	2.77	2.70	2.63	2.58	2.53	2.48
20	993.10	39.45	14.17	8.56	6.33	5.17	4.47	4.00	3.67	3.42	3.23	3.07	2.95	2.84	2.76	2.68	2.62	2.56	2.51	2.46
21	994.28	39.45	14.16	8.55	6.31	5.15	4.45	3.98	3.65	3.40	3.21	3.06	2.93	2.83	2.74	2.67	2.60	2.54	2.49	2.45
22	995.36	39.45	14.14	8.53	6.30	5.14	4.44	3.97	3.64	3.39	3.20	3.04	2.92	2.81	2.73	2.65	2.59	2.53	2.48	2.43
23	996.35	39.45	14.13	8.52	6.29	5.13	4.43	3.96	3.63	3.38	3.18	3.03	2.91	2.80	2.71	2.64	2.57	2.52	2.46	2.42
24	997.25	39.46	14.12	8.51	6.28	5.12	4.41	3.95	3.61	3.37	3.17	3.02	2.89	2.79	2.70	2.63	2.56	2.50	2.45	2.41
25	998.08	39.46	14.12	8.50	6.27	5.11	4.40	3.94	3.60	3.35	3.16	3.01	2.88	2.78	2.69	2.61	2.55	2.49	2.44	2.40
26	998.85	39.46	14.11	8.49	6.26	5.10	4.39	3.93	3.59	3.34	3.15	3.00	2.87	2.77	2.68	2.60	2.54	2.48	2.43	2.39
27	999.55	39.46	14.10	8.48	6.25	5.09	4.39	3.92	3.58	3.34	3.14	2.99	2.86	2.76	2.67	2.59	2.53	2.47	2.42	2.38
28	1000.21	39.46	14.09	8.48	6.24	5.08	4.38	3.91	3.58	3.33	3.13	2.98	2.85	2.75	2.66	2.58	2.52	2.46	2.41	2.37
29	1000.84	39.46	14.09	8.47	6.23	5.07	4.37	3.90	3.57	3.32	3.13	2.97	2.85	2.74	2.65	2.58	2.51	2.45	2.40	2.36
30	1001.42	39.46	14.08	8.46	6.23	5.07	4.36	3.89	3.56	3.31	3.12	2.96	2.84	2.73	2.64	2.57	2.50	2.44	2.39	2.35
40	1005.60	39.47	14.04	8.41	6.18	5.01	4.31	3.84	3.51	3.26	3.06	2.91	2.78	2.67	2.59	2.51	2.44	2.38	2.33	2.29
50	1008.11	39.48	14.01	8.38	6.14	4.98	4.28	3.81	3.47	3.22	3.03	2.87	2.74	2.64	2.55	2.47	2.41	2.35	2.30	2.25
60	1009.80	39.48	13.99	8.36	6.12	4.96	4.25	3.78	3.45	3.20	3.00	2.85	2.72	2.61	2.52	2.45	2.38	2.32	2.27	2.22
70	1011.00	39.48	13.98	8.35	6.11	4.94	4.24	3.77	3.43	3.18	2.99	2.83	2.70	2.60	2.51	2.43	2.36	2.30	2.25	2.20
80	1011.90	39.49	13.97	8.33	6.10	4.93	4.23	3.76	3.42	3.17	2.97	2.82	2.69	2.58	2.49	2.42	2.35	2.29	2.24	2.19
90	1012.61	39.49	13.96	8.33	6.09	4.92	4.22	3.75	3.41	3.16	2.96	2.81	2.68	2.57	2.48	2.40	2.34	2.28	2.23	2.18
100	1013.17	39.49	13.96	8.32	6.08	4.92	4.21	3.74	3.40	3.15	2.96	2.80	2.67	2.56	2.47	2.40	2.33	2.27	2.22	2.17
120	1014.03	39.49	13.95	8.31	6.07	4.90	4.20	3.73	3.39	3.14	2.94	2.79	2.66	2.55	2.46	2.38	2.32	2.26	2.20	2.16
∞	1018.21	39.50	13.90	8.26	6.02	4.85	4.14	3.67	3.33	3.08	2.88	2.73	2.60	2.49	2.40	2.32	2.25	2.19	2.13	2.09

Tabla VI.3 (cont.): Distribución $F(m, n)$. $\alpha = 0.025$

1^a fila= g.l. denominador. 1^a columna= g.l. numerador.

	21	22	23	24	25	26	27	28	29	30	40	50	60	70	80	90	100	110	120	∞
1	5.83	5.79	5.75	5.72	5.69	5.66	5.63	5.61	5.59	5.57	5.42	5.34	5.29	5.25	5.22	5.20	5.18	5.16	5.15	5.03
2	4.42	4.38	4.35	4.32	4.29	4.27	4.24	4.22	4.20	4.18	4.05	3.97	3.93	3.89	3.86	3.84	3.83	3.82	3.80	3.69
3	3.82	3.78	3.75	3.72	3.69	3.67	3.65	3.63	3.61	3.59	3.46	3.39	3.34	3.31	3.28	3.26	3.25	3.24	3.23	3.12
4	3.48	3.44	3.41	3.38	3.35	3.33	3.31	3.29	3.27	3.25	3.13	3.05	3.01	2.97	2.95	2.93	2.92	2.90	2.89	2.79
5	3.25	3.22	3.18	3.15	3.13	3.10	3.08	3.06	3.04	3.03	2.90	2.83	2.79	2.75	2.73	2.71	2.70	2.68	2.67	2.57
6	3.09	3.05	3.02	2.99	2.97	2.94	2.92	2.90	2.88	2.87	2.74	2.67	2.63	2.59	2.57	2.55	2.54	2.53	2.52	2.41
7	2.97	2.93	2.90	2.87	2.85	2.82	2.80	2.78	2.76	2.75	2.62	2.55	2.51	2.47	2.45	2.43	2.42	2.40	2.39	2.29
8	2.87	2.84	2.81	2.78	2.75	2.73	2.71	2.69	2.67	2.65	2.53	2.46	2.41	2.38	2.35	2.34	2.32	2.31	2.30	2.19
9	2.80	2.76	2.73	2.70	2.68	2.65	2.63	2.61	2.59	2.57	2.45	2.38	2.33	2.30	2.28	2.26	2.24	2.23	2.22	2.11
10	2.73	2.70	2.67	2.64	2.61	2.59	2.57	2.55	2.53	2.51	2.39	2.32	2.27	2.24	2.21	2.19	2.18	2.17	2.16	2.05
11	2.68	2.65	2.62	2.59	2.56	2.54	2.51	2.49	2.48	2.46	2.33	2.26	2.22	2.18	2.16	2.14	2.12	2.11	2.10	1.99
12	2.64	2.60	2.57	2.54	2.51	2.49	2.47	2.45	2.43	2.41	2.29	2.22	2.17	2.14	2.11	2.09	2.08	2.07	2.05	1.95
13	2.60	2.56	2.53	2.50	2.48	2.45	2.43	2.41	2.39	2.37	2.25	2.18	2.13	2.10	2.07	2.05	2.04	2.02	2.01	1.90
14	2.56	2.53	2.50	2.47	2.44	2.42	2.39	2.37	2.36	2.34	2.21	2.14	2.09	2.06	2.03	2.02	2.00	1.99	1.98	1.87
15	2.53	2.50	2.47	2.44	2.41	2.39	2.36	2.34	2.32	2.31	2.18	2.11	2.06	2.03	2.00	1.98	1.97	1.96	1.94	1.83
16	2.51	2.47	2.44	2.41	2.38	2.36	2.34	2.32	2.30	2.28	2.15	2.08	2.03	2.00	1.97	1.95	1.94	1.93	1.92	1.80
17	2.48	2.45	2.42	2.39	2.36	2.34	2.31	2.29	2.27	2.26	2.13	2.06	2.01	1.97	1.95	1.93	1.91	1.90	1.89	1.78
18	2.46	2.43	2.39	2.36	2.34	2.31	2.29	2.27	2.25	2.23	2.11	2.03	1.98	1.95	1.92	1.91	1.89	1.88	1.87	1.75
19	2.44	2.41	2.37	2.35	2.32	2.29	2.27	2.25	2.23	2.21	2.09	2.01	1.96	1.93	1.90	1.88	1.87	1.86	1.84	1.73
20	2.42	2.39	2.36	2.33	2.30	2.28	2.25	2.23	2.21	2.20	2.07	1.99	1.94	1.91	1.88	1.86	1.85	1.84	1.82	1.71
21	2.41	2.37	2.34	2.31	2.28	2.26	2.24	2.22	2.20	2.18	2.05	1.98	1.93	1.89	1.87	1.85	1.83	1.82	1.81	1.69
22	2.39	2.36	2.33	2.30	2.27	2.24	2.22	2.20	2.18	2.16	2.03	1.96	1.91	1.88	1.85	1.83	1.81	1.80	1.79	1.67
23	2.38	2.34	2.31	2.28	2.26	2.23	2.21	2.19	2.17	2.15	2.02	1.95	1.90	1.86	1.83	1.81	1.80	1.79	1.77	1.66
24	2.37	2.33	2.30	2.27	2.24	2.22	2.19	2.17	2.15	2.14	2.01	1.93	1.88	1.85	1.82	1.80	1.78	1.77	1.76	1.64
25	2.36	2.32	2.29	2.26	2.23	2.21	2.18	2.16	2.14	2.12	1.99	1.92	1.87	1.83	1.81	1.79	1.77	1.76	1.75	1.63
26	2.34	2.31	2.28	2.25	2.22	2.19	2.17	2.15	2.13	2.11	1.98	1.91	1.86	1.82	1.79	1.77	1.76	1.74	1.73	1.61
27	2.33	2.30	2.27	2.24	2.21	2.18	2.16	2.14	2.12	2.10	1.97	1.90	1.85	1.81	1.78	1.76	1.75	1.73	1.72	1.60
28	2.33	2.29	2.26	2.23	2.20	2.17	2.15	2.13	2.11	2.09	1.96	1.89	1.83	1.80	1.77	1.75	1.74	1.72	1.71	1.59
29	2.32	2.28	2.25	2.22	2.19	2.17	2.14	2.12	2.10	2.08	1.95	1.88	1.82	1.79	1.76	1.74	1.72	1.71	1.70	1.58
30	2.31	2.27	2.24	2.21	2.18	2.16	2.13	2.11	2.09	2.07	1.94	1.87	1.82	1.78	1.75	1.73	1.71	1.70	1.69	1.57
40	2.25	2.21	2.18	2.15	2.12	2.09	2.07	2.05	2.03	2.01	1.88	1.80	1.74	1.71	1.68	1.66	1.64	1.63	1.61	1.56
50	2.21	2.17	2.14	2.11	2.08	2.05	2.03	2.01	1.99	1.97	1.83	1.75	1.70	1.66	1.63	1.61	1.59	1.58	1.56	1.55
60	2.18	2.14	2.11	2.08	2.05	2.03	2.00	1.98	1.96	1.94	1.80	1.72	1.67	1.63	1.60	1.58	1.56	1.54	1.53	1.54
70	2.16	2.13	2.09	2.06	2.03	2.01	1.98	1.96	1.94	1.92	1.78	1.70	1.64	1.60	1.57	1.55	1.53	1.52	1.50	1.53
80	2.15	2.11	2.08	2.05	2.02	1.99	1.97	1.94	1.92	1.90	1.76	1.68	1.63	1.59	1.55	1.53	1.51	1.50	1.48	1.52
90	2.14	2.10	2.07	2.03	2.01	1.98	1.95	1.93	1.91	1.89	1.75	1.67	1.61	1.57	1.54	1.52	1.50	1.48	1.47	1.51
100	2.13	2.09	2.06	2.02	2.00	1.97	1.94	1.92	1.90	1.88	1.74	1.66	1.60	1.56	1.53	1.50	1.48	1.47	1.45	1.51
120	2.11	2.08	2.04	2.01	1.98	1.95	1.93	1.91	1.89	1.87	1.72	1.64	1.58	1.54	1.51	1.48	1.46	1.45	1.43	1.49
∞	2.04	2.00	1.97	1.94	1.91	1.88	1.85	1.83	1.81	1.79	1.77	1.75	1.73	1.72	1.70	1.69	1.67	1.66	1.65	1.04

Tabla VI.4: Distribución $F(m, n)$. $\alpha = 0.01$

1ª fila= g.l. denominador. 1ª columna= g.l. numerador.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	4052.12	98.50	34.12	21.20	16.26	13.75	12.25	11.26	10.56	10.04	9.65	9.33	9.07	8.86	8.68	8.53	8.40	8.29	8.18	8.10
2	4999.45	99.00	30.82	18.00	13.27	10.92	9.55	8.65	8.02	7.56	7.21	6.93	6.70	6.51	6.36	6.23	6.11	6.01	5.93	5.85
3	5403.34	99.17	29.46	16.69	12.06	9.78	8.45	7.59	6.99	6.55	6.22	5.95	5.74	5.56	5.42	5.29	5.18	5.09	5.01	4.94
4	5624.56	99.25	28.71	15.98	11.39	9.15	7.85	7.01	6.42	5.99	5.67	5.41	5.21	5.04	4.89	4.77	4.67	4.58	4.50	4.43
5	5763.59	99.30	28.24	15.52	10.97	8.75	7.46	6.63	6.06	5.64	5.32	5.06	4.86	4.69	4.56	4.44	4.34	4.25	4.17	4.10
6	5858.98	99.33	27.91	15.21	10.67	8.47	7.19	6.37	5.80	5.39	5.07	4.82	4.62	4.46	4.32	4.20	4.10	4.01	3.94	3.87
7	5928.39	99.36	27.67	14.98	10.46	8.26	6.99	6.18	5.61	5.20	4.89	4.64	4.44	4.28	4.14	4.03	3.93	3.84	3.77	3.70
8	5981.16	99.37	27.49	14.80	10.29	8.10	6.84	6.03	5.47	5.06	4.74	4.50	4.30	4.14	4.00	3.89	3.79	3.71	3.63	3.56
9	6022.56	99.39	27.35	14.66	10.16	7.98	6.72	5.91	5.35	4.94	4.63	4.39	4.19	4.03	3.89	3.78	3.68	3.60	3.52	3.46
10	6055.85	99.40	27.23	14.55	10.05	7.87	6.62	5.81	5.26	4.85	4.54	4.30	4.10	3.94	3.80	3.69	3.59	3.51	3.43	3.37
11	6083.25	99.41	27.13	14.45	9.96	7.79	6.54	5.73	5.18	4.77	4.46	4.22	4.02	3.86	3.73	3.62	3.52	3.43	3.36	3.29
12	6106.38	99.42	27.05	14.37	9.89	7.72	6.47	5.67	5.11	4.71	4.40	4.16	3.96	3.80	3.67	3.55	3.46	3.37	3.30	3.23
13	6125.87	99.42	26.98	14.31	9.82	7.66	6.41	5.61	5.05	4.65	4.34	4.10	3.91	3.75	3.61	3.50	3.40	3.32	3.24	3.18
14	6142.71	99.43	26.92	14.25	9.77	7.60	6.36	5.56	5.01	4.60	4.29	4.05	3.86	3.70	3.56	3.45	3.35	3.27	3.19	3.13
15	6157.33	99.43	26.87	14.20	9.72	7.56	6.31	5.52	4.96	4.56	4.25	4.01	3.82	3.66	3.52	3.41	3.31	3.23	3.15	3.09
16	6170.11	99.44	26.83	14.15	9.68	7.52	6.28	5.48	4.92	4.52	4.21	3.97	3.78	3.62	3.49	3.37	3.27	3.19	3.12	3.05
17	6181.48	99.44	26.79	14.11	9.64	7.48	6.24	5.44	4.89	4.49	4.18	3.94	3.75	3.59	3.45	3.34	3.24	3.16	3.08	3.02
18	6191.42	99.44	26.75	14.08	9.61	7.45	6.21	5.41	4.86	4.46	4.15	3.91	3.72	3.56	3.42	3.31	3.21	3.13	3.05	2.99
19	6200.56	99.45	26.72	14.05	9.58	7.42	6.18	5.38	4.83	4.43	4.12	3.88	3.69	3.53	3.40	3.28	3.19	3.10	3.03	2.96
20	6208.67	99.45	26.69	14.02	9.55	7.40	6.16	5.36	4.81	4.41	4.10	3.86	3.66	3.51	3.37	3.26	3.16	3.08	3.00	2.94
21	6216.18	99.45	26.66	13.99	9.53	7.37	6.13	5.34	4.79	4.38	4.08	3.84	3.64	3.48	3.35	3.24	3.14	3.05	2.98	2.92
22	6222.88	99.45	26.64	13.97	9.51	7.35	6.11	5.32	4.77	4.36	4.06	3.82	3.62	3.46	3.33	3.22	3.12	3.03	2.96	2.90
23	6228.97	99.45	26.62	13.95	9.49	7.33	6.09	5.30	4.75	4.34	4.04	3.80	3.60	3.44	3.31	3.20	3.10	3.02	2.94	2.88
24	6234.65	99.46	26.60	13.93	9.47	7.31	6.07	5.28	4.73	4.33	4.02	3.78	3.59	3.43	3.29	3.18	3.08	3.00	2.92	2.86
25	6239.73	99.46	26.58	13.91	9.45	7.30	6.06	5.26	4.71	4.31	4.01	3.76	3.57	3.41	3.28	3.16	3.07	2.98	2.91	2.84
26	6244.60	99.46	26.56	13.89	9.43	7.28	6.04	5.25	4.70	4.30	3.99	3.75	3.56	3.40	3.26	3.15	3.05	2.97	2.89	2.83
27	6249.06	99.46	26.55	13.88	9.42	7.27	6.03	5.23	4.68	4.28	3.98	3.74	3.54	3.38	3.25	3.14	3.04	2.95	2.88	2.81
28	6253.12	99.46	26.53	13.86	9.40	7.25	6.02	5.22	4.67	4.27	3.96	3.72	3.53	3.37	3.24	3.12	3.03	2.94	2.87	2.80
29	6256.98	99.46	26.52	13.85	9.39	7.24	6.00	5.21	4.66	4.26	3.95	3.71	3.52	3.36	3.23	3.11	3.01	2.93	2.86	2.79
30	6260.63	99.47	26.50	13.84	9.38	7.23	5.99	5.20	4.65	4.25	3.94	3.70	3.51	3.35	3.21	3.10	3.00	2.92	2.84	2.78
40	6286.81	99.47	26.41	13.75	9.29	7.14	5.91	5.12	4.57	4.17	3.86	3.62	3.43	3.27	3.13	3.02	2.92	2.84	2.76	2.69
50	6302.44	99.48	26.35	13.69	9.24	7.09	5.86	5.07	4.52	4.12	3.81	3.57	3.38	3.22	3.08	2.97	2.87	2.78	2.71	2.64
60	6313.00	99.48	26.32	13.65	9.20	7.06	5.82	5.03	4.48	4.08	3.78	3.54	3.34	3.18	3.05	2.93	2.83	2.75	2.67	2.61
70	6320.50	99.48	26.29	13.63	9.18	7.03	5.80	5.01	4.46	4.06	3.75	3.51	3.32	3.16	3.02	2.91	2.81	2.72	2.65	2.58
80	6326.19	99.49	26.27	13.61	9.16	7.01	5.78	4.99	4.44	4.04	3.73	3.49	3.30	3.14	3.00	2.89	2.79	2.70	2.63	2.56
90	6330.65	99.49	26.25	13.59	9.14	7.00	5.77	4.97	4.43	4.03	3.72	3.48	3.28	3.12	2.99	2.87	2.78	2.69	2.61	2.55
100	6334.10	99.49	26.24	13.58	9.13	6.99	5.75	4.96	4.41	4.01	3.71	3.47	3.27	3.11	2.98	2.86	2.76	2.68	2.60	2.54
120	6339.38	99.49	26.22	13.56	9.11	6.97	5.74	4.95	4.40	4.00	3.69	3.45	3.25	3.09	2.96	2.84	2.75	2.66	2.58	2.52
∞	6365.56	99.50	26.13	13.46	9.02	6.88	5.65	4.86	4.31	3.91	3.60	3.36	3.17	3.01	2.87	2.75	2.65	2.57	2.49	2.42

Tabla VI.4 (cont.): Distribución $F(m, n)$. $\alpha = 0.01$

1^a fila= g.l. denominador. 1^a columna= g.l. numerador.

