UDS

Mi Universidad

LIBRO

Probabilidad y estadística

Ing. En sistemas computacionales

Cuatrimestre

Segundo cuatrimestre

Periódo

Enero - abril

Marco Estratégico de Referencia

Antecedentes históricos

Nuestra Universidad tiene sus antecedentes de formación en el año de 1979 con el inicio de actividades de la normal de educadoras "Edgar Robledo Santiago", que en su momento marcó un nuevo rumbo para la educación de Comitán y del estado de Chiapas. Nuestra escuela fue fundada por el Profesor Manuel Albores Salazar con la idea de traer educación a Comitán, ya que esto representaba una forma de apoyar a muchas familias de la región para que siguieran estudiando.

En el año 1984 inicia actividades el CBTiS Moctezuma Ilhuicamina, que fue el primer bachillerato tecnológico particular del estado de Chiapas, manteniendo con esto la visión en grande de traer educación a nuestro municipio, esta institución fue creada para que la gente que trabajaba por la mañana tuviera la opción de estudiar por las tardes.

La Maestra Martha Ruth Alcázar Mellanes es la madre de los tres integrantes de la familia Albores Alcázar que se fueron integrando poco a poco a la escuela formada por su padre, el Profesor Manuel Albores Salazar; Víctor Manuel Albores Alcázar en julio de 1996 como chofer de transporte escolar, Karla Fabiola Albores Alcázar se integró en la docencia en 1998, Martha Patricia Albores Alcázar en el departamento de cobranza en 1999.

En el año 2002, Víctor Manuel Albores Alcázar formó el Grupo Educativo Albores Alcázar S.C. para darle un nuevo rumbo y sentido empresarial al negocio familiar y en el año 2004 funda la Universidad Del Sureste.

La formación de nuestra Universidad se da principalmente porque en Comitán y en toda la región no existía una verdadera oferta Educativa, por lo que se veía urgente la creación de una institución de Educación superior, pero que estuviera a la altura de las exigencias de los jóvenes que tenían intención de seguir estudiando o de los profesionistas para seguir preparándose a través de estudios de posgrado.

Nuestra Universidad inició sus actividades el 18 de agosto del 2004 en las instalaciones de la 4ª avenida oriente sur no. 24, con la licenciatura en Puericultura, contando con dos grupos de cuarenta alumnos cada uno. En el año 2005 nos trasladamos a nuestras propias instalaciones en la carretera Comitán – Tzimol km. 57 donde actualmente se encuentra el campus Comitán y el corporativo UDS, este último, es el encargado de estandarizar y controlar todos los procesos operativos y educativos de los diferentes campus, así como de crear los diferentes planes estratégicos de expansión de la marca.

Misión

Satisfacer la necesidad de Educación que promueva el espíritu emprendedor, aplicando altos estándares de calidad académica, que propicien el desarrollo de nuestros alumnos, Profesores, colaboradores y la sociedad, a través de la incorporación de tecnologías en el proceso de enseñanza-aprendizaje.

Visión

Ser la mejor oferta académica en cada región de influencia, y a través de nuestra plataforma virtual tener una cobertura global, con un crecimiento sostenible y las ofertas académicas innovadoras con pertinencia para la sociedad.

Valores

- Disciplina
- Honestidad
- Equidad
- Libertad

Escudo

El escudo del Grupo Educativo Albores Alcázar S.C. está constituido por tres líneas curvas que nacen de izquierda a derecha formando los escalones al éxito. En la parte superior está situado un cuadro motivo de la abstracción de la forma de un libro abierto.

Eslogan

“Mi Universidad”

ALBORES

Es nuestra mascota, un Jaguar. Su piel es negra y se distingue por ser líder, trabaja en equipo y obtiene lo que desea. El ímpetu, extremo valor y fortaleza son los rasgos que distinguen.

Objetivo de la materia:

Resolver problemas que involucren fenómenos aleatorios, aplicando los modelos probabilísticos más adecuados. Asimismo, podrá organizar y analizar la información a su alcance para la toma de decisiones.