	21	22	23	24	25	26	27	28	29	30	40	50	60	70	80	90	100	110	120	∞
1	8.02	7.95	7.88	7.82	7.77	7.72	7.68	7.64	7.60	7.56	7.31	7.17	7.08	7.01	6.96	6.93	6.90	6.87	6.85	6.64
2	5.78	5.72	5.66	5.61	5.57	5.53	5.49	5.45	5.42	5.39	5.18	5.06	4.98	4.92	4.88	4.85	4.82	4.80	4.79	4.61
3	4.87	4.82	4.76	4.72	4.68	4.64	4.60	4.57	4.54	4.51	4.31	4.20	4.13	4.07	4.04	4.01	3.98	3.96	3.95	3.78
4	4.37	4.31	4.26	4.22	4.18	4.14	4.11	4.07	4.04	4.02	3.83	3.72	3.65	3.60	3.56	3.53	3.51	3.49	3.48	3.32
5	4.04	3.99	3.94	3.90	3.85	3.82	3.78	3.75	3.73	3.70	3.51	3.41	3.34	3.29	3.26	3.23	3.21	3.19	3.17	3.02
6	3.81	3.76	3.71	3.67	3.63	3.59	3.56	3.53	3.50	3.47	3.29	3.19	3.12	3.07	3.04	3.01	2.99	2.97	2.96	2.80
7	3.64	3.59	3.54	3.50	3.46	3.42	3.39	3.36	3.33	3.30	3.12	3.02	2.95	2.91	2.87	2.84	2.82	2.81	2.79	2.64
8	3.51	3.45	3.41	3.36	3.32	3.29	3.26	3.23	3.20	3.17	2.99	2.89	2.82	2.78	2.74	2.72	2.69	2.68	2.66	2.51
9	3.40	3.35	3.30	3.26	3.22	3.18	3.15	3.12	3.09	3.07	2.89	2.78	2.72	2.67	2.64	2.61	2.59	2.57	2.56	2.41
10	3.31	3.26	3.21	3.17	3.13	3.09	3.06	3.03	3.00	2.98	2.80	2.70	2.63	2.59	2.55	2.52	2.50	2.49	2.47	2.32
11	3.24	3.18	3.14	3.09	3.06	3.02	2.99	2.96	2.93	2.91	2.73	2.63	2.56	2.51	2.48	2.45	2.43	2.41	2.40	2.25
12	3.17	3.12	3.07	3.03	2.99	2.96	2.93	2.90	2.87	2.84	2.66	2.56	2.50	2.45	2.42	2.39	2.37	2.35	2.34	2.19
13	3.12	3.07	3.02	2.98	2.94	2.90	2.87	2.84	2.81	2.79	2.61	2.51	2.44	2.40	2.36	2.33	2.31	2.30	2.28	2.13
14	3.07	3.02	2.97	2.93	2.89	2.86	2.82	2.79	2.77	2.74	2.56	2.46	2.39	2.35	2.31	2.29	2.27	2.25	2.23	2.08
15	3.03	2.98	2.93	2.89	2.85	2.81	2.78	2.75	2.73	2.70	2.52	2.42	2.35	2.31	2.27	2.24	2.22	2.21	2.19	2.04
16	2.99	2.94	2.89	2.85	2.81	2.78	2.75	2.72	2.69	2.66	2.48	2.38	2.31	2.27	2.23	2.21	2.19	2.17	2.15	2.00
17	2.96	2.91	2.86	2.82	2.78	2.75	2.71	2.68	2.66	2.63	2.45	2.35	2.28	2.23	2.20	2.17	2.15	2.13	2.12	1.97
18	2.93	2.88	2.83	2.79	2.75	2.72	2.68	2.65	2.63	2.60	2.42	2.32	2.25	2.20	2.17	2.14	2.12	2.10	2.09	1.94
19	2.90	2.85	2.80	2.76	2.72	2.69	2.66	2.63	2.60	2.57	2.39	2.29	2.22	2.18	2.14	2.11	2.09	2.07	2.06	1.91
20	2.88	2.83	2.78	2.74	2.70	2.66	2.63	2.60	2.57	2.55	2.37	2.27	2.20	2.15	2.12	2.09	2.07	2.05	2.03	1.88
21	2.86	2.81	2.76	2.72	2.68	2.64	2.61	2.58	2.55	2.53	2.35	2.24	2.17	2.13	2.09	2.06	2.04	2.03	2.01	1.86
22	2.84	2.78	2.74	2.70	2.66	2.62	2.59	2.56	2.53	2.51	2.33	2.22	2.15	2.11	2.07	2.04	2.02	2.00	1.99	1.83
23	2.82	2.77	2.72	2.68	2.64	2.60	2.57	2.54	2.51	2.49	2.31	2.20	2.13	2.09	2.05	2.02	2.00	1.98	1.97	1.81
24	2.80	2.75	2.70	2.66	2.62	2.58	2.55	2.52	2.49	2.47	2.29	2.18	2.12	2.07	2.03	2.00	1.98	1.96	1.95	1.79
25	2.79	2.73	2.69	2.64	2.60	2.57	2.54	2.51	2.48	2.45	2.27	2.17	2.10	2.05	2.01	1.99	1.97	1.95	1.93	1.77
26	2.77	2.72	2.67	2.63	2.59	2.55	2.52	2.49	2.46	2.44	2.26	2.15	2.08	2.03	2.00	1.97	1.95	1.93	1.92	1.76
27	2.76	2.70	2.66	2.61	2.58	2.54	2.51	2.48	2.45	2.42	2.24	2.14	2.07	2.02	1.98	1.96	1.93	1.92	1.90	1.74
28	2.74	2.69	2.64	2.60	2.56	2.53	2.49	2.46	2.44	2.41	2.23	2.12	2.05	2.01	1.97	1.94	1.92	1.90	1.89	1.73
29	2.73	2.68	2.63	2.59	2.55	2.51	2.48	2.45	2.42	2.40	2.22	2.11	2.04	1.99	1.96	1.93	1.91	1.89	1.87	1.71
30	2.72	2.67	2.62	2.58	2.54	2.50	2.47	2.44	2.41	2.39	2.20	2.10	2.03	1.98	1.94	1.92	1.89	1.88	1.86	1.70
40	2.64	2.58	2.54	2.49	2.45	2.42	2.38	2.35	2.33	2.30	2.11	2.01	1.94	1.89	1.85	1.82	1.80	1.78	1.76	1.69
50	2.58	2.53	2.48	2.44	2.40	2.36	2.33	2.30	2.27	2.25	2.06	1.95	1.88	1.83	1.79	1.76	1.74	1.72	1.70	1.67
60	2.55	2.50	2.45	2.40	2.36	2.33	2.29	2.26	2.23	2.21	2.02	1.91	1.84	1.78	1.75	1.72	1.69	1.67	1.66	1.66
70	2.52	2.47	2.42	2.38	2.34	2.30	2.27	2.24	2.21	2.18	1.99	1.88	1.81	1.75	1.71	1.68	1.66	1.64	1.62	1.65
80	2.50	2.45	2.40	2.36	2.32	2.28	2.25	2.22	2.19	2.16	1.97	1.86	1.78	1.73	1.69	1.66	1.63	1.61	1.60	1.64
90	2.49	2.43	2.39	2.34	2.30	2.26	2.23	2.20	2.17	2.14	1.95	1.84	1.76	1.71	1.67	1.64	1.61	1.59	1.58	1.63
100	2.48	2.42	2.37	2.33	2.29	2.25	2.22	2.19	2.16	2.13	1.94	1.82	1.75	1.70	1.65	1.62	1.60	1.58	1.56	1.62
120	2.46	2.40	2.35	2.31	2.27	2.23	2.20	2.17	2.14	2.11	1.92	1.80	1.73	1.67	1.63	1.60	1.57	1.55	1.53	1.60
∞	2.36	2.31	2.26	2.21	2.17	2.13	2.10	2.07	2.04	2.01	1.98	1.96	1.93	1.91	1.89	1.87	1.86	1.84	1.82	1.05

Tabla VI.5: Distribución $F(m, n)$. $\alpha = 0.005$

1^a fila= g.l. denominador. 1^a columna= g.l. numerador.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	16210.78	198.50	55.55	31.33	22.78	18.63	16.24	14.69	13.61	12.83	12.23	11.75	11.37	11.06	10.80	10.58	10.38	10.22	10.07	9.94
2	19998.73	199.00	49.80	26.28	18.31	14.54	12.40	11.04	10.11	9.43	8.91	8.51	8.19	7.92	7.70	7.51	7.35	7.21	7.09	6.99
3	21615.29	199.17	47.47	24.26	16.53	12.92	10.88	9.60	8.72	8.08	7.60	7.23	6.93	6.68	6.48	6.30	6.16	6.03	5.92	5.82
4	22499.60	199.25	46.19	23.15	15.56	12.03	10.05	8.81	7.96	7.34	6.88	6.52	6.23	6.00	5.80	5.64	5.50	5.37	5.27	5.17
5	23056.49	199.30	45.39	22.46	14.94	11.46	9.52	8.30	7.47	6.87	6.42	6.07	5.79	5.56	5.37	5.21	5.07	4.96	4.85	4.76
6	23436.65	199.33	44.84	21.97	14.51	11.07	9.16	7.95	7.13	6.54	6.10	5.76	5.48	5.26	5.07	4.91	4.78	4.66	4.56	4.47
7	23714.07	199.36	44.43	21.62	14.20	10.79	8.89	7.69	6.88	6.30	5.86	5.52	5.25	5.03	4.85	4.69	4.56	4.44	4.34	4.26
8	23925.73	199.37	44.13	21.35	13.96	10.57	8.68	7.50	6.69	6.12	5.68	5.35	5.08	4.86	4.67	4.52	4.39	4.28	4.18	4.09
9	24090.12	199.39	43.88	21.14	13.77	10.39	8.51	7.34	6.54	5.97	5.54	5.20	4.94	4.72	4.54	4.38	4.25	4.14	4.04	3.96
10	24223.70	199.40	43.69	20.97	13.62	10.25	8.38	7.21	6.42	5.85	5.42	5.09	4.82	4.60	4.42	4.27	4.14	4.03	3.93	3.85
11	24334.66	199.41	43.52	20.82	13.49	10.13	8.27	7.10	6.31	5.75	5.32	4.99	4.72	4.51	4.33	4.18	4.05	3.94	3.84	3.76
12	24427.14	199.42	43.39	20.70	13.38	10.03	8.18	7.01	6.23	5.66	5.24	4.91	4.64	4.43	4.25	4.10	3.97	3.86	3.76	3.68
13	24505.22	199.43	43.27	20.60	13.29	9.95	8.10	6.94	6.15	5.59	5.16	4.84	4.57	4.36	4.18	4.03	3.90	3.79	3.70	3.61
14	24570.98	199.43	43.17	20.51	13.21	9.88	8.03	6.87	6.09	5.53	5.10	4.77	4.51	4.30	4.12	3.97	3.84	3.73	3.64	3.55
15	24630.57	199.43	43.08	20.44	13.15	9.81	7.97	6.81	6.03	5.47	5.05	4.72	4.46	4.25	4.07	3.92	3.79	3.68	3.59	3.50
16	24681.95	199.43	43.01	20.37	13.09	9.76	7.91	6.76	5.98	5.42	5.00	4.67	4.41	4.20	4.02	3.87	3.75	3.64	3.54	3.46
17	24727.16	199.44	42.94	20.31	13.03	9.71	7.87	6.72	5.94	5.38	4.96	4.63	4.37	4.16	3.98	3.83	3.71	3.60	3.50	3.42
18	24766.20	199.44	42.88	20.26	12.99	9.66	7.83	6.68	5.90	5.34	4.92	4.59	4.33	4.12	3.95	3.80	3.67	3.56	3.46	3.38
19	24803.19	199.45	42.83	20.21	12.94	9.62	7.79	6.64	5.86	5.31	4.89	4.56	4.30	4.09	3.91	3.76	3.64	3.53	3.43	3.35
20	24836.07	199.45	42.78	20.17	12.90	9.59	7.75	6.61	5.83	5.27	4.86	4.53	4.27	4.06	3.88	3.73	3.61	3.50	3.40	3.32
21	24864.84	199.45	42.73	20.13	12.87	9.56	7.72	6.58	5.80	5.25	4.83	4.50	4.24	4.03	3.86	3.71	3.58	3.47	3.37	3.29
22	24891.55	199.45	42.69	20.09	12.84	9.53	7.69	6.55	5.78	5.22	4.80	4.48	4.22	4.01	3.83	3.68	3.56	3.45	3.35	3.27
23	24916.21	199.46	42.66	20.06	12.81	9.50	7.67	6.53	5.75	5.20	4.78	4.45	4.19	3.98	3.81	3.66	3.53	3.42	3.33	3.24
24	24938.82	199.46	42.62	20.03	12.78	9.47	7.64	6.50	5.73	5.17	4.76	4.43	4.17	3.96	3.79	3.64	3.51	3.40	3.31	3.22
25	24959.37	199.46	42.59	20.00	12.76	9.45	7.62	6.48	5.71	5.15	4.74	4.41	4.15	3.94	3.77	3.62	3.49	3.38	3.29	3.20
26	24979.92	199.46	42.56	19.98	12.73	9.43	7.60	6.46	5.69	5.13	4.72	4.39	4.13	3.92	3.75	3.60	3.47	3.36	3.27	3.18
27	24996.36	199.47	42.54	19.95	12.71	9.41	7.58	6.44	5.67	5.12	4.70	4.38	4.12	3.91	3.73	3.58	3.46	3.35	3.25	3.17
28	25013.82	199.47	42.51	19.93	12.69	9.39	7.57	6.43	5.65	5.10	4.68	4.36	4.10	3.89	3.72	3.57	3.44	3.33	3.24	3.15
29	25029.23	199.47	42.49	19.91	12.67	9.37	7.55	6.41	5.64	5.08	4.67	4.34	4.09	3.88	3.70	3.55	3.43	3.32	3.22	3.14
30	25043.62	199.47	42.47	19.89	12.66	9.36	7.53	6.40	5.62	5.07	4.65	4.33	4.07	3.86	3.69	3.54	3.41	3.30	3.21	3.12
40	25148.42	199.47	42.31	19.75	12.53	9.24	7.42	6.29	5.52	4.97	4.55	4.23	3.97	3.76	3.58	3.44	3.31	3.20	3.11	3.02
50	25211.10	199.48	42.21	19.67	12.45	9.17	7.35	6.22	5.45	4.90	4.49	4.17	3.91	3.70	3.52	3.37	3.25	3.14	3.04	2.96
60	25253.22	199.48	42.15	19.61	12.40	9.12	7.31	6.18	5.41	4.86	4.44	4.12	3.87	3.66	3.48	3.33	3.21	3.10	3.00	2.92
70	25284.05	199.48	42.10	19.57	12.37	9.09	7.28	6.15	5.38	4.83	4.41	4.09	3.84	3.62	3.45	3.30	3.18	3.07	2.97	2.88
80	25306.65	199.49	42.07	19.54	12.34	9.06	7.25	6.12	5.36	4.80	4.39	4.07	3.81	3.60	3.43	3.28	3.15	3.04	2.95	2.86
90	25323.09	199.49	42.04	19.52	12.32	9.04	7.23	6.10	5.34	4.79	4.37	4.05	3.79	3.58	3.41	3.26	3.13	3.02	2.93	2.84
100	25337.48	199.49	42.02	19.50	12.30	9.03	7.22	6.09	5.32	4.77	4.36	4.04	3.78	3.57	3.39	3.25	3.12	3.01	2.91	2.83
120	25358.03	199.49	41.99	19.47	12.27	9.00	7.19	6.06	5.30	4.75	4.34	4.01	3.76	3.55	3.37	3.22	3.10	2.99	2.89	2.81
∞	25462.83	199.50	41.83	19.33	12.15	8.88	7.08	5.95	5.19	4.64	4.23	3.91	3.65	3.44	3.26	3.11	2.99	2.87	2.78	2.69

Tabla VI.5 (cont.): Distribución $F(m, n)$. $\alpha = 0.005$

1^a fila= g.l. denominador. 1^a columna= g.l. numerador.

	21	22	23	24	25	26	27	28	29	30	40	50	60	70	80	90	100	110	120	∞
1	9.83	9.73	9.63	9.55	9.48	9.41	9.34	9.28	9.23	9.18	8.83	8.63	8.49	8.40	8.33	8.28	8.24	8.21	8.18	7.88
2	6.89	6.81	6.73	6.66	6.60	6.54	6.49	6.44	6.40	6.35	6.07	5.90	5.80	5.72	5.67	5.62	5.59	5.56	5.54	5.30
3	5.73	5.65	5.58	5.52	5.46	5.41	5.36	5.32	5.28	5.24	4.98	4.83	4.73	4.66	4.61	4.57	4.54	4.52	4.50	4.28
4	5.09	5.02	4.95	4.89	4.84	4.79	4.74	4.70	4.66	4.62	4.37	4.23	4.14	4.08	4.03	3.99	3.96	3.94	3.92	3.72
5	4.68	4.61	4.54	4.49	4.43	4.38	4.34	4.30	4.26	4.23	3.99	3.85	3.76	3.70	3.65	3.62	3.59	3.57	3.55	3.35
6	4.39	4.32	4.26	4.20	4.15	4.10	4.06	4.02	3.98	3.95	3.71	3.58	3.49	3.43	3.39	3.35	3.33	3.30	3.28	3.09
7	4.18	4.11	4.05	3.99	3.94	3.89	3.85	3.81	3.77	3.74	3.51	3.38	3.29	3.23	3.19	3.15	3.13	3.11	3.09	2.90
8	4.01	3.94	3.88	3.83	3.78	3.73	3.69	3.65	3.61	3.58	3.35	3.22	3.13	3.08	3.03	3.00	2.97	2.95	2.93	2.75
9	3.88	3.81	3.75	3.69	3.64	3.60	3.56	3.52	3.48	3.45	3.22	3.09	3.01	2.95	2.91	2.87	2.85	2.83	2.81	2.62
10	3.77	3.70	3.64	3.59	3.54	3.49	3.45	3.41	3.38	3.34	3.12	2.99	2.90	2.85	2.80	2.77	2.74	2.72	2.71	2.52
11	3.68	3.61	3.55	3.50	3.45	3.40	3.36	3.32	3.29	3.25	3.03	2.90	2.82	2.76	2.72	2.68	2.66	2.64	2.62	2.43
12	3.60	3.54	3.47	3.42	3.37	3.33	3.28	3.25	3.21	3.18	2.95	2.82	2.74	2.68	2.64	2.61	2.58	2.56	2.54	2.36
13	3.54	3.47	3.41	3.35	3.30	3.26	3.22	3.18	3.15	3.11	2.89	2.76	2.68	2.62	2.58	2.54	2.52	2.50	2.48	2.30
14	3.48	3.41	3.35	3.30	3.25	3.20	3.16	3.12	3.09	3.06	2.83	2.70	2.62	2.56	2.52	2.49	2.46	2.44	2.42	2.24
15	3.43	3.36	3.30	3.25	3.20	3.15	3.11	3.07	3.04	3.01	2.78	2.65	2.57	2.51	2.47	2.44	2.41	2.39	2.37	2.19
16	3.38	3.31	3.25	3.20	3.15	3.11	3.07	3.03	2.99	2.96	2.74	2.61	2.53	2.47	2.43	2.39	2.37	2.35	2.33	2.14
17	3.34	3.27	3.21	3.16	3.11	3.07	3.03	2.99	2.95	2.92	2.70	2.57	2.49	2.43	2.39	2.35	2.33	2.31	2.29	2.10
18	3.31	3.24	3.18	3.12	3.08	3.03	2.99	2.95	2.92	2.89	2.66	2.53	2.45	2.39	2.35	2.32	2.29	2.27	2.25	2.07
19	3.27	3.21	3.15	3.09	3.04	3.00	2.96	2.92	2.88	2.85	2.63	2.50	2.42	2.36	2.32	2.28	2.26	2.24	2.22	2.03
20	3.24	3.18	3.12	3.06	3.01	2.97	2.93	2.89	2.86	2.82	2.60	2.47	2.39	2.33	2.29	2.25	2.23	2.21	2.19	2.00
21	3.22	3.15	3.09	3.04	2.99	2.94	2.90	2.86	2.83	2.80	2.57	2.44	2.36	2.30	2.26	2.23	2.20	2.18	2.16	1.97
22	3.19	3.12	3.06	3.01	2.96	2.92	2.88	2.84	2.80	2.77	2.55	2.42	2.33	2.28	2.23	2.20	2.17	2.15	2.13	1.95
23	3.17	3.10	3.04	2.99	2.94	2.89	2.85	2.82	2.78	2.75	2.52	2.39	2.31	2.25	2.21	2.18	2.15	2.13	2.11	1.92
24	3.15	3.08	3.02	2.97	2.92	2.87	2.83	2.79	2.76	2.73	2.50	2.37	2.29	2.23	2.19	2.15	2.13	2.11	2.09	1.90
25	3.13	3.06	3.00	2.95	2.90	2.85	2.81	2.77	2.74	2.71	2.48	2.35	2.27	2.21	2.17	2.13	2.11	2.09	2.07	1.88
26	3.11	3.04	2.98	2.93	2.88	2.84	2.79	2.76	2.72	2.69	2.46	2.33	2.25	2.19	2.15	2.12	2.09	2.07	2.05	1.86
27	3.09	3.03	2.97	2.91	2.86	2.82	2.78	2.74	2.70	2.67	2.45	2.32	2.23	2.17	2.13	2.10	2.07	2.05	2.03	1.84
28	3.08	3.01	2.95	2.90	2.85	2.80	2.76	2.72	2.69	2.66	2.43	2.30	2.22	2.16	2.11	2.08	2.05	2.03	2.01	1.82
29	3.06	3.00	2.94	2.88	2.83	2.79	2.75	2.71	2.67	2.64	2.42	2.29	2.20	2.14	2.10	2.07	2.04	2.02	2.00	1.81
30	3.05	2.98	2.92	2.87	2.82	2.77	2.73	2.69	2.66	2.63	2.40	2.27	2.19	2.13	2.08	2.05	2.02	2.00	1.98	1.79
40	2.95	2.88	2.82	2.77	2.72	2.67	2.63	2.59	2.56	2.52	2.30	2.16	2.08	2.02	1.97	1.94	1.91	1.89	1.87	1.78
50	2.88	2.82	2.76	2.70	2.65	2.61	2.57	2.53	2.49	2.46	2.23	2.10	2.01	1.95	1.90	1.87	1.84	1.82	1.80	1.76
60	2.84	2.77	2.71	2.66	2.61	2.56	2.52	2.48	2.45	2.42	2.18	2.05	1.96	1.90	1.85	1.82	1.79	1.77	1.75	1.75
70	2.81	2.74	2.68	2.63	2.58	2.53	2.49	2.45	2.42	2.38	2.15	2.02	1.93	1.86	1.82	1.78	1.75	1.73	1.71	1.74
80	2.79	2.72	2.66	2.60	2.55	2.51	2.47	2.43	2.39	2.36	2.12	1.99	1.90	1.84	1.79	1.75	1.72	1.70	1.68	1.72
90	2.77	2.70	2.64	2.58	2.53	2.49	2.45	2.41	2.37	2.34	2.10	1.97	1.88	1.81	1.77	1.73	1.70	1.68	1.66	1.71
100	2.75	2.69	2.62	2.57	2.52	2.47	2.43	2.39	2.36	2.32	2.09	1.95	1.86	1.80	1.75	1.71	1.68	1.66	1.64	1.70
120	2.73	2.66	2.60	2.55	2.50	2.45	2.41	2.37	2.33	2.30	2.06	1.93	1.83	1.77	1.72	1.68	1.65	1.63	1.61	1.68
∞	2.62	2.55	2.49	2.43	2.38	2.33	2.29	2.25	2.21	2.18	2.15	2.12	2.09	2.06	2.04	2.01	1.99	1.97	1.95	1.05

Tabla VI.6: Distribución $F(m, n)$. $\alpha = 0.001$

1ª fila= g.l. denominador. 1ª columna= g.l. numerador.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	405317.5	998.41	167.04	74.14	47.18	35.51	29.25	25.41	22.86	21.04	19.69	18.64	17.82	17.14	16.59	16.12	15.72	15.38	15.08	14.82
2	499975.9	999.05	148.5	61.25	37.12	27	21.69	18.49	16.39	14.91	13.81	12.97	12.31	11.78	11.34	10.97	10.66	10.39	10.16	9.95
3	540423.4	999.21	141.11	56.18	33.2	23.7	18.77	15.83	13.9	12.55	11.56	10.8	10.21	9.73	9.34	9.01	8.73	8.49	8.28	8.1
4	562472.9	999.21	137.1	53.44	31.09	21.92	17.2	14.39	12.56	11.28	10.35	9.63	9.07	8.62	8.25	7.94	7.68	7.46	7.27	7.1
5	576376.7	999.37	134.58	51.71	29.75	20.8	16.21	13.48	11.71	10.48	9.58	8.89	8.35	7.92	7.57	7.27	7.02	6.81	6.62	6.46
6	585856.6	999.37	132.85	50.53	28.83	20.03	15.52	12.86	11.13	9.93	9.05	8.38	7.86	7.44	7.09	6.8	6.56	6.35	6.18	6.02
7	592808.5	999.37	131.58	49.66	28.16	19.46	15.02	12.4	10.7	9.52	8.66	8	7.49	7.08	6.74	6.46	6.22	6.02	5.85	5.69
8	598075.1	999.37	130.61	49	27.65	19.03	14.63	12.05	10.37	9.2	8.35	7.71	7.21	6.8	6.47	6.2	5.96	5.76	5.59	5.44
9	602288.4	999.37	129.85	48.47	27.25	18.69	14.33	11.77	10.11	8.96	8.12	7.48	6.98	6.58	6.26	5.98	5.75	5.56	5.39	5.24
10	605659	999.37	129.24	48.05	26.92	18.41	14.08	11.54	9.89	8.75	7.92	7.29	6.8	6.4	6.08	5.81	5.58	5.39	5.22	5.08
11	608397.6	999.37	128.73	47.7	26.65	18.18	13.88	11.35	9.72	8.59	7.76	7.14	6.65	6.26	5.94	5.67	5.44	5.25	5.08	4.94
12	610715	999.37	128.31	47.41	26.42	17.99	13.71	11.19	9.57	8.45	7.63	7	6.52	6.13	5.81	5.55	5.32	5.13	4.97	4.82
13	612610.9	999.37	127.95	47.16	26.22	17.82	13.56	11.06	9.44	8.32	7.51	6.89	6.41	6.02	5.71	5.44	5.22	5.03	4.87	4.72
14	614296.2	999.37	127.65	46.95	26.06	17.68	13.43	10.94	9.33	8.22	7.41	6.79	6.31	5.93	5.62	5.35	5.13	4.94	4.78	4.64
15	615770.9	999.37	127.38	46.76	25.91	17.56	13.32	10.84	9.24	8.13	7.32	6.71	6.23	5.85	5.54	5.27	5.05	4.87	4.7	4.56
16	617034.9	999.37	127.13	46.6	25.78	17.45	13.23	10.75	9.15	8.05	7.24	6.63	6.16	5.78	5.46	5.2	4.99	4.8	4.64	4.49
17	618088.2	999.37	126.93	46.45	25.67	17.35	13.14	10.67	9.08	7.98	7.17	6.57	6.09	5.71	5.4	5.14	4.92	4.74	4.58	4.44
18	619141.5	999.37	126.74	46.32	25.57	17.27	13.06	10.6	9.01	7.91	7.11	6.51	6.03	5.66	5.35	5.09	4.87	4.68	4.52	4.38
19	620194.8	999.37	126.57	46.21	25.48	17.19	12.99	10.54	8.95	7.86	7.06	6.45	5.98	5.6	5.29	5.04	4.82	4.63	4.47	4.33
20	620826.8	999.37	126.41	46.1	25.4	17.12	12.93	10.48	8.9	7.8	7.01	6.4	5.93	5.56	5.25	4.99	4.78	4.59	4.43	4.29
21	621669.5	999.37	126.28	46.01	25.32	17.06	12.87	10.43	8.85	7.76	6.96	6.36	5.89	5.51	5.21	4.95	4.73	4.55	4.39	4.25
22	622301.5	999.37	126.16	45.92	25.25	17	12.82	10.38	8.8	7.71	6.92	6.32	5.85	5.48	5.17	4.91	4.7	4.51	4.35	4.21
23	622933.5	999.37	126.04	45.84	25.19	16.95	12.78	10.34	8.76	7.67	6.88	6.28	5.81	5.44	5.13	4.88	4.66	4.48	4.32	4.18
24	623565.5	999.37	125.93	45.76	25.13	16.9	12.73	10.3	8.72	7.64	6.85	6.25	5.78	5.41	5.1	4.85	4.63	4.45	4.29	4.15
25	623986.8	999.37	125.84	45.7	25.08	16.85	12.69	10.26	8.69	7.6	6.81	6.22	5.75	5.38	5.07	4.82	4.6	4.42	4.26	4.12
26	624408.1	999.37	125.76	45.64	25.03	16.81	12.66	10.22	8.66	7.57	6.78	6.19	5.72	5.35	5.04	4.79	4.57	4.39	4.23	4.09
27	624829.4	999.37	125.67	45.58	24.99	16.77	12.62	10.19	8.63	7.54	6.76	6.16	5.7	5.32	5.02	4.76	4.55	4.37	4.21	4.07
28	625250.8	999.37	125.59	45.53	24.94	16.74	12.59	10.16	8.6	7.52	6.73	6.14	5.67	5.3	4.99	4.74	4.53	4.34	4.18	4.05
29	625672.1	999.37	125.52	45.48	24.91	16.7	12.56	10.13	8.57	7.49	6.71	6.11	5.65	5.28	4.97	4.72	4.5	4.32	4.16	4.03
30	626093.4	999.37	125.45	45.43	24.87	16.67	12.53	10.11	8.55	7.47	6.68	6.09	5.63	5.25	4.95	4.7	4.48	4.3	4.14	4
40	628621.4	999.53	124.95	45.09	24.6	16.44	12.33	9.92	8.37	7.3	6.52	5.93	5.47	5.1	4.8	4.54	4.33	4.15	3.99	3.86
50	630306.7	999.53	124.66	44.88	24.44	16.31	12.2	9.8	8.26	7.19	6.42	5.83	5.37	5	4.7	4.45	4.24	4.06	3.9	3.77
60	631360	999.53	124.47	44.74	24.33	16.21	12.12	9.73	8.19	7.12	6.35	5.76	5.3	4.94	4.64	4.39	4.18	4	3.84	3.7
70	631992	999.53	124.32	44.65	24.25	16.15	12.06	9.67	8.13	7.07	6.3	5.71	5.26	4.89	4.59	4.34	4.13	3.95	3.79	3.66
80	632624	999.53	124.21	44.58	24.2	16.1	12.01	9.63	8.09	7.03	6.26	5.68	5.22	4.86	4.56	4.31	4.1	3.92	3.76	3.62
90	633045.3	999.53	124.13	44.52	24.15	16.06	11.98	9.6	8.06	7	6.23	5.65	5.19	4.83	4.53	4.28	4.07	3.89	3.73	3.6
100	633466.7	999.53	124.07	44.47	24.11	16.03	11.95	9.57	8.04	6.98	6.21	5.63	5.17	4.81	4.51	4.26	4.05	3.87	3.71	3.58
120	634098.7	999.53	123.97	44.4	24.06	15.98	11.91	9.53	8	6.94	6.18	5.59	5.14	4.77	4.48	4.23	4.02	3.84	3.68	3.54
∞	636626.6	999.53	123.47	44.06	23.79	15.75	11.7	9.34	7.81	6.76	6	5.42	4.97	4.61	4.31	4.06	3.85	3.67	3.52	3.38

Tabla VI.6 (cont.): Distribución $F(m, n)$. $\alpha = 0.001$

1^a fila= g.l. denominador. 1^a columna= g.l. numerador.