Criterios de evaluación:

No	Concepto	Porcentaje
1	Actividades en plataforma educativa	30%
2	Actividades áulicas	20%
3	Examen	50%
4	Total	100%
5	Escala de calificación	7- 10
6	Mínima aprobatoria	7

Contenido.

UNIDAD III

TEORÍA DE LA ESTIMACIÓN

3.1 Estimación insesgada y estimación ineficiente y viceversa.

3.2 Confiabilidad. Estimación por intervalo de confianza para parámetros de la población.

3.3 Intervalo de confianza para medias y proporciones.

3.4 Intervalo para diferencias de medias y de proporciones.

3.5 Curva de ajuste, regresión y correlación.

3.6 Pruebas de hipótesis para medias, proporciones, diferencia de medias.

3.7 Diferencia de proporciones.

3.8 Errores Tipo I y tipo II de significancia, prueba de una y dos colas.

3.9 Datos bivariados

3.10 Gráficas para variables cualitativas

3.11 Gráficas de dispersión para dos variables cuantitativas

3.12 Medidas numéricas para datos cuantitativos bivariados

UNIDAD IV

PROBABILIDAD Y DISTRIBUCIONES DE PROBABILIDAD

4.1 El papel de la probabilidad en estadística

4.2 Eventos y el espacio muestral

4.3 Cálculo de probabilidades con el uso de eventos sencillos

4.4 Reglas útiles de conteo

4.5 Relaciones de evento y reglas de probabilidad

4.6 Independencia, probabilidad condicional y la regla de la multiplicación

4.7 Regla de Bayes (opcional)

4.8 Variables aleatorias discretas y sus distribuciones de probabilidad

4.9 Variables aleatorias

4.10 Distribuciones de probabilidad

4.11 La media y desviación estándar para una variable aleatoria discreta

UNIDAD I

CONCEPTOS BÁSICOS

1.1.- Definición de probabilidad, espacio de probabilidad.

El término probabilidad proviene de lo *probable*, o sea, de aquello que es más posible que ocurra, y se entiende como **el mayor o menor grado de posibilidad de que un evento aleatorio ocurra**, expresado en una cifra entre 1 (posibilidad total) y 0 (imposibilidad absoluta), o bien en porcentajes entre el 100% o el 0%, respectivamente.

Para obtener la probabilidad de un suceso, generalmente se determina la frecuencia con la que ocurre (en experimentos aleatorios bajo condiciones estables), y se procede a realizar cálculos teóricos.

Para ello se sigue lo establecido por la Teoría de la probabilidad, una rama de las matemáticas dedicada al estudio de la probabilidad. Esta disciplina **es largamente empleada por otras ciencias naturales y sociales como disciplina auxiliar**, ya que les permite manejar escenarios posibles en base a generalizaciones.

El origen de la probabilidad **reside en la necesidad del ser humano de anticiparse a los hechos**, y de predecir en cierta medida el futuro. Así, en su empeño por percibir patrones y conexiones en la realidad, se enfrentó constantemente al azar, o sea, a lo que carece de orden.

Las primeras consideraciones formales sobre esta materia provienen del siglo XVII, específicamente de la correspondencia entre Pierre de Fermat y Blaise Pascal en 1654, o de los estudios de Christiaan Huygens en 1657 y de la *Kybeia* de Juan Caramuel en 1649, texto hoy en día perdido.

Puede servirte: Polígono de frecuencias.

Tipos de probabilidad

Existen los siguientes tipos de probabilidad:

- **Frecuencial.** Aquella que determina la cantidad de veces que un fenómeno puede ocurrir, considerando un número determinado de oportunidades, a través de la experimentación.
- **Matemática.** Pertenece al ámbito de la aritmética, y aspira al cálculo en cifras de la probabilidad de que determinados eventos aleatorios tengan lugar, a partir de la lógica formal y no de su experimentación.
- **Binomial.** Aquella en la que se estudia el éxito o fracaso de un evento, o cualquier otro tipo de escenario probable que tenga dos posibles resultados únicamente.
- **Objetiva.** Se denomina así a toda probabilidad en la que conocemos de antemano la frecuencia de un evento, y simplemente se dan a conocer los casos probables de que ocurra dicho evento.
- **Subjetiva.** Contrapuesta a la matemática, se sustenta en ciertas eventualidades que permiten inferir la probabilidad de un evento, aunque alejada de una probabilidad certera o calculable. De allí su subjetividad.
- **Hipergeométrica.** Aquella que se obtiene gracias a técnicas de muestreo, creando grupos de eventos según su aparición.
- **Lógica.** La que posee como rasgo característico que establece la posibilidad de ocurrencia de un hecho a partir de las leyes de la lógica inductiva.
- **Condicionada.** Aquella que se emplea para comprender la causalidad entre dos hechos distintos, cuando puede determinarse la ocurrencia de uno tras la ocurrencia del otro.

1.2.- Teoremas básicos de probabilidad

La probabilidad indica el grado de certidumbre o certeza de un suceso o fenómeno estudiado, en la investigación científica existen muchos fenómenos en los cuales es necesario determinar la probabilidad de que un evento(s) ocurra(n) o dejen de ocurrir, para lo cual el estudio de este campo, es necesario, además tiene aplicaciones muy importantes en investigación; dado que es base para la inferencia estadística que permite el estudio de muestras con el objetivo de inferir o extrapolar características de estas a una población.

(Hanke John E & Arthur Reitsch, 1997)

AXIOMAS

Un axioma es el elemento básico de un sistema de lógica formal y junto con las reglas de inferencia definen un sistema deductivo.

La palabra viene del griego αξιοειν (axioein) que significa "valorar", que a su vez procede de αξιος (axios) que significa "valuable" o "digno". Entre los antiguos filósofos griegos, un axioma era aquello que parecía ser verdadero sin ninguna necesidad de prueba.

La lógica del axioma es partir de una premisa calificada verdadera por sí misma (el axioma) e inferir sobre esta otras proposiciones por medio del método deductivo, obteniendo conclusiones coherentes con el axioma. Los axiomas han de cumplir sólo un requisito: de ellos, y sólo de ellos, han de deducirse todas las demás proposiciones de la teoría dada. Los axiomas de probabilidad son las condiciones mínimas que deben verificarse para que una función que definimos sobre unos sucesos determine consistentemente valores de probabilidad sobre dichos sucesos.

Los axiomas de la formulación moderna de la teoría de la probabilidad constituyen una base para deducir a partir de ellas un amplio número de resultados.

La letra **P** se utiliza para designar la probabilidad de un evento, siendo **P(A)** la probabilidad de ocurrencia de un evento **A** en un experimento.

AXIOMA 1: Si **A** es un evento de **S**, entonces la probabilidad del evento **A** es:

$$0 \leq P(A) \leq 1$$

Figura 2.1.1 Probabilidad De Evento A

Como no podemos obtener menos de cero éxitos ni más de **n** éxitos en **n** experimentos, la probabilidad de cualquier evento **A**, se representa mediante un valor que puede variar de **0** a **1**.

AXIOMA 2: Si dos eventos son mutuamente excluyentes, la probabilidad de obtener **A** o **B** es igual a la probabilidad de obtener **A** más la probabilidad de obtener **B**.

$$\mathbf{P(A \cup B) = P(A) + P(B)}$$

Figura 2.1.2 Eventos Mutuamente Excluyentes

Excluirse mutuamente quiere decir que **A** y **B** no pueden ocurrir simultáneamente en el mismo experimento. Así, la probabilidad de obtener águila o sol en la misma tirada de una moneda será:

$$\mathbf{P(A \cup B) = P(A) + P(B)}$$

$$\mathbf{P(A \cup B)} = 1/2 + 1/2 \ = 1.$$

Figura 2.1.3 Propiedades De La Exclusión

En general podemos decir que la suma de las probabilidades de todos los posibles eventos mutuamente excluyentes es igual a 1:

$$P(A_1) + P(A_2) + P(A_3) + \ldots + P(A_n) = 1$$

AXIOMA 3: Si **A** es un evento cualquiera de un experimento aleatorio y **A'** es el complemento de **A**, entonces:

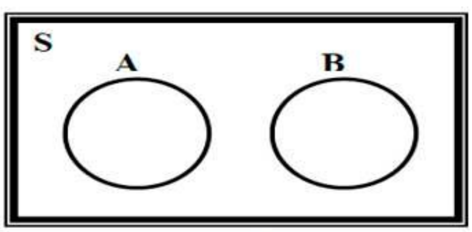
$$P(A') = 1 - P(A)$$

Figura 2.1.5 Propiedades Del Axioma 3

Es decir, la probabilidad de que el evento **A** no ocurra, es igual a 1 menos la probabilidad de que ocurra.

TEOREMAS

Suponiendo que **P(A)** y **P(B)** representan las probabilidades para los dos eventos **A** y **B**, entonces **P(A U B)** significa la probabilidad de que ocurran **A** o **B**. Si representamos los eventos **A** y **B** en un Diagrama de Venn con:



$$A \cap B = \varnothing$$

Figura 2.1.6 Representación de Los Eventos A y B en un Diagrama de Venn

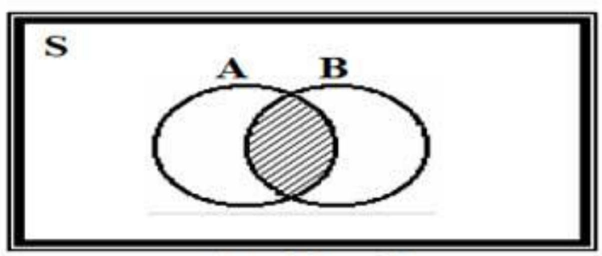
Entonces **A** y **B** son conjuntos disjuntos o mutuamente excluyentes, o sea que no pueden ocurrir en forma simultánea

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

Figura 2.1.7 Fórmula Para La Union De A y B

En cambio, si ambos eventos tienen puntos muestrales en común

$$\mathbf{A} \cap \mathbf{B} \neq \varnothing$$



$$\mathbf{A} \cap \mathbf{B} \neq \varnothing$$

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

Figura 2.1.8 Representación De la Interacción Entre A y B

1.3.-Teorema de adición.

Se puede generalizar para situaciones en las que los eventos pueden no ser necesariamente mutuamente excluyentes. Para cualquier par de eventos *A* y *B* , la probabilidad de que *A* o *B* es la suma de la probabilidad de que *A* y la probabilidad de *B* menos la probabilidad compartida de ambos *A* y *B* :

P (*A* o *B*) = *P* (*A*) + *P* (*B*) - *P* (*A* y *B*)

A veces, la palabra "y" se sustituye por ∩, que es el símbolo de la teoría de conjuntos que indica la intersección de dos conjuntos .

La regla de la suma de eventos mutuamente excluyentes es realmente un caso especial de la regla generalizada. Esto es porque si *A* y *B* son mutuamente excluyentes, entonces la probabilidad de ambos *A* y *B* es cero.

Ejemplo 1

Veremos ejemplos de cómo utilizar estas reglas de suma. Supóngase que se roba una carta de un pozo barajado- baraja de cartas . Queremos determinar la probabilidad de que la carta robada es una de dos o una tarjeta de la cara. El evento "una tarjeta de la cara se dibuja" es mutuamente excluyente con el evento "dos se dibuja," por lo que sólo tendrá que añadir las probabilidades de estos dos eventos juntos.

Hay un total de 12 tarjetas de la cara, y así la probabilidad de sacar una tarjeta de la cara es $^{12}/_{52}$. Hay cuatro grupos de dos en la cubierta, y por lo que la probabilidad de sacar un dos es $^{4}/_{52}$. Esto significa que la probabilidad de sacar una o una tarjeta de dos cara es $^{12}/_{52} + {}^{4}/_{52} = {}^{16}/_{52}$.