	21	22	23	24	25	26	27	28	29	30	40	50	60	70	80	90	100	110	120	∞
1	14.59	14.38	14.19	14.03	13.88	13.74	13.61	13.5	13.39	13.29	12.61	12.22	11.97	11.8	11.67	11.57	11.5	11.43	11.38	10.83
2	9.77	9.61	9.47	9.34	9.22	9.12	9.02	8.93	8.85	8.77	8.25	7.96	7.77	7.64	7.54	7.47	7.41	7.36	7.32	6.91
3	7.94	7.8	7.67	7.55	7.45	7.36	7.27	7.19	7.12	7.05	6.59	6.34	6.17	6.06	5.97	5.91	5.86	5.82	5.78	5.43
4	6.95	6.81	6.7	6.59	6.49	6.41	6.33	6.25	6.19	6.12	5.7	5.46	5.31	5.2	5.12	5.06	5.02	4.98	4.95	4.62
5	6.32	6.19	6.08	5.98	5.89	5.8	5.73	5.66	5.59	5.53	5.13	4.9	4.76	4.66	4.58	4.53	4.48	4.45	4.42	4.11
6	5.88	5.76	5.65	5.55	5.46	5.38	5.31	5.24	5.18	5.12	4.73	4.51	4.37	4.28	4.2	4.15	4.11	4.07	4.04	3.75
7	5.56	5.44	5.33	5.23	5.15	5.07	5	4.93	4.87	4.82	4.44	4.22	4.09	3.99	3.92	3.87	3.83	3.79	3.77	3.48
8	5.31	5.19	5.09	4.99	4.91	4.83	4.76	4.69	4.64	4.58	4.21	4	3.86	3.77	3.7	3.65	3.61	3.58	3.55	3.27
9	5.11	4.99	4.89	4.8	4.71	4.64	4.57	4.5	4.45	4.39	4.02	3.82	3.69	3.6	3.53	3.48	3.44	3.41	3.38	3.1
10	4.95	4.83	4.73	4.64	4.56	4.48	4.41	4.35	4.29	4.24	3.87	3.67	3.54	3.45	3.39	3.34	3.3	3.26	3.24	2.96
11	4.81	4.7	4.6	4.51	4.42	4.35	4.28	4.22	4.16	4.11	3.75	3.55	3.42	3.33	3.27	3.22	3.18	3.14	3.12	2.85
12	4.7	4.58	4.48	4.39	4.31	4.24	4.17	4.11	4.05	4	3.64	3.44	3.32	3.23	3.16	3.11	3.07	3.04	3.02	2.75
13	4.6	4.49	4.39	4.3	4.22	4.14	4.08	4.01	3.96	3.91	3.55	3.35	3.23	3.14	3.07	3.02	2.99	2.95	2.93	2.66
14	4.51	4.4	4.3	4.21	4.13	4.06	3.99	3.93	3.88	3.82	3.47	3.27	3.15	3.06	3	2.95	2.91	2.88	2.85	2.58
15	4.44	4.33	4.23	4.14	4.06	3.99	3.92	3.86	3.8	3.75	3.4	3.2	3.08	2.99	2.93	2.88	2.84	2.81	2.78	2.52
16	4.37	4.26	4.16	4.07	3.99	3.92	3.86	3.8	3.74	3.69	3.34	3.14	3.02	2.93	2.87	2.82	2.78	2.75	2.72	2.46
17	4.31	4.2	4.1	4.02	3.94	3.86	3.8	3.74	3.68	3.63	3.28	3.09	2.96	2.88	2.81	2.76	2.73	2.69	2.67	2.4
18	4.26	4.15	4.05	3.96	3.88	3.81	3.75	3.69	3.63	3.58	3.23	3.04	2.91	2.83	2.76	2.71	2.68	2.65	2.62	2.35
19	4.21	4.1	4	3.92	3.84	3.77	3.7	3.64	3.59	3.53	3.19	2.99	2.87	2.78	2.72	2.67	2.63	2.6	2.58	2.31
20	4.17	4.06	3.96	3.87	3.79	3.72	3.66	3.6	3.54	3.49	3.14	2.95	2.83	2.74	2.68	2.63	2.59	2.56	2.53	2.27
21	4.13	4.02	3.92	3.83	3.76	3.68	3.62	3.56	3.5	3.45	3.11	2.91	2.79	2.7	2.64	2.59	2.55	2.52	2.5	2.23
22	4.09	3.98	3.89	3.8	3.72	3.65	3.58	3.52	3.47	3.42	3.07	2.88	2.75	2.67	2.61	2.56	2.52	2.49	2.46	2.2
23	4.06	3.95	3.85	3.77	3.69	3.62	3.55	3.49	3.44	3.39	3.04	2.85	2.72	2.64	2.57	2.53	2.49	2.46	2.43	2.17
24	4.03	3.92	3.82	3.74	3.66	3.59	3.52	3.46	3.41	3.36	3.01	2.82	2.69	2.61	2.54	2.5	2.46	2.43	2.4	2.14
25	4	3.89	3.79	3.71	3.63	3.56	3.49	3.43	3.38	3.33	2.98	2.79	2.67	2.58	2.52	2.47	2.43	2.4	2.37	2.11
26	3.97	3.86	3.77	3.68	3.6	3.53	3.47	3.41	3.35	3.3	2.96	2.76	2.64	2.56	2.49	2.44	2.41	2.37	2.35	2.08
27	3.95	3.84	3.74	3.66	3.58	3.51	3.44	3.38	3.33	3.28	2.93	2.74	2.62	2.53	2.47	2.42	2.38	2.35	2.33	2.06
28	3.93	3.82	3.72	3.63	3.56	3.49	3.42	3.36	3.31	3.26	2.91	2.72	2.6	2.51	2.45	2.4	2.36	2.33	2.3	2.04
29	3.9	3.8	3.7	3.61	3.54	3.46	3.4	3.34	3.29	3.24	2.89	2.7	2.57	2.49	2.43	2.38	2.34	2.31	2.28	2.01
30	3.88	3.78	3.68	3.59	3.52	3.44	3.38	3.32	3.27	3.22	2.87	2.68	2.55	2.47	2.41	2.36	2.32	2.29	2.26	1.99
40	3.74	3.63	3.53	3.45	3.37	3.3	3.23	3.18	3.12	3.07	2.73	2.53	2.41	2.32	2.26	2.21	2.17	2.14	2.11	1.97
50	3.64	3.54	3.44	3.36	3.28	3.21	3.14	3.09	3.03	2.98	2.64	2.44	2.32	2.23	2.16	2.11	2.08	2.04	2.02	1.96
60	3.58	3.48	3.38	3.29	3.22	3.15	3.08	3.02	2.97	2.92	2.57	2.38	2.25	2.16	2.1	2.05	2.01	1.98	1.95	1.94
70	3.54	3.43	3.34	3.25	3.17	3.1	3.04	2.98	2.92	2.87	2.53	2.33	2.21	2.12	2.05	2	1.96	1.93	1.9	1.92
80	3.5	3.4	3.3	3.22	3.14	3.07	3	2.94	2.89	2.84	2.49	2.3	2.17	2.08	2.01	1.96	1.92	1.89	1.86	1.91
90	3.48	3.37	3.28	3.19	3.11	3.04	2.98	2.92	2.86	2.81	2.47	2.27	2.14	2.05	1.98	1.93	1.89	1.86	1.83	1.89
100	3.46	3.35	3.25	3.17	3.09	3.02	2.96	2.9	2.84	2.79	2.44	2.25	2.12	2.03	1.96	1.91	1.87	1.83	1.81	1.88
120	3.42	3.32	3.22	3.14	3.06	2.99	2.92	2.86	2.81	2.76	2.41	2.21	2.08	1.99	1.92	1.87	1.83	1.8	1.77	1.85
∞	3.26	3.15	3.06	2.97	2.89	2.82	2.76	2.7	2.64	2.59	2.54	2.5	2.46	2.42	2.39	2.35	2.32	2.29	2.26	1.06

Tabla VII: Test de normalidad de D'Agostino (p. 168)

n	α				
	0.20	0.10	0.05	0.02	0.01
10	0.2632—0.2835	0.2573—0.2843	0.2513—0.2849	0.2436—0.2855	0.2379—0.2857
12	0.2653—0.2841	0.2598—0.2849	0.2544—0.2854	0.2473—0.2859	0.2420—0.2862
14	0.2669—0.2846	0.2618—0.2853	0.2568—0.2858	0.2503—0.2862	0.2455—0.2865
16	0.2081—0.2848	0.2634—0.2855	0.2587—0.2860	0.2527—0.2865	0.2482—0.2867
18	0.2690—0.2850	0.2646—0.2857	0.2603—0.2862	0.2547—0.2866	0.2505—0.2868
20	0.2699—0.2852	0.2657—0.2859	0.2617—0.2863	0.2564—0.2867	0.2525—0.2869
22	0.2705—0.2853	0.2667—0.2860	0.2629—0.2864	0.2579—0.2869	0.2542—0.2870
24	0.2711—0.2853	0.2675—0.2861	0.2639—0.2865	0.2591—0.2869	0.2557—0.2871
26	0.2717—0.2854	0.2682—0.2861	0.2647—0.2866	0.2603—0.2870	0.2570—0.2872
28	0.2721—0.2854	0.2688—0.2861	0.2655—0.2866	0.2612—0.2870	0.2581—0.2873
30	0.2725—0.2854	0.2693—0.2862	0.2662—0.2866	0.2622—0.2871	0.2592—0.2872
32	0.2729—0.2854	0.2698—0.2862	0.2668—0.2867	0.2630—0.2871	0.2600—0.2873
34	0.2732—0.2854	0.2703—0.2862	0.2674—0.2867	0.2636—0.2871	0.2609—0.2873
36	0.2735—0.2854	0.2707—0.2862	0.2679—0.2867	0.2643—0.2871	0.2617—0.2873
38	0.2738—0.2854	0.2710—0.2862	0.2683—0.2867	0.2649—0.2871	0.2623—0.2873
40	0.2740—0.2854	0.2714—0.2862	0.2688—0.2867	0.2655—0.2871	0.2630—0.2874
42	0.2743—0.2854	0.2717—0.2861	0.2691—0.2867	0.2659—0.2871	0.2636—0.2874
44	0.2745—0.2854	0.2720—0.2861	0.2695—0.2867	0.2664—0.2871	0.2641—0.2874
46	0.2747—0.2854	0.2722—0.2861	0.2698—0.2866	0.2668—0.2871	0.2646—0.2874
48	0.2749—0.2854	0.2725—0.2861	0.2702—0.2866	0.2672—0.2871	0.2651—0.2874
50	0.2751—0.2853	0.2727—0.2861	0.2705—0.2866	0.2676—0.2871	0.2655—0.2874
60	0.2757—0.2852	0.2737—0.2860	0.2717—0.2865	0.2692—0.2870	0.2673—0.2873
70	0.2763—0.2851	0.2744—0.2859	0.2726—0.2864	0.2704—0.2869	0.2687—0.2872
80	0.2768—0.2850	0.2750—0.2857	0.2734—0.2863	0.2713—0.2868	0.2698—0.2871
90	0.2771—0.2849	0.2756—0.2856	0.2741—0.2861	0.2721—0.2866	0.2707—0.2870
100	0.2774—0.2849	0.2759—0.2855	0.2745—0.2860	0.2727—0.2865	0.2714—0.2869
120	0.2779—0.2847	0.2765—0.2853	0.2752—0.2858	0.2737—0.2863	0.2725—0.2866
140	0.2782—0.2846	0.2770—0.2852	0.2758—0.2856	0.2744—0.2862	0.2734—0.2865
150	0.2784—0.2845	0.2772—0.2851	0.2761—0.2856	0.2747—0.2861	0.2737—0.2864
160	0.2785—0.2845	0.2774—0.2851	0.2763—0.2855	0.2750—0.2860	0.2741—0.2863
180	0.2787—0.2844	0.2777—0.2850	0.2767—0.2854	0.2755—0.2859	0.2745—0.2862
200	0.2789—0.2843	0.2779—0.2848	0.2770—0.2853	0.2759—0.2857	0.2751—0.2860
250	0.2793—0.2841	0.2784—0.2846	0.2776—0.2850	0.2767—0.2855	0.2760—0.2857
300	0.2796—0.2840	0.2788—0.2844	0.2781—0.2847	0.2772—0.2853	0.2766—0.2855
350	0.2798—0.2839	0.2791—0.2843	0.2784—0.2847	0.2776—0.2851	0.2771—0.2853
400	0.2799—0.2838	0.2793—0.2842	0.2787—0.2845	0.2780—0.2849	0.2775—0.2852
450	0.2801—0.2837	0.2795—0.2841	0.2789—0.2844	0.2782—0.2848	0.2778—0.2850
500	0.2802—0.2836	0.2796—0.2840	0.2791—0.2843	0.2785—0.2847	0.2780—0.2849
550	0.2803—0.2835	0.2797—0.2839	0.2792—0.2842	0.2787—0.2846	0.2782—0.2848
600	0.2804—0.2835	0.2799—0.2839	0.2794—0.2842	0.2788—0.2845	0.2784—0.2847
650	0.2804—0.2834	0.2799—0.2838	0.2795—0.2841	0.2790—0.2844	0.2786—0.2846
700	0.2805—0.2834	0.2800—0.2837	0.2796—0.2840	0.2791—0.2844	0.2787—0.2846
750	0.2806—0.2834	0.2801—0.2837	0.2797—0.2840	0.2792—0.2843	0.2789—0.2845
800	0.2806—0.2833	0.2802—0.2837	0.2798—0.2839	0.2793—0.2842	0.2790—0.2844
850	0.2807—0.2833	0.2802—0.2836	0.2799—0.2839	0.2794—0.2842	0.2791—0.2844
900	0.2807—0.2833	0.2803—0.2836	0.2799—0.2838	0.2795—0.2841	0.2792—0.2843
950	0.2807—0.2832	0.2803—0.2835	0.2800—0.2838	0.2796—0.2841	0.2793—0.2843
1000	0.2808—0.2832	0.2804—0.2835	0.2800—0.2838	0.2796—0.2840	0.2793—0.2842
1500	0.2810—0.2830	0.2807—0.2833	0.2805—0.2835	0.2801—0.2837	0.2799—0.2839
2000	0.2812—0.2829	0.2809—0.2831	0.2807—0.2833	0.2804—0.2835	0.2802—0.2837

Nota: Esta tabla ha sido tomada de Martín, Luna (1990). Proporciona dos valores $a_1(n, \alpha) < a_2(n, \alpha)$.

Tabla VIII: (★) Test de Mann-Whitney-Wilcoxon: región aceptación (p. 182)

m	n	α		
		0.10	0.05	0.01
4	4	11—25	10—26	—
4	5	12—28	11—29	—
4	6	13—31	12—32	10—34
4	7	14—34	13—35	10—38
4	8	15—37	14—38	11—41
4	9	16—40	14—42	11—45
4	10	17—43	15—45	12—48
4	11	18—46	16—48	12—52
4	12	19—49	17—51	13—55
4	13	20—52	18—54	13—59
4	14	21—55	19—57	14—62
4	15	22—58	20—60	15—65
4	16	24—60	21—63	15—69
4	17	25—63	21—67	16—72
4	18	26—66	22—70	16—76
4	19	27—69	23—73	17—79
4	20	28—72	24—76	18—82
4	21	29—75	25—79	18—86
4	22	30—78	26—82	19—89
4	23	31—81	27—85	19—93
4	24	32—84	27—89	20—96
4	25	33—87	28—92	20—100
4	26	34—90	29—95	21—103
5	5	19—36	17—38	15—40
5	6	20—40	18—42	16—44
5	7	21—44	20—45	16—49
5	8	23—47	21—49	17—53
5	9	24—51	22—53	18—57
5	10	26—54	23—57	19—61
5	11	27—58	24—61	20—65
5	12	28—62	26—64	21—69
5	13	30—65	27—68	22—73
5	14	31—69	28—72	22—78
5	15	33—72	29—76	23—82
5	16	34—76	30—80	24—86
5	17	35—80	32—83	25—90
5	18	37—83	33—87	26—94
5	19	38—87	34—91	27—98
5	20	40—90	35—95	28—102
5	21	41—94	37—98	29—106
5	22	43—97	38—102	29—111
5	23	44—101	39—106	30—115
5	24	45—105	40—110	31—119
5	25	47—108	42—113	32—123
6	6	28—50	26—52	23—55
6	7	29—55	27—57	24—60
6	8	31—59	29—61	25—65
6	9	33—63	31—65	26—70
6	10	35—67	32—70	27—75
6	11	37—71	34—74	28—80
6	12	38—76	35—79	30—84
6	13	40—80	37—83	31—89
6	14	42—84	38—88	32—94
6	15	44—88	40—92	33—99
6	16	46—92	42—96	34—104
6	17	47—97	43—101	36—108
6	18	49—101	45—105	37—113
6	19	51—105	46—110	38—118
6	20	53—109	48—114	39—123
6	21	55—113	50—118	40—128
6	22	57—117	51—123	42—132
6	23	58—122	53—127	43—137
6	24	60—126	54—132	44—142
7	7	39—66	36—69	32—73
7	8	41—71	38—74	34—78
7	9	43—76	40—79	35—84
7	10	45—81	42—84	37—89
7	11	47—86	44—89	38—95
7	12	49—91	46—94	40—100
7	13	52—95	48—99	41—106
1	14	54—100	50—104	43—111
7	15	56—105	52—109	44—117

m	n	α		
		0.10	0.05	0.01
7	16	58—110	54—114	46—122
7	17	61—114	56—119	47—128
7	18	63—119	58—124	49—133
7	19	65—124	60—129	50—139
7	20	67—129	62—134	52—144
7	21	69—134	64—139	53—150
7	22	72—138	66—144	55—155
7	23	74—143	68—149	57—160
8	8	51—85	49—87	43—93
8	9	54—90	51—93	45—99
8	10	56—96	53—99	47—105
8	11	59—101	55—105	49—111
8	12	62—106	58—110	51—117
8	13	64—112	60—116	53—123
8	14	67—117	62—122	54—130
8	15	69—123	65—127	56—136
8	16	72—128	67—133	58—142
8	17	75—133	70—138	60—148
8	18	77—139	72—144	62—154
8	19	80—144	74—150	64—160
8	20	83—149	77—155	66—166
8	21	85—155	79—161	68—172
8	22	88—160	81—167	70—178
9	9	66—105	62—109	56—115
9	10	69—111	65—115	58—122
9	11	72—117	68—121	61—128
9	12	75—123	71—127	63—135
9	13	78—129	73—134	65—142
9	14	81—135	76—140	67—149
9	15	84—141	79—146	69—156
9	16	87—147	82—152	72—162
9	17	90—153	84—159	74—169
9	18	93—159	87—165	76—176
9	19	96—165	90—171	78—183
9	20	99—171	93—177	81—189
9	21	102—177	95—184	83—196
10	10	82—128	78—132	71—139
10	11	86—134	81—139	73—147
10	12	89—141	84—146	76—154
10	13	92—148	88—152	79—161
10	14	96—154	91—159	81—169
10	15	99—161	94—166	84—176
10	16	103—167	97—173	86—184
10	17	106—174	100—180	89—191
10	18	110—180	103—187	92—198
10	19	113—187	107—193	94—206
10	20	117—193	110—200	97—213
11	11	100—153	96—157	87—166
11	12	104—160	99—165	90—174
11	13	108—167	103—172	93—182
11	14	112—174	106—180	96—190
11	15	116—181	110—187	99—198
11	16	120—188	113—195	102—206
11	17	123—196	117—202	105—214
11	18	127—203	121—209	108—222
11	19	131—210	124—217	111—230
12	12	120—180	115—185	105—195
12	13	125—187	119—193	109—203
12	14	129—195	123—201	112—212
12	15	133—203	127—209	115—221
12	16	138—210	131—217	119—229
12	17	142—218	135—225	122—238
12	18	146—226	139—233	125—247
13	13	142—209	136—215	125—226
13	14	147—217	141—223	129—235
13	15	152—225	145—232	133—244
13	16	156—234	150—240	136—254
13	17	161—242	154—249	140—263
14	14	166—240	160—246	147—259
14	15	171—249	164—256	151—269
14	16	176—258	169—265	155—279
15	15	191—274	184—281	171—294

Nota: La ausencia de límites se indica con un guión. Esta tabla ha sido extraída de Martín, Luna (1990).

Tabla IX: (★) Distribución T de Wilcoxon para muestras relacionadas (p. 185)

$T_{\alpha}(n)$	α							
	0.50	0.20	0.10	0.05	0.02	0.01	0.005	0.001
4	2	0	—	—	—	—	—	—
5	4	2	0	—	—	—	—	—
6	6	3	2	0	—	—	—	—
7	9	5	3	2	0	—	—	—
8	12	8	5	3	1	0	—	—
9	16	10	8	5	3	1	0	—
10	20	14	10	8	5	3	1	—
11	24	17	13	10	7	5	3	0
12	29	21	17	13	9	7	5	1
13	35	26	21	17	12	9	7	2
14	40	31	25	21	15	12	9	4
15	47	36	30	25	19	15	12	6
16	54	42	35	29	23	19	15	8
17	61	48	41	34	27	23	19	11
18	69	55	47	40	32	27	23	14
19	77	62	53	46	37	32	27	18
20	86	69	60	52	43	37	32	21
21	95	77	67	58	49	42	37	25
22	104	86	75	65	55	48	42	30
23	114	94	83	73	62	54	48	35
24	125	104	91	81	69	61	54	40
25	136	113	100	89	76	68	60	45
26	148	124	110	98	84	75	67	51
27	160	134	119	107	92	83	74	57
28	172	145	130	116	101	91	82	64
29	185	157	140	126	110	100	90	71
30	198	169	151	137	120	109	98	78
31	212	181	163	147	130	118	107	86
32	226	194	175	159	140	128	116	94
33	241	207	187	170	151	138	126	102
34	257	221	200	182	162	148	136	111
35	272	235	213	195	173	159	146	120
36	289	250	227	208	185	171	157	130
37	305	265	241	221	198	182	168	140
38	323	281	256	235	211	194	180	150
39	340	297	271	249	224	207	192	161
40	358	313	286	264	238	220	204	172
41	377	330	302	279	252	233	217	183
42	396	348	319	294	266	247	230	195
43	416	365	336	310	281	261	244	207
44	436	384	353	327	296	276	258	220
45	456	402	371	343	312	291	272	233
46	477	422	389	361	328	307	287	246
47	499	441	407	378	345	322	302	260
48	521	462	426	396	362	339	318	274
49	543	482	446	415	379	355	334	289
50	566	503	466	434	397	373	350	304
52	613	547	507	473	434	408	384	335
54	668	592	550	514	473	445	420	368
56	714	639	595	557	514	484	457	402
58	767	688	642	602	556	525	497	438
60	822	739	690	648	600	567	537	476
62	879	792	741	697	646	611	580	515
64	938	847	793	747	693	657	624	556
66	998	903	847	798	742	705	670	599
68	1061	962	903	852	793	754	718	643
70	1126	1022	960	907	846	805	767	689
72	1192	1084	1020	964	901	858	818	736
74	1261	1148	1081	1023	957	912	871	786
76	1331	1214	1144	1084	1015	968	925	836
78	1403	1282	1209	1147	1075	1026	981	889
80	1478	1351	1276	1211	1136	1086	1039	943
82	1554	1423	1345	1277	1200	1147	1099	999
84	1632	1496	1415	1345	1265	1210	1160	1057
86	1712	1571	1487	1415	1332	1275	1223	1116
88	1794	1648	1561	1487	1400	1342	1288	1177
90	1878	1727	1638	1560	1471	1410	1355	1240
92	1964	1808	1715	1635	1543	1480	1423	1304
94	2052	1891	1795	1712	1617	1552	1493	1370
96	2142	1976	1877	1791	1693	1626	1565	1438
98	2233	2062	1960	1872	1770	1702	1638	1507
100	2327	2151	2045	1955	1850	1779	1714	1578

Nota: La ausencia de límites se indica con un guión. Esta tabla ha sido extraída de la tabla B.12 de Zar (1999).

Tabla X.1: Distribución q de Tukey para comparaciones múltiples (p. 211)

$\alpha = 0.1$

$q_\alpha(n, k)$	k													
n	2	3	4	5	6	7	8	9	10	11	12	13	14	
1	8.93	13.4	16.4	18.5	20.2	21.5	22.6	23.6	24.5	25.2	25.9	26.5	27.1	
2	4.13	5.73	6.77	1.54	8.14	8.63	9.05	9.41	9.73	10.0	10.3	10.5	10.7	
3	3.33	4.47	5.2	5.74	6.16	6.51	6.81	7.06	7.29	7.49	7.67	7.83	7.98	
4	3.02	3.97	4.59	5.04	5.39	5.68	5.93	6.14	6.33	6.50	6.65	6.78	6.91	
5	2.85	3.72	4.26	4.66	4.98	5.24	5.46	5.65	5.82	5.97	6.10	6.22	6.34	
6	2.75	3.56	4.06	4.44	4.73	4.97	5.17	5.34	5.50	5.64	5.76	5.88	5.98	
7	2.68	3.45	3.93	4.28	4.56	4.78	4.97	5.14	5.28	5.41	5.53	5.64	5.74	
8	2.63	3.37	3.83	4.17	4.43	4.65	4.83	4.99	5.13	5.25	5.36	5.46	5.56	
9	2.59	3.32	3.76	4.08	4.34	4.55	4.72	4.87	5.01	5.13	5.23	5.33	5.42	
10	2.56	3.27	3.70	4.02	4.26	4.47	4.64	4.78	4.91	5.03	5.13	5.23	5.32	
11	2.54	3.23	3.66	3.97	4.21	4.40	4.57	4.71	4.84	4.95	5.05	5.15	5.23	
12	2.52	3.20	3.62	3.92	4.16	4.35	4.51	4.65	4.78	4.89	4.99	5.08	5.16	
13	2.51	3.18	3.59	3.89	4.12	4.31	4.46	4.60	4.72	4.83	4.93	5.02	5.10	
14	2.49	3.16	3.56	3.85	4.08	4.27	4.42	4.56	4.68	4.79	4.88	4.97	5.05	
15	2.48	3.14	3.54	3.83	4.05	4.24	4.39	4.52	4.64	4.75	4.84	4.93	5.01	
16	2.47	3.12	3.52	3.80	4.03	4.21	4.36	4.49	4.61	4.71	4.81	4.89	4.97	
17	2.46	3.11	3.50	3.78	4.00	4.18	4.33	4.46	4.58	4.68	4.77	4.86	4.94	
18	2.45	3.10	3.49	3.77	3.98	4.16	4.31	4.44	4.55	4.66	4.75	4.83	4.91	
19	2.45	3.09	3.47	3.75	3.97	4.14	4.29	4.42	4.53	4.63	4.72	4.80	4.88	
20	2.44	3.08	3.46	3.74	3.95	4.12	4.27	4.40	4.51	4.61	4.70	4.78	4.86	
24	2.42	3.05	3.42	3.69	3.90	4.07	4.21	4.34	4.45	4.54	4.63	4.71	4.78	
30	2.40	3.02	3.39	3.65	3.85	4.02	4.16	4.28	4.38	4.47	4.56	4.64	4.71	
40	2.38	2.99	3.35	3.61	3.80	3.96	4.10	4.22	4.32	4.41	4.49	4.56	4.63	
60	2.36	2.96	3.31	3.56	3.76	3.91	4.04	4.16	4.25	4.34	4.42	4.49	4.56	
120	2.34	2.93	3.28	3.52	3.70	3.86	3.99	4.10	4.19	4.28	4.35	4.42	4.49	
∞	2.33	2.90	3.24	3.48	3.66	3.81	3.93	4.04	4.13	4.21	4.29	4.35	4.41	

$q_\alpha(n, k)$	k													
n	15	16	17	18	19	20	30	40	50	60	70	80	90	100
1	27.6	28.1	28.5	29.0	29.4	29.7	32.5	34.4	35.8	36.9	37.8	38.6	39.3	39.9
2	10.9	11.1	11.2	11.4	11.5	11.7	12.7	13.4	14.0	14.4	14.8	15.1	15.3	15.5
3	8.12	8.25	8.37	8.48	8.58	8.68	9.44	9.95	10.3	10.7	10.9	11.1	11.3	11.5
4	7.03	7.13	7.23	7.33	7.41	7.50	8.14	8.57	8.90	9.16	9.37	9.56	9.72	9.86
5	6.44	6.54	6.63	6.71	6.79	6.86	7.44	7.83	8.12	8.35	8.55	8.72	8.86	8.99
6	6.08	6.16	6.25	6.33	6.40	6.47	7.00	7.36	7.63	7.85	8.03	8.18	8.32	8.44
7	5.83	5.91	5.99	6.06	6.13	6.20	6.70	7.04	7.29	7.50	7.67	7.82	7.95	8.06
8	5.64	5.72	5.80	5.87	5.94	6.00	6.48	6.80	7.05	7.25	7.41	7.55	7.67	7.78
9	5.51	5.58	5.66	5.72	5.79	5.85	6.31	6.62	6.86	7.05	7.21	7.34	7.46	7.57
10	5.40	5.47	5.54	5.61	5.67	5.73	6.17	6.48	6.71	6.90	7.05	7.18	7.30	7.40
11	5.31	5.38	5.45	5.51	5.57	5.63	6.07	6.36	6.59	6.77	6.92	7.05	7.16	7.26
12	5.24	5.31	5.37	5.44	5.50	5.55	5.98	6.27	6.49	6.66	6.81	6.94	7.05	7.14
13	5.18	5.25	5.31	5.37	5.43	5.48	5.90	6.19	6.40	6.58	6.72	6.84	6.95	7.05
14	5.12	5.19	5.26	5.32	5.37	5.43	5.84	6.12	6.33	6.50	6.64	6.76	6.87	6.96
15	5.08	5.15	5.21	5.27	5.32	5.38	5.78	6.06	6.27	6.43	6.57	6.69	6.80	6.89
16	5.04	5.11	5.17	5.23	5.28	5.33	5.73	6.00	6.21	6.38	6.51	6.63	6.73	6.83
17	5.01	5.07	5.13	5.19	5.24	5.30	5.69	5.96	6.16	6.33	6.46	6.58	6.68	6.77
18	4.98	5.04	5.10	5.16	5.21	5.26	5.65	5.92	6.12	6.28	6.41	6.53	6.63	6.72
19	4.95	5.01	5.07	5.13	5.18	5.23	5.62	5.88	6.08	6.24	6.37	6.49	6.59	6.67
20	4.92	4.99	5.05	5.10	5.16	5.21	5.59	5.85	6.04	6.20	6.34	6.45	6.55	6.63
24	4.85	4.91	4.97	5.02	5.07	5.12	5.49	5.74	5.93	6.09	6.21	6.32	6.42	6.5
30	4.77	4.83	4.89	4.94	4.99	5.03	5.39	5.64	5.82	5.97	6.09	6.20	6.29	6.37
40	4.70	4.75	4.81	4.86	4.91	4.95	5.29	5.53	5.71	5.85	5.97	6.07	6.16	6.24
60	4.62	4.68	4.73	4.78	4.82	4.86	5.20	5.42	5.59	5.73	5.84	5.94	6.03	6.1
120	4.54	4.60	4.65	4.69	4.74	4.78	5.10	5.31	5.48	5.61	5.72	5.81	5.89	5.96
∞	4.47	4.52	4.57	4.61	4.65	4.69	5.00	5.20	5.36	5.48	5.58	5.67	5.75	5.81

Nota: Esta tabla y las dos siguientes han sido extraídas de la tabla B.5 de Zar (1999).

Tabla X.2: Distribución q de Tukey para comparaciones múltiples

$\alpha = 0.05$

$q_\alpha(n, k)$	k													
n	2	3	4	5	6	7	8	9	10	11	12	13	14	
1	18.0	27.0	32.8	37.1	40.4	43.1	45.4	47.4	49.1	50.6	52.0	53.2	54.3	
2	6.09	8.33	9.80	10.9	11.7	12.4	13.0	13.5	14.0	14.4	14.8	15.1	15.4	
3	4.50	5.91	6.83	7.50	8.04	8.48	8.85	9.18	9.46	9.72	9.95	10.2	10.4	
4	3.93	5.04	5.76	6.29	6.71	7.05	7.35	7.60	7.83	8.03	8.21	8.37	8.53	
5	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	7.00	7.17	7.32	7.47	7.60	
6	3.46	4.34	4.90	5.31	5.63	5.90	6.12	6.32	6.49	6.65	6.79	6.92	7.03	
7	3.34	4.17	4.68	5.06	5.36	5.61	5.82	6.00	6.16	6.30	6.43	6.55	6.66	
8	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05	6.18	6.29	6.39	
9	3.20	3.95	4.42	4.76	5.02	5.24	5.43	5.60	5.74	5.87	5.98	6.09	6.19	
10	3.15	3.88	4.33	4.65	4.91	5.12	5.31	5.46	5.60	5.72	5.83	5.94	6.03	
11	3.11	3.82	4.26	4.57	4.82	5.03	5.20	5.35	5.49	5.61	5.71	5.81	5.90	
12	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.40	5.51	5.62	5.71	5.80	
13	3.06	3.74	4.15	4.45	4.69	4.89	5.05	5.19	5.32	5.43	5.53	5.63	5.71	
14	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25	5.36	5.46	5.55	5.64	
15	3.01	3.67	4.08	4.37	4.60	4.78	4.94	5.08	5.20	5.31	5.40	5.49	5.57	
16	3.00	3.65	4.05	4.33	4.56	4.74	4.90	5.03	5.15	5.26	5.35	5.44	5.52	
17	2.98	3.63	4.02	4.30	4.52	4.71	4.86	4.99	5.11	5.21	5.31	5.39	5.47	
18	2.97	3.61	4.00	4.28	4.50	4.67	4.82	4.96	5.07	5.17	5.27	5.35	5.43	
19	2.96	3.59	3.98	4.25	4.47	4.65	4.79	4.92	5.04	5.14	5.23	5.32	5.39	
20	2.95	3.58	3.96	4.23	4.45	4.62	4.77	4.90	5.01	5.11	5.20	5.28	5.36	
24	2.92	3.53	3.90	4.17	4.37	4.54	4.68	4.81	4.92	5.01	5.10	5.18	5.25	
30	2.89	3.49	3.85	4.10	4.30	4.46	4.60	4.72	4.82	4.92	5.00	5.08	5.15	
40	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.64	4.74	4.82	4.90	4.98	5.04	
60	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65	4.73	4.81	4.88	4.94	
120	2.80	3.36	3.69	3.92	4.10	4.24	4.36	4.47	4.56	4.64	4.71	4.78	4.84	
∞	2.77	3.31	3.63	3.86	4.03	4.17	4.29	4.39	4.47	4.55	4.62	4.69	4.74	

$q_\alpha(n, k)$	k													
n	15	16	17	18	19	20	30	40	50	60	70	80	90	100
1	55.4	56.3	57.2	58.0	58.8	59.6	65.2	68.9	71.7	74.0	75.8	77.4	78.8	80.0
2	15.7	15.9	16.1	16.4	16.6	16.8	18.3	19.3	20.1	20.7	21.2	21.6	22.0	22.3
3	10.5	10.7	10.8	11.0	11.1	11.2	12.2	12.9	13.4	13.8	14.1	14.4	14.6	14.8
4	8.66	8.79	8.91	9.03	9.13	9.23	10.0	10.5	10.9	11.2	11.5	11.7	11.9	12.1
5	7.72	7.83	7.93	8.03	8.12	8.21	8.88	9.33	9.67	9.95	10.2	10.4	10.5	10.7
6	7.14	7.24	7.34	7.43	7.51	7.59	8.19	8.60	8.91	9.16	9.37	9.55	9.70	9.84
7	6.76	6.85	6.94	7.02	7.10	7.17	7.73	8.11	8.40	8.63	8.82	8.99	9.13	9.26
8	6.48	6.57	6.65	6.73	6.80	6.87	7.40	7.76	8.03	8.25	8.43	8.59	8.72	8.84
9	6.28	6.36	6.44	6.51	6.58	6.64	7.15	7.49	7.75	7.96	8.13	8.28	8.41	8.53
10	6.11	6.19	6.27	6.34	6.41	6.47	6.95	7.28	7.53	7.73	7.90	8.04	8.17	8.28
11	5.98	6.06	6.13	6.20	6.27	6.33	6.79	7.11	7.35	7.55	7.71	7.85	7.97	8.08
12	5.88	5.95	6.02	6.09	6.15	6.21	6.66	6.97	7.21	7.39	7.55	7.69	7.80	7.91
13	5.79	5.86	5.93	6.00	6.06	6.11	6.55	6.85	7.08	7.27	7.42	7.55	7.67	7.77
14	5.71	5.79	5.85	5.92	5.97	6.03	6.46	6.75	6.98	7.16	7.31	7.44	7.55	7.65
15	5.65	5.72	5.79	5.85	5.90	5.96	6.38	6.67	6.89	7.07	7.21	7.34	7.45	7.55
16	5.59	5.66	5.73	5.79	5.84	5.90	6.31	6.59	6.81	6.98	7.13	7.25	7.36	7.46
17	5.54	5.61	5.68	5.73	5.79	5.84	6.25	6.53	6.74	6.91	7.05	7.18	7.28	7.38
18	5.50	5.57	5.63	5.69	5.74	5.79	6.20	6.47	6.68	6.85	6.99	7.11	7.21	7.31
19	5.46	5.53	5.59	5.65	5.70	5.75	6.15	6.42	6.63	6.79	6.93	7.05	7.15	7.24
20	5.43	5.49	5.55	5.61	5.66	5.71	6.10	6.37	6.58	6.74	6.88	6.99	7.10	7.19
24	5.32	5.38	5.44	5.49	5.55	5.59	5.97	6.23	6.42	6.58	6.71	6.82	6.92	7.01
30	5.21	5.27	5.33	5.38	5.43	5.48	5.83	6.08	6.27	6.42	6.54	6.65	6.74	6.83
40	5.11	5.16	5.22	5.27	5.31	5.36	5.70	5.93	6.11	6.26	6.38	6.48	6.57	6.65
60	5.00	5.06	5.11	5.15	5.20	5.24	5.57	5.79	5.96	6.09	6.21	6.30	6.39	6.46
120	4.90	4.95	5.00	5.04	5.09	5.13	5.43	5.64	5.80	5.93	6.04	6.13	6.21	6.28
∞	4.80	4.85	4.89	4.93	4.97	5.01	5.30	5.50	5.65	5.76	5.86	5.95	6.02	6.09

Tabla X.3: Distribución q de Tukey para comparaciones múltiples

$\alpha = 0.01$

$q_\alpha(n, k)$	k													
n	2	3	4	5	6	7	8	9	10	11	12	13	14	
1	90.0	135.0	164.3	185.6	202.2	215.8	227.2	237.0	245.6	253.2	260	266.2	271.8	
2	14.0	19.0	22.3	24.7	26.6	28.2	29.5	30.7	31.7	32.6	33.4	34.13	34.81	
3	8.26	10.6	12.2	13.3	14.2	15.0	15.6	16.2	16.7	17.1	17.5	17.9	18.2	
4	6.51	8.12	9.17	10.0	10.6	11.1	11.6	11.9	12.3	12.6	12.8	13.1	13.3	
5	5.70	6.98	7.80	8.42	8.91	9.32	9.67	10.0	10.2	10.5	10.7	10.9	11.1	
6	5.24	6.33	7.03	7.56	7.97	8.32	8.61	8.87	9.10	9.30	9.49	9.65	9.81	
7	4.95	5.92	6.54	7.01	7.37	7.68	7.94	8.17	8.37	8.55	8.71	8.86	9.00	
8	4.75	5.64	6.20	6.63	6.96	7.24	7.47	7.68	7.86	8.03	8.18	8.31	8.44	
9	4.60	5.43	5.96	6.35	6.66	6.92	7.13	7.33	7.50	7.65	7.78	7.91	8.03	
10	4.48	5.27	5.77	6.14	6.43	6.67	6.88	7.06	7.21	7.36	7.49	7.60	7.71	
11	4.39	5.15	5.62	5.97	6.25	6.48	6.67	6.84	6.99	7.13	7.25	7.36	7.47	
12	4.32	5.05	5.50	5.84	6.10	6.32	6.51	6.67	6.81	6.94	7.06	7.17	7.27	
13	4.26	4.96	5.40	5.73	5.98	6.19	6.37	6.53	6.67	6.79	6.90	7.01	7.10	
14	4.21	4.90	5.32	5.63	5.88	6.09	6.26	6.41	6.54	6.66	6.77	6.87	6.96	
15	4.17	4.84	5.25	5.56	5.80	5.99	6.16	6.31	6.44	6.56	6.66	6.76	6.85	
16	4.13	4.79	5.19	5.49	5.72	5.92	6.08	6.22	6.35	6.46	6.56	6.66	6.74	
17	4.10	4.74	5.14	5.43	5.66	5.85	6.01	6.15	6.27	6.38	6.48	6.57	6.66	
18	4.07	4.70	5.09	5.38	5.60	5.79	5.94	6.08	6.20	6.31	6.41	6.50	6.58	
19	4.05	4.67	5.05	5.33	5.55	5.74	5.89	6.02	6.14	6.25	6.34	6.43	6.51	
20	4.02	4.64	5.02	5.29	5.51	5.69	5.84	5.97	6.09	6.19	6.29	6.37	6.45	
24	3.96	4.55	4.91	5.17	5.37	5.54	5.69	5.81	5.92	6.02	6.11	6.19	6.26	
30	3.89	4.46	4.80	5.05	5.24	5.40	5.54	5.65	5.76	5.85	5.93	6.01	6.08	
40	3.83	4.37	4.70	4.93	5.11	5.27	5.39	5.50	5.56	5.69	5.76	5.84	5.90	
60	3.76	4.28	4.60	4.82	4.99	5.13	5.25	5.36	5.45	5.53	5.60	5.67	5.73	
120	3.70	4.20	4.50	4.71	4.87	5.01	5.12	5.21	5.30	5.38	5.44	5.51	5.56	
∞	3.64	4.12	4.40	4.60	4.76	4.88	4.99	5.08	5.16	5.23	5.29	5.35	5.40	

$q_\alpha(n, k)$	k													
n	15	16	17	18	19	20	30	40	50	60	70	80	90	100
1	277	281.8	286.3	290.4	294.3	298	326	344.8	358.9	370.1	379.4	387.3	394.1	400.1
2	35.43	36	36.53	37.03	37.5	38.0	41.3	43.6	45.3	46.7	47.8	48.8	49.6	50.4
3	18.5	18.8	19.1	19.3	19.6	19.8	21.4	22.6	23.5	24.1	24.7	25.2	25.6	26.0
4	13.5	13.7	13.9	14.1	14.2	14.4	15.6	16.4	17.0	17.5	17.9	18.0	18.5	18.8
5	11.2	11.4	11.6	11.7	11.8	11.9	12.9	13.5	14.0	14.4	14.7	15.0	15.2	15.5
6	9.95	10.1	10.2	10.3	10.4	10.5	11.3	11.9	12.3	12.7	12.9	13.2	13.4	13.6
7	9.12	9.24	9.35	9.46	9.55	9.65	10.4	10.9	11.2	11.5	11.8	12.0	12.2	12.3
8	8.55	8.66	8.76	8.85	8.94	9.03	9.68	10.1	10.5	10.8	11.0	11.2	11.3	11.5
9	8.13	8.23	8.33	8.41	8.50	8.57	9.18	9.59	9.91	10.2	10.4	10.6	10.7	10.9
10	7.81	7.91	7.99	8.08	8.15	8.23	8.79	9.19	9.49	9.73	9.93	10.1	10.3	10.4
11	7.56	7.65	7.73	7.81	7.88	7.95	8.49	8.86	9.15	9.38	9.57	9.73	9.88	10
12	7.36	7.44	7.52	7.59	7.67	7.73	8.25	8.60	8.88	9.09	9.28	9.43	9.57	9.69
13	7.19	7.27	7.35	7.42	7.49	7.55	8.04	8.39	8.65	8.86	9.04	9.19	9.32	9.44
14	7.05	7.13	7.20	7.27	7.33	7.40	7.87	8.20	8.46	8.66	8.83	8.98	9.11	9.22
15	6.93	7.00	7.07	7.14	7.20	7.26	7.73	8.05	8.30	8.49	8.66	8.80	8.92	9.04
16	6.82	6.90	6.97	7.03	7.09	7.15	7.60	7.92	8.15	8.35	8.51	8.65	8.77	8.87
17	6.73	6.81	6.87	6.94	7.00	7.05	7.49	7.80	8.03	8.22	8.38	8.51	8.63	8.74
18	6.66	6.73	6.79	6.85	6.91	6.97	7.40	7.70	7.92	8.11	8.26	8.39	8.51	8.61
19	6.59	6.65	6.72	6.78	6.84	6.89	7.31	7.61	7.83	8.01	8.16	8.29	8.40	8.50
20	6.52	6.59	6.65	6.71	6.77	6.82	7.24	7.52	7.74	7.92	8.07	8.19	8.31	8.40
24	6.33	6.39	6.45	6.51	6.56	6.61	7.00	7.27	7.48	7.64	7.78	7.90	8.00	8.10
30	6.14	6.20	6.26	6.31	6.36	6.41	6.77	7.02	7.22	7.37	7.50	7.61	7.71	7.80
40	5.96	6.02	6.07	6.12	6.17	6.21	6.55	6.78	6.96	7.10	7.23	7.33	7.42	7.50
60	5.79	5.84	5.89	5.93	5.97	6.02	6.33	6.55	6.71	6.84	6.95	7.05	7.13	7.21
120	5.61	5.66	5.71	5.75	5.79	5.83	6.12	6.32	6.47	6.59	6.69	6.78	6.85	6.92
∞	5.45	5.49	5.54	5.57	5.61	5.65	5.91	6.09	6.23	6.34	6.43	6.51	6.58	6.64

Tabla XI: (★) Distribución Q para comparaciones múltiples no paramétricas (p. 215)

$Q_\alpha(k)$	α								
k	0.50	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
2	0.674	1.282	1.645	1.960	2.327	2.576	2.807	3.091	3.291
3	1.383	1.834	2.128	2.394	2.713	2.936	3.144	3.403	3.588
4	1.732	2.128	2.394	2.639	2.936	3.144	3.342	3.588	3.765
5	1.960	2.327	2.576	2.807	3.091	3.291	3.481	3.719	3.891
6	2.128	2.475	2.713	2.936	3.209	3.403	3.588	3.820	3.988
7	2.261	2.593	2.823	3.038	3.304	3.494	3.675	3.902	4.067
8	2.369	2.690	2.914	3.124	3.384	3.570	3.748	3.972	4.134
9	2.461	2.773	2.992	3.197	3.453	3.635	3.810	4.031	4.191
10	2.540	2.845	3.059	3.261	3.512	3.692	3.865	4.083	4.241
11	2.609	2.908	3.119	3.317	3.565	3.743	3.914	4.129	4.286
12	2.671	2.965	3.172	3.368	3.613	3.789	3.957	4.171	4.326
13	2.726	3.016	3.220	3.414	3.656	3.830	3.997	4.209	4.363
14	2.777	3.062	3.264	3.456	3.695	3.868	4.034	4.244	4.397
15	2.823	3.105	3.304	3.494	3.731	3.902	4.067	4.276	4.428
16	2.866	3.144	3.342	3.529	3.765	3.935	4.098	4.305	4.456
17	2.905	3.181	3.376	3.562	3.796	3.965	4.127	4.333	4.483
18	2.942	3.215	3.409	3.593	3.825	3.993	4.154	4.359	4.508
19	2.976	3.246	3.439	3.622	3.852	4.019	4.179	4.383	4.532
20	3.008	3.276	3.467	3.649	3.878	4.044	4.203	4.406	4.554
21	3.038	3.304	3.494	3.675	3.902	4.067	4.226	4.428	4.575
22	3.067	3.331	3.519	3.699	3.925	4.089	4.247	4.448	4.595
23	3.094	3.356	3.543	3.722	3.947	4.110	4.268	4.468	4.614
24	3.120	3.380	3.566	3.744	3.968	4.130	4.287	4.486	4.632
25	3.144	3.403	3.588	3.765	3.988	4.149	4.305	4.504	4.649

Nota: Tomada de la tabla B.15 de Zar (1999).

Tabla XII: (★) Distribución de Bonferroni (p. 212)

$\alpha = 0.1$

$b_\alpha(n, s)$	s													
	1	2	3	4	5	6	7	8	9	10	20	30	40	50
2	4.13	6.08	7.55	8.78	9.84	10.82	11.71	12.53	13.31	14.03	19.93	24.44	28.23	31.58
3	3.32	4.50	5.29	5.91	6.42	6.87	7.27	7.62	7.96	8.26	10.54	12.13	13.38	14.44
4	3.01	3.93	4.51	4.95	5.30	5.60	5.87	6.10	6.32	6.51	7.92	8.84	9.56	10.14
5	2.86	3.63	4.12	4.47	4.75	4.99	5.20	5.39	5.56	5.70	6.75	7.42	7.92	8.33
6	2.74	3.46	3.89	4.20	4.44	4.65	4.82	4.98	5.12	5.25	6.11	6.65	7.04	7.37
7	2.67	3.34	3.73	4.02	4.24	4.43	4.58	4.72	4.84	4.95	5.70	6.17	6.49	6.77
8	2.63	3.27	3.63	3.89	4.10	4.27	4.41	4.54	4.64	4.75	5.42	5.83	6.12	6.36
9	2.59	3.20	3.55	3.80	3.99	4.14	4.29	4.40	4.50	4.60	5.22	5.59	5.87	6.08
10	2.56	3.15	3.49	3.72	3.90	4.06	4.19	4.30	4.40	4.48	5.06	5.42	5.66	5.85
11	2.55	3.11	3.44	3.66	3.85	3.99	4.12	4.21	4.31	4.40	4.95	5.28	5.50	5.69
12	2.52	3.08	3.39	3.62	3.79	3.93	4.04	4.14	4.24	4.31	4.85	5.16	5.39	5.56
13	2.50	3.05	3.37	3.58	3.75	3.89	4.00	4.10	4.19	4.26	4.77	5.06	5.28	5.44
14	2.49	3.03	3.34	3.55	3.71	3.85	3.96	4.04	4.13	4.21	4.71	4.99	5.19	5.36
15	2.47	3.01	3.31	3.52	3.68	3.80	3.92	4.02	4.09	4.17	4.65	4.92	5.12	5.28
16	2.47	3.00	3.30	3.49	3.65	3.78	3.89	3.97	4.06	4.13	4.60	4.86	5.06	5.22
17	2.46	2.98	3.28	3.48	3.63	3.75	3.86	3.95	4.03	4.10	4.55	4.82	5.01	5.16
18	2.45	2.97	3.25	3.46	3.61	3.73	3.83	3.92	4.00	4.07	4.53	4.78	4.96	5.11
19	2.45	2.96	3.24	3.44	3.59	3.72	3.82	3.90	3.97	4.04	4.48	4.74	4.92	5.06
20	2.43	2.96	3.24	3.42	3.58	3.69	3.79	3.87	3.96	4.03	4.45	4.71	4.89	5.02
24	2.42	2.91	3.20	3.38	3.52	3.63	3.73	3.82	3.89	3.96	4.37	4.61	4.78	4.91
30	2.40	2.88	3.15	3.34	3.48	3.59	3.68	3.76	3.83	3.89	4.29	4.51	4.67	4.79
40	2.38	2.86	3.11	3.30	3.42	3.54	3.62	3.71	3.76	3.82	4.20	4.41	4.57	4.68
60	2.36	2.83	3.08	3.25	3.38	3.48	3.56	3.65	3.71	3.76	4.12	4.33	4.47	4.57
120	2.35	2.80	3.04	3.21	3.34	3.44	3.52	3.59	3.65	3.71	4.04	4.24	4.37	4.47
∞	2.32	2.77	3.01	3.17	3.30	3.38	3.46	3.54	3.59	3.65	3.97	4.16	4.27	4.37

$\alpha = 0.05$

$b_\alpha(n, s)$	s													
	1	2	3	4	5	6	7	8	9	10	20	30	40	50
2	6.08	8.78	10.82	12.53	14.03	15.40	16.65	17.80	18.89	19.93	28.23	34.59	39.97	44.69
3	4.50	5.91	6.87	7.62	8.26	8.81	9.31	9.76	10.15	10.54	13.38	15.37	16.94	18.27
4	3.93	4.95	5.60	6.10	6.51	6.86	7.17	7.44	7.69	7.92	9.56	10.65	11.48	12.18
5	3.63	4.47	4.99	5.39	5.70	5.97	6.19	6.41	6.59	6.75	7.92	8.68	9.25	9.72
6	3.46	4.20	4.65	4.98	5.25	5.46	5.66	5.83	5.97	6.11	7.04	7.64	8.08	8.43
7	3.34	4.02	4.43	4.72	4.95	5.15	5.30	5.46	5.59	5.70	6.49	6.99	7.35	7.65
8	3.27	3.89	4.27	4.54	4.75	4.92	5.06	5.20	5.32	5.42	6.12	6.56	6.87	7.13
9	3.20	3.80	4.14	4.40	4.60	4.75	4.89	5.02	5.12	5.22	5.87	6.25	6.53	6.76
10	3.15	3.72	4.06	4.30	4.48	4.64	4.77	4.88	4.98	5.06	5.66	6.02	6.28	6.49
11	3.11	3.66	3.99	4.21	4.40	4.54	4.65	4.77	4.86	4.95	5.50	5.84	6.08	6.28
12	3.08	3.62	3.93	4.14	4.31	4.45	4.58	4.68	4.77	4.85	5.39	5.70	5.93	6.11
13	3.05	3.58	3.89	4.10	4.26	4.40	4.51	4.61	4.70	4.77	5.28	5.59	5.80	5.97
14	3.03	3.55	3.85	4.04	4.21	4.34	4.45	4.54	4.62	4.71	5.19	5.49	5.70	5.85
15	3.01	3.52	3.80	4.02	4.17	4.30	4.40	4.50	4.57	4.65	5.12	5.40	5.60	5.76
16	3.00	3.49	3.78	3.97	4.13	4.26	4.36	4.45	4.53	4.60	5.06	5.33	5.53	5.67
17	2.98	3.48	3.75	3.95	4.10	4.21	4.33	4.41	4.48	4.55	5.01	5.28	5.46	5.61
18	2.97	3.46	3.73	3.92	4.07	4.19	4.29	4.37	4.45	4.53	4.96	5.22	5.40	5.54
19	2.96	3.44	3.72	3.90	4.04	4.16	4.26	4.34	4.43	4.48	4.92	5.18	5.36	5.49
20	2.96	3.42	3.69	3.87	4.03	4.14	4.24	4.33	4.40	4.45	4.89	5.13	5.30	5.44
24	2.91	3.38	3.63	3.82	3.96	4.07	4.16	4.24	4.31	4.37	4.78	5.01	5.18	5.30
30	2.88	3.34	3.59	3.76	3.89	3.99	4.09	4.16	4.23	4.29	4.67	4.88	5.03	5.16
40	2.86	3.30	3.54	3.71	3.82	3.93	4.02	4.09	4.14	4.20	4.57	4.77	4.91	5.02
60	2.83	3.25	3.48	3.65	3.76	3.86	3.95	4.00	4.07	4.12	4.47	4.65	4.79	4.89
120	2.80	3.21	3.44	3.59	3.71	3.79	3.87	3.93	3.99	4.04	4.37	4.55	4.68	4.77
∞	2.77	3.17	3.38	3.54	3.65	3.73	3.80	3.86	3.92	3.97	4.27	4.44	4.57	4.65

Tabla XII (cont.): Distribución de Bonferroni

$$\alpha = 0.01$$

$b_\alpha(n, s)$	s													
n	1	2	3	4	5	6	7	8	9	10	20	30	40	50
2	14.03	19.93	24.44	28.23	31.58	34.59	37.38	39.97	42.38	44.69	63.22	77.44	89.42	99.98
3	8.26	10.54	12.13	13.38	14.44	15.37	16.19	16.94	17.64	18.27	23.09	26.46	29.13	31.40
4	6.51	7.92	8.84	9.56	10.14	10.65	11.09	11.48	11.85	12.18	14.58	16.18	17.41	18.43
5	5.70	6.75	7.42	7.92	8.33	8.68	8.98	9.25	9.49	9.72	11.29	12.29	13.07	13.69
6	5.25	6.11	6.65	7.04	7.37	7.64	7.86	8.08	8.26	8.43	9.60	10.34	10.90	11.34
7	4.95	5.70	6.17	6.49	6.77	6.99	7.18	7.35	7.51	7.65	8.60	9.19	9.63	9.98
8	4.75	5.42	5.83	6.12	6.36	6.56	6.73	6.87	7.01	7.13	7.95	8.44	8.81	9.11
9	4.60	5.22	5.59	5.87	6.08	6.25	6.41	6.53	6.66	6.76	7.48	7.92	8.24	8.50
10	4.48	5.06	5.42	5.66	5.85	6.02	6.17	6.28	6.39	6.49	7.14	7.54	7.82	8.05
11	4.40	4.95	5.28	5.50	5.69	5.84	5.97	6.08	6.18	6.28	6.87	7.24	7.51	7.71
12	4.31	4.85	5.16	5.39	5.56	5.70	5.83	5.93	6.02	6.11	6.68	7.01	7.25	7.44
13	4.26	4.77	5.06	5.28	5.44	5.59	5.70	5.80	5.88	5.97	6.51	6.82	7.04	7.23
14	4.21	4.71	4.99	5.19	5.36	5.49	5.60	5.70	5.78	5.85	6.36	6.66	6.89	7.06
15	4.17	4.65	4.92	5.12	5.28	5.40	5.52	5.60	5.69	5.76	6.25	6.53	6.75	6.90
16	4.13	4.60	4.86	5.06	5.22	5.33	5.44	5.53	5.60	5.67	6.15	6.42	6.62	6.77
17	4.10	4.55	4.82	5.01	5.16	5.28	5.37	5.46	5.54	5.61	6.07	6.32	6.52	6.66
18	4.07	4.53	4.78	4.96	5.11	5.22	5.32	5.40	5.47	5.54	5.98	6.25	6.43	6.58
19	4.04	4.48	4.74	4.92	5.06	5.18	5.28	5.36	5.43	5.49	5.93	6.17	6.35	6.49
20	4.03	4.45	4.71	4.89	5.02	5.13	5.23	5.30	5.37	5.44	5.87	6.11	6.28	6.42
24	3.96	4.37	4.61	4.78	4.91	5.01	5.09	5.18	5.23	5.30	5.69	5.91	6.07	6.19
30	3.89	4.29	4.51	4.67	4.79	4.88	4.96	5.03	5.11	5.16	5.52	5.73	5.87	5.98
40	3.82	4.20	4.41	4.57	4.68	4.77	4.85	4.91	4.96	5.02	5.36	5.54	5.69	5.78
60	3.76	4.12	4.33	4.47	4.57	4.65	4.72	4.79	4.85	4.89	5.20	5.39	5.50	5.60
120	3.71	4.04	4.24	4.37	4.47	4.55	4.61	4.68	4.72	4.77	5.06	5.22	5.35	5.43
∞	3.65	3.97	4.16	4.27	4.37	4.44	4.51	4.57	4.61	4.65	4.92	5.08	5.18	5.26

Tabla XIII: (★) Tabla para el coeficiente de correlación de Spearman(p. 252)

$r_\alpha(n)$	α						
	0.20	0.10	0.05	0.02	0.01	0.002	0.001
4	1.000	1.000	—	—	—	—	—
5	0.800	0.900	1.000	1.000	—	—	—
6	0.657	0.829	0.886	0.943	1.000	—	—
7	0.571	0.714	0.786	0.893	0.929	1.000	1.000
8	0.524	0.643	0.738	0.833	0.881	0.952	0.976
9	0.483	0.600	0.700	0.783	0.833	0.917	0.933
10	0.455	0.564	0.648	0.745	0.794	0.879	0.903
11	0.427	0.536	0.618	0.709	0.755	0.845	0.873
12	0.406	0.503	0.587	0.678	0.727	0.818	0.846
13	0.385	0.484	0.560	0.648	0.703	0.791	0.824
14	0.367	0.464	0.538	0.626	0.679	0.771	0.802
15	0.354	0.446	0.521	0.604	0.654	0.750	0.779
16	0.341	0.429	0.503	0.582	0.635	0.729	0.762
17	0.328	0.414	0.485	0.566	0.615	0.713	0.748
18	0.317	0.401	0.472	0.550	0.600	0.695	0.728
19	0.309	0.391	0.460	0.535	0.584	0.677	0.712
20	0.299	0.380	0.447	0.520	0.570	0.662	0.696
21	0.292	0.370	0.435	0.508	0.556	0.648	0.681
22	0.284	0.361	0.425	0.496	0.544	0.634	0.667
23	0.278	0.353	0.415	0.486	0.532	0.622	0.654
24	0.271	0.344	0.406	0.476	0.521	0.610	0.642
25	0.265	0.337	0.398	0.466	0.511	0.598	0.630
26	0.259	0.331	0.390	0.457	0.501	0.587	0.619
27	0.255	0.324	0.382	0.448	0.491	0.577	0.608
28	0.250	0.317	0.375	0.440	0.483	0.567	0.598
29	0.245	0.312	0.368	0.433	0.475	0.558	0.589
30	0.240	0.306	0.362	0.425	0.467	0.549	0.580
31	0.236	0.301	0.356	0.418	0.459	0.541	0.571
32	0.232	0.296	0.350	0.412	0.452	0.533	0.563
33	0.229	0.291	0.345	0.405	0.446	0.525	0.554
34	0.225	0.287	0.340	0.399	0.439	0.517	0.547
35	0.222	0.283	0.335	0.394	0.433	0.510	0.539
36	0.219	0.279	0.330	0.388	0.427	0.504	0.533
37	0.216	0.275	0.325	0.383	0.421	0.497	0.526
38	0.212	0.271	0.321	0.378	0.415	0.491	0.519
39	0.210	0.267	0.317	0.373	0.410	0.485	0.513
40	0.207	0.264	0.313	0.368	0.405	0.479	0.507
41	0.204	0.261	0.309	0.364	0.400	0.473	0.501
42	0.202	0.257	0.305	0.359	0.395	0.468	0.495
43	0.199	0.254	0.301	0.355	0.391	0.463	0.490
44	0.197	0.251	0.298	0.351	0.386	0.458	0.484
45	0.194	0.248	0.294	0.347	0.382	0.453	0.479
46	0.192	0.246	0.291	0.343	0.378	0.448	0.474
47	0.190	0.243	0.288	0.340	0.374	0.443	0.469
48	0.188	0.240	0.285	0.336	0.370	0.439	0.465
49	0.186	0.238	0.282	0.333	0.366	0.434	0.460
50	0.184	0.235	0.279	0.329	0.363	0.430	0.456

Nota: La ausencia de límites se indica con un guión. Esta tabla ha sido extraída de la tabla B.20 de Zar (1999).

Tabla XIV: Factores K para límites de tolerancia bilaterales para distribuciones normales (p. 166)

n	$1 - \alpha = 0.95$			$1 - \alpha = 0.99$		
	π			π		
	0.90	0.95	0.99	0.90	0.95	0.99
2	32.019	37.674	48.430	160.193	188.491	242.300
3	8.380	9.916	12.861	18.930	22.401	29.055
4	5.369	6.370	8.299	9.398	11.150	14.527
5	4.275	5.079	6.634	6.612	7.855	10.260
6	3.712	4.414	5.775	5.337	6.345	8.301
7	3.369	4.007	5.248	4.613	5.488	7.187
8	3.136	3.732	4.891	4.147	4.936	6.468
9	2.967	3.532	4.631	3.822	4.550	5.966
10	2.839	3.379	4.433	3.582	4.265	5.594
11	2.737	3.259	4.277	3.397	4.045	5.308
12	2.655	3.162	4.150	3.250	3.870	5.079
13	2.587	3.081	4.044	3.130	3.727	4.893
14	2.529	3.012	3.955	3.029	3.608	4.737
15	2.480	2.954	3.878	2.954	3.507	4.605
16	2.437	2.903	3.812	2.872	3.421	4.492
17	2.400	2.858	3.754	2.808	3.345	4.393
18	2.366	2.819	3.702	2.753	3.279	4.307
11	2.337	2.784	3.656	2.703	3.221	4.230
20	2.310	2.752	3.615	2.659	3.168	4.161
21	2.286	2.723	3.577	2.620	3.121	4.100
22	2.264	2.697	3.543	2.584	3.078	4.044
23	2.244	2.673	3.512	2.551	3.040	3.993
24	2.225	2.651	3.483	2.522	3.004	3.947
25	2.208	2.631	3.457	2.494	2.972	3.904
26	2.193	2.612	3.432	2.469	2.941	3.865
27	2.178	2.595	3.409	2.446	2.914	3.828
30	2.140	2.549	3.350	2.385	2.841	3.733
35	2.090	2.490	3.272	2.306	2.748	3.611
40	2.052	2.445	3.213	2.247	2.677	3.518
45	2.021	2.408	3.165	2.200	2.621	3.444
50	1.996	2.379	3.126	2.162	2.576	3.385
55	1.976	2.354	3.094	2.130	2.538	3.335
60	1.958	2.333	3.066	2.103	2.506	3.293
65	1.943	2.315	3.042	2.080	2.478	3.257
70	1.929	2.299	3.021	2.060	2.454	3.225
75	1.917	2.285	3.002	2.042	2.433	3.197
80	1.907	2.272	2.986	2.026	2.414	3.173
85	1.897	2.261	2.971	2.012	2.397	3.150
90	1.889	2.251	2.958	1.999	2.382	3.130
95	1.881	2.241	2.945	1.987	2.368	3.112
100	1.874	2.233	2.934	1.977	2.355	3.096
110	1.861	2.218	2.915	1.958	2.333	3.066
120	1.850	2.205	2.898	1.942	2.314	3.041
130	1.841	2.194	2.883	1.928	2.298	3.019
140	1.833	2.184	2.870	1.916	2.283	3.000
150	1.825	2.175	2.859	1.905	2.270	2.983
160	1.819	2.167	2.848	1.896	2.259	2.968
170	1.813	2.160	2.839	1.887	2.248	2.955
180	1.808	2.154	2.831	1.879	2.239	2.942
190	1.803	2.148	2.823	1.872	2.230	2.931
200	1.798	2.143	2.816	1.865	2.222	2.921
250	1.780	2.121	2.788	1.839	2.191	2.880
300	1.767	2.106	2.767	1.820	2.169	2.850
400	1.749	2.084	2.739	1.794	2.138	2.809
500	1.737	2.070	2.721	1.777	2.117	2.783
600	1.729	2.060	2.701	1.764	2.102	2.763
700	1.722	2.052	2.697	1.755	2.091	2.748
800	1.717	2.046	2.688	1.747	2.082	2.736
900	1.712	2.040	2.682	1.741	2.075	2.726
1000	1.709	2.036	2.676	1.736	2.068	2.718
∞	1.645	1.960	2.576	1.645	1.960	2.576

Nota: Esta tabla ha sido tomada de Remington, Schork (1974).

Soluciones para Algunas de las Cuestiones ¿Verdadero o Falso?

TEMA 0:

VF0.3: V VF0.6: V VF0.9: V VF0.12: F

TEMA 1:

VF1.3: V VF1.6: V VF1.9: V VF1.12: V VF1.15: V
VF1.18: F

TEMA 2:

VF2.3: V VF2.6: V VF2.9: F VF2.12: V VF2.15: F
VF2.18: F VF2.21: V VF2.24: F VF2.27: V

TEMAS 3 y 4:

VF3y4.3: V VF3y4.6: V VF3y4.9: F VF3y4.12: F VF3y4.15: V
VF3y4.18: V VF3y4.21: V VF3y4.24: V VF3y4.27: V

TEMA 5:

VF5.3: F VF5.6: V

TEMA 6:

VF6.3: V VF6.6: V VF6.9: V VF6.12: F VF6.15: F
VF6.18: F

Soluciones para Algunos de los Problemas

TEMA 0: Prob0.3: a) 35; b) 210; c) 105; d) 210.

TEMA 1: Prob1.3: a) Sí; b) 0.27. Prob1.6: $P(E|T) = 0,126$; $P(E^c|T^c) = 0,971$

TEMA 2: Prob2.3: 0.417. Prob2.6: No. Prob2.9: a) 0.75; b) 0.683. Prob2.12:
a) $P(\{k\}) = 1/6$, $0 \leq k \leq 5$ y $P(\{1, 3, 5\}) = 0,5$; b) $P(\{0\}) = 0,031$, $P(\{1\}) = 0,156$,
 $P(\{2\}) = 0,313$, $P(\{3\}) = 0,313$, $P(\{4\}) = 0,156$, $P(\{5\}) = 0,031$ y $P(\{1, 3, 5\}) = 0,5$;
b) $P(\{0\}) = 0,590$, $P(\{1\}) = 0,328$, $P(\{2\}) = 0,073$, $P(\{3\}) = 0,008$, $P(\{4\}) = 0,000$,
 $P(\{5\}) = 0,000$ y $P(\{1, 3, 5\}) = 0,336$. Prob2.15: a) 0.9826; b) 0.895.

TEMAS 3 y 4: Prob3y4.3: Comparación de dos proporciones. Unilateral. Muestras independientes. Se rechaza la hipótesis de igualdad de las 2 proporciones comparadas. Los datos avalan la hipótesis formulada. Prob3y4.6: Test de McNemar. Hay diferencias estadísticamente muy significativas entre ambos tratamientos. Prob3y4.9: a) Sí; b) Sí; c) No; d) $[-31,35, -19]$; e) No

TEMA 5: Prob5.3: a) Sí; b) Sí

TEMA 6: Prob6.3: a) Sí; b) Estimación puntual: $\widehat{OR} = 2,18$, Intervalo de confianza: $[1'28, 3'72]$. Prob6.6: $P < 0,001$. $C = 0,307$.

Índice alfabético

A

afijación, 104
afijación óptima, 104
afijación proporcional, 104
afijación uniforme, 104
aleatoriedad, 43
aleatorización, 174
análisis de la varianza de una vía, 203
análisis de la varianza para el modelo
de regresión lineal, 241
análisis exploratorio de datos, 29
apareamiento artificial, 174
apareamiento natural, 174
aplicación biyectiva, 19
aplicación entre dos conjuntos, 19
aproximación del muestreo sin reempla-
zamiento por el muestreo con
reemplazamiento, 100
asimetría, 294
asimetría, coeficiente de, 31

B

Bartlett, test, 210

Bartlett, test de comparación de varias vari-
anzas, 210

Bayes, teorema de, 54

Bernoulli, distribución, 67, 72

bilateral, intervalo de confianza, 144

binomial, distribución, 70, 72, 99

bioestadística, 14

bondad de ajuste, 171

Bonferroni, método de comparaciones múlti-
ples, 212

C

χ^2 , distribución de Pearson, 95, 166, 254

χ^2 , test de comparación de varias propor-
ciones, 222

χ^2 , test de independencia de dos variables
cualitativas, 254

χ^2 . test de bondad de ajuste, 171

cálculo de probabilidades e inferencia estadís-
tica, 131

campana de Gauss, 81

cantidades esperadas en una tabla de contin-
gencias, 221

- centro de gravedad, 23, 74, 86
- Chebyshev, desigualdad, 86
- Cochran, test C , 210
- Cochran, test de comparación de varias proporciones relacionadas, 225
- cociente de dos varianzas, intervalo de confianza para el, 176
- coeficiente ϕ en tablas 2×2 , 258
- coeficiente de asimetría, 31, 294
- coeficiente de contingencia de Pearson, 258
- coeficiente de correlación de Spearman, 251
- coeficiente de correlación entre dos v.a., 234, 248
- coeficiente de correlación muestral o de Pearson, 248
- coeficiente de correlación poblacional, 248
- coeficiente de curtosis, 31, 294
- coeficiente de determinación en análisis de la varianza, 207, 209
- coeficiente de determinación muestral, 247
- coeficiente de variación, 35
- coeficientes de regresión, 233
- coeficientes de regresión muestrales, 235
- combinatoria, 20
- comparación de varias muestras cualitativas, 220
- comparación de dos proporciones, aproximación normal en el caso de dos muestras independientes, 187
- comparación de dos varianzas, 176
- comparación de varias muestras independientes, 203
- comparación de varias proporciones independientes, 223
- comparación de varias proporciones relacionadas, 225
- comparación de varias proporciones, test χ^2 , 222
- comparación de varias varianzas, 210
- comparaciones múltiples en análisis de la varianza, 210
- comparaciones múltiples en el caso no paramétrico, test de Dunn, 215
- comparaciones múltiples, método de Bonferroni, 212
- comparaciones múltiples, método de Dunnett, 213
- comparaciones múltiples, método de la mínima diferencia significativa, 213
- comparaciones múltiples, método de Tukey, 210
- comparaciones múltiples, método de Tukey-Kramer, 210
- complementario de un conjunto, 18
- confianza, nivel de, 141
- conjunto producto, 18

- contraimagen de un conjunto por una aplicación, 19
- contraimagen de un elemento por una aplicación, 18
- contraste de hipótesis, 134, 146
- contraste de la hipótesis de independencia de dos variables cualitativas, 254
- contraste de la hipótesis de independencia lineal entre dos variables cuantitativas, 249
- contraste de la hipótesis de normalidad: test de D'Agostino, 167
- contraste de la independencia lineal de dos variables cuantitativas, 242
- contraste de la media de Y dado un valor de X , 244
- contraste de la varianza de una distribución normal, 166
- contraste de una proporción, aproximación normal, 169
- contraste no paramétrico de la hipótesis de independencia de dos variables cuantitativas, 252
- corrección del test de Mann-Whitney-Wilcoxon en el caso de que haya empates, 182
- corrección del test de rangos con signo de Wilcoxon en caso de empates, 185
- correlación, 234, 248
- correlación no paramétrica, 251
- correlaciones espúreas, 249
- correspondencia entre conjuntos, 18
- cota inferior de confianza, 145
- cota superior de confianza, 144
- cuantil de la distribución χ^2 , 166
- cuantil de la distribución t , 352
- cuantil de un conjunto de datos, 23
- cuantil muestral, 294
- cuartiles, 24
- curtosis, 294
- curtosis, coeficiente de, 31
- ## D
- D'Agostino, test, 209, 294
- D'Agostino, test de normalidad, 167
- dato anómalo, 31
- dato anómalo extremo, 31
- datos categóricos, 290
- definición axiomática de probabilidad, 47
- densidad, función de, 80
- desigualdad de Chebyshev, 86
- desviación media de un conjunto de datos, 35
- desviación típica de un conjunto de datos, 24
- desviación típica de una distribución continua, 81
- desviación típica de una distribución discreta, 74
- determinismo, 43

- diagrama de barras, 28
 diagrama de caja, 29
 diagrama de sectores, 34
 diagrama de tallos y hojas, 32
 diagrama de Venn, 19
 diferencia de conjuntos, 18
 diferencia de riesgos, 259
 diseño de bloques al azar, 217
 diseño de experimentos, 13, 291
 diseño de muestras independientes, 173
 distribución binomial, 70, 72
 distribución binomial: aproximación normal, 99
 distribución χ^2 de Pearson, 95, 166, 254
 distribución de Bernoulli, 67, 72
 distribución de probabilidad de una variable aleatoria, 66
 distribución F de Fisher, 95, 175, 205
 distribución hipergeométrica, 100
 distribución normal, 81
 distribución normal bidimensional, 234
 distribución t de Student, 95
 distribución T de Wilcoxon para muestras relacionadas, 185
 distribución uniforme en un conjunto finito de puntos, 72
 Dunn, test de comparaciones múltiples en el caso no paramétrico, 215
 Dunnet, método de comparaciones múltiples, 213
 efecto límite central, 101
 efecto placebo, 174
 equiprobabilidad, 49
 error de tipo I, 150, 154
 error de tipo II, 150, 155
 errores de diagnóstico, 57
 espacio de probabilidad, 49
 espacio muestral de un experimento aleatorio, 43
 especificidad de un test de diagnóstico, 55
 estadística aplicada, 14
 estadística descriptiva, 13
 estadística matemática, 14
 estadístico, 132, 133
 estadístico de contraste, 149
 estadístico de rango estudentizado, 211
 estadísticos descriptivos de un conjunto de datos, 22
 estimación conjuntista, 134, 139
 estimación de la media de una distribución normal, 136
 estimación de un parámetro, 135
 estimación de una proporción, 137
 estimación por intervalos, 139
 estimación puntual, 134, 135
 estimador de la varianza común en análisis de la varianza, 206

estimador de la varianza de
 regresión, 238

estimador de un parámetro, 135

estrato, 104

estructura estadística, 132, 148

estudio prospectivo, 260

estudios caso control, 262

estudios retrospectivos, 262

F

ϕ , coeficiente en tablas 2×2 , 258

F , distribución, 95, 175, 205

F , test de comparación de dos varianzas, 176

F , test de comparación de varias medias, 205

factor de riesgo, 259

factorial de un número, 20

falso negativo, 55

falso positivo, 55

fenómeno aleatorio, 43

fenómeno determinista, 43

fórmula de Laplace, 49

frecuencia absoluta de una clase, 26

frecuencia relativa de una clase, 26

frecuencias acumuladas, 27

frecuencias, polígono de, 34

Friedman, test, 219

función de densidad de una variable
 aleatoria continua, 80

función de distribución, 181

función de probabilidad de una variable
 aleatoria discreta, 71

función entre dos conjuntos, 19

G

Galton, máquina de, 101

ganancia del negativo de un test de diagnóstico, 56

ganancia del positivo de un test de diagnóstico, 56

Gauss, campana, 81

gráfico de probabilidad normal, 294

gráficos engañosos, 34

grupo de control, 173

grupo de tratamiento, 173

H

hipergeométrica, distribución, 100

hipótesis, 146

hipótesis alternativa, 146, 149

hipótesis nula, 146, 149

histograma, 28

homocedasticidad, 243

I

imagen de un elemento por una aplicación,
 18

inclusión, 17

independencia, 293

independencia de sucesos, 52

- independencia de variables aleatorias, 68
- independencia lineal, test de, 249
- indicios de significación, 154
- individuo, 14
- inferencia estadística, 13
- inferencia no paramétrica, 132
- inferencia paramétrica, 132
- inferencia sobre la media de la variable dependiente dado un valor de la variable independiente, 243
- interpolación lineal, 97
- intersección de conjuntos, 17
- intervalo, 16
- intervalo de confianza bilateral, 144
- intervalo de confianza para el cociente de dos varianzas, 176
- intervalo de confianza para la diferencia de dos proporciones, 188
- intervalo de confianza para la diferencia de dos proporciones en el caso de muestras relacionadas, 190
- intervalo de confianza para la diferencia de las medias de dos distribuciones normales, caso de muestras independientes, 180
- intervalo de confianza para la media de Y dado un valor de X , 243
- intervalo de confianza para la media de una distribución normal de varianza conocida, 140
- intervalo de confianza para la media de una distribución normal de varianza desconocida, 165
- intervalo de confianza para la ordenada en el origen poblacional, 244
- intervalo de confianza para la pendiente de regresión, 242
- intervalo de confianza para la varianza de una distribución normal, 166
- intervalo de confianza para una proporción, aproximación normal, 169
- intervalo de confianza unilateral, 145
- intervalo de normalidad, 87, 165
- intervalo de predicción para el valor de Y dado un valor de X , 245
- intervalos de clase, 27
- intervalos de confianza y contraste de hipótesis, 158
- ## K
- Kolmogorov, 47, 139
- Kolmogorov-Smirnov, test, 294
- Kolmogorov-Smirnov-Lilliefors, test, 210
- Kruskal-Wallis, test, 214
- ## L
- Laplace, fórmula de, 49
- leptocúrtico, conjunto de datos, 32
- Levene, test, 210
- ley de los grandes números, 47

ley fuerte de los grandes números, 139

Lilliefors, test, 294

límites de normalidad, 87, 165

M

m.a.s., 90, 292

máquina de Galton, 101

marca de clase, 27

McNemar, test, 189, 225

meda, 35

media armónica, 35

media cuadrática “dentro” en análisis
de la varianza, 205

media cuadrática “entre” en análisis
de la varianza, 205

media de un conjunto de datos cuanti-
tativos, 22

media de una distribución continua, 80

media de una distribución discreta, 74

media de una distribución normal,
intervalo de confianza (caso de
varianza conocida), 140

media de una distribución normal,
estimación de la, 136

media de una distribución normal,
intervalo de confianza (caso de
varianza desconocida), 165

media geométrica, 34

media muestral, 94, 133, 136

media ponderada, 35

mediana de un conjunto de datos, 23

medidas de asociación entre variables
cualitativas, 258

mesocúrtico, conjunto de datos, 32

método de Bonferroni de comparaciones
múltiples, 212

método de Dunnet de comparaciones múlti-
ples, 213

método de la mínima diferencia significativa
de comparaciones múltiples, 213

método de Tukey de comparaciones múlti-
ples, 210

método de Tukey-Kramer de comparaciones
múltiples, 210

métodos no paramétricos, 294

mínimos cuadrados, 235

modelo de clasificación simple en análisis de
la varianza, 203

modelo de regresión lineal, 237

modelo estadístico, 132

momento de inercia, 25, 75, 86

muestra, 14

muestra, definición probabilística, 92

muestras apareadas, 173

muestras independientes, diseño de, 173

muestras relacionadas, 173

muestreo, 90

muestreo aleatorio simple, 90, 292

muestreo aleatorio simple con reempla-

muestreo aleatorio simple sin reemplazamiento, 91

muestreo aleatorio simple con reemplazamiento, 91

muestreo estratificado, 104

N

N , 16

nivel de confianza, 141

nivel de significación de un test, 150

nivel mínimo de significación de un test, 153

normal bivalente, distribución, 234

normal, distribución, 81

normalidad, 293

nube de puntos, 235

número combinatorio, 20

números aleatorios, 92

O

observación, 43

odds ratio, 260

ordenada en el origen de la recta de regresión, 233

ordenada en el origen de la recta de regresión muestral, 237

ordenada en el origen poblacional, 238

ordenada en el origen poblacional, contraste, 244

ordenada en el origen poblacional, intervalo de confianza, 244

P

P -valor, 153

parámetro, 131

partición de tablas de contingencia, 223

pendiente de regresión, 233

pendiente de regresión muestral, 235

pendiente de regresión, intervalo de confianza, 242

pendiente poblacional, 238

percentil de un conjunto de datos, 24

placebo, 174

platicúrtico, conjunto de datos, 32

población, 14

población de muestreo, 93

población objetivo, 93

polígono de frecuencias, 34

potencia de un test de hipótesis, 150

predicción, 244

prevalencia, 55

principio de los mínimos cuadrados, 235

principio de máxima verosimilitud, 64

probabilidad, 47

probabilidad condicionada, 51

probabilidad de falso negativo, 55

probabilidad de falso positivo, 55

probabilidad de significación de un test, 153

probabilidad total, teorema de la, 54

problema de regresión lineal, 233

problema general de regresión, 233

proporción muestral, 137
proporción, contraste de una (aproximación normal), 169
proporción, estimación de una, 137
proporción, intervalo de confianza para una (aproximación normal), 169

Q

quasi-varianza, 25

R

\mathbb{R} , 16
rango de un conjunto de datos, 35
rango intercuartílico de un conjunto de datos, 35
razón del producto cruzado, 260
redondeo, 17
región crítica de un test, 147, 150
región de aceptación, 147
región de aceptación de un test, 150
regresión, 233
regresión lineal, 233, 237
regresión lineal múltiple, 234
regresión no lineal, 234
relación entre la media y la mediana, 31
relación entre la media, la mediana y la desviación típica, 25
residuos en el modelo de regresión lineal, 238
riesgo relativo, 259

S

sectores, diagrama de, 34
sensibilidad de un test de diagnóstico, 55
Shapiro-Wilks, test, 210, 294
Spearman, coeficiente de correlación, 251
Student, test t de comparación de dos medias, caso de muestras relacionadas, 184
Student, test t de comparación de dos medias, muestras independientes, 178
suceso, 44
suceso elemental, 43
suceso imposible, 44
suceso seguro, 44
suma de cuadrados debida a la regresión lineal, 241
suma de cuadrados residual en el modelo de regresión lineal, 241
sumatorio, 21

T

t , distribución, 95
 t , test de Student de comparación de dos medias, caso de muestras relacionadas, 184
 t , test de Student de comparación de dos medias, muestras independientes, 178
 $t_\alpha(n-1)$, cuantil de orden $1 - \alpha/2$ de la distribución $t(n-1)$, 163

- tabla de análisis de la varianza de una vía, 205
- tabla de análisis de la varianza para el modelo de clasificación doble sin interacciones y una observación por celda, 219
- tabla de contingencias, 220, 253
- tablas de contingencia, partición de, 223
- tablas de frecuencias, 26
- tallos y hojas, diagrama de, 32
- tamaño de muestra necesario para conseguir una precisión dada en la estimación por intervalos, 144
- tamaño de muestra necesario para detectar alternativas, 156
- Tartaglia, triángulo, 101
- teorema de Bayes, 54
- teorema de la probabilidad total, 54
- teorema del límite central, 98
- teorema del límite central, uso del, 156
- test χ^2 de bondad de ajuste, 171
- test χ^2 de independencia de dos variables cualitativas, 254
- test χ^2 para la comparación de varias proporciones, 222
- test C de Cochran, 210
- test F de comparación de dos varianzas, 176
- test F para el problema de comparación de varias medias, 205
- test t de Student de comparación de dos medias, caso de muestras relacionadas, 184
- test t de Student de comparación de dos medias, muestras independientes, 178
- test altamente significativo, 154
- test de Bartlett, 210
- test de Bartlett de comparación de varias varianzas, 210
- test de Cochran, 225
- test de Dunn, comparaciones múltiples en el caso no paramétrico, 215
- test de Friedman, 219
- test de hipótesis, 146
- test de independencia lineal de dos variables cuantitativas, 242, 249
- test de Kolmogorov-Smirnov, 294
- test de Kruskal-Wallis, 214
- test de Levene, 210
- test de Lilliefors, 294
- test de Mann-Whitney-Wilcoxon, 182
- test de McNemar, 225
- test de McNemar de comparación de dos proporciones en el caso de muestras relacionadas, 189
- test de normalidad de D'Agostino, 167, 209

- test de normalidad de Kolmogorov-Smirnov-Lilliefors, 210
- test de normalidad de Shapiro-Wilks, 210
- test de rangos con signo de Wilcoxon, 185
- test de Shapiro-Wilks, 294
- test de suma de rangos de Wilcoxon, 182
- test de Welch para la comparación de dos medias, 179
- test muy significativo, 154
- test no paramétrico de comparación de dos muestras independientes, 182
- test no paramétrico de comparación de dos muestras relacionadas, 185
- test no paramétrico de independencia de dos variables cuantitativas, 252
- test no paramétrico para la comparación de varias muestras, 214
- test significativo, 154
- tests unilaterales, 154
- tipificación de la distribución normal, 82
- triángulo de Tartaglia, 101
- Tukey-Kramer, método de comparaciones múltiples, 210
- U**
- uniforme, distribución, 72
- unilateral, intervalo de confianza, 145
- unión de conjuntos, 17
- V**
- v.a., 66
- valor P , 153
- valor experimental de un test, 149
- valor predictivo negativo de un test de diagnóstico, 56
- valor predictivo positivo de un test de diagnóstico, 56
- variabilidad biológica, 15
- variable, 14
- variable aleatoria, 66
- variable aleatoria continua, 79
- variable aleatoria discreta, 71
- variable continua, 14
- variable cualitativa, 14
- variable cuantitativa, 14
- variable discreta, 14
- variación “dentro” en análisis de la varianza, 204
- variación “entre” en análisis de la varianza, 204
- variación total, 204
- varianza de regresión, estimador, 238
- varianza de un conjunto de datos, 24
- varianza de una distribución continua, 80
- varianza de una distribución discreta, 74
- varianza de una distribución normal, intervalo de confianza, 166
- varianza muestral, 94, 133, 136
- Venn, diagrama, 19

W

Welch, test, 179

Wilcoxon, test de rangos con signos de, 185

Wilcoxon, test de suma de rangos, 182

Z

z_α , cuantil de la distribución normal, 85