Ejemplo # 2

Ahora supongamos que robar una carta de una baraja bien barajado de cartas. Ahora queremos determinar la probabilidad de sacar una tarjeta roja o un as. En este caso, los dos eventos no son mutuamente excluyentes. El as de corazones y el as de diamantes son elementos del conjunto de tarjetas rojas y el conjunto de ases.

Consideramos tres probabilidades y luego combinarlas utilizando la regla de la suma generalizada:

- La probabilidad de sacar una tarjeta roja es $^{26}/_{52}$
- La probabilidad de sacar un as es $^{4}/_{52}$
- La probabilidad de sacar una tarjeta de color rojo y un as es $^{2}/_{52}$

Esto significa que la probabilidad de sacar una tarjeta de color rojo o un as es $^{26}/_{52} + {}^{4}/_{52} - {}^{2}/_{52} = {}^{28}/_{52}$.

1.4.- Teorema de probabilidad total.

Sean los sucesos A1,A2,…,AnA1,A2,…,An una partición del espacio muestral ΩΩ, y sea BB un suceso cualquiera.

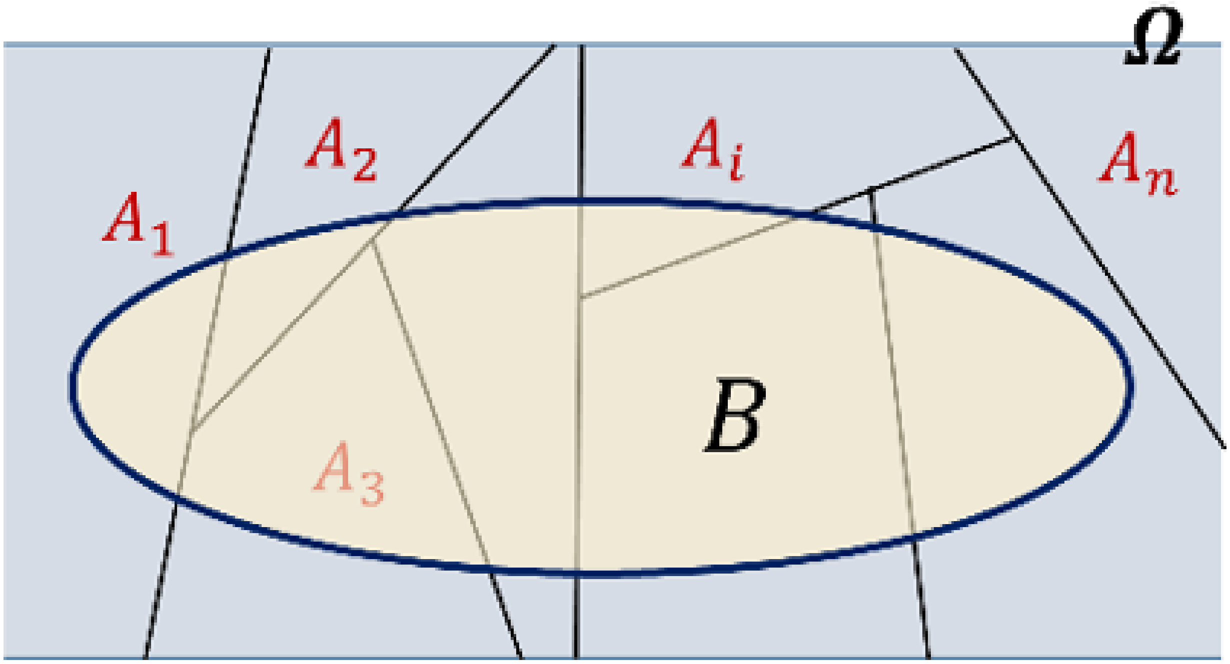


Figura 4.30: Partición: sucesos disjuntos que recubren el total.

Se cumple que la probabilidad del suceso BB puede expresarse en función de los sucesos AiAi de la siguiente manera (fórmula de la probabilidad total):$P(B)=P(B|A1)\cdot P(A1)+P(B|A2)\cdot P(A2)+\cdots+P(B|An)\cdot P(An)P(B)=P(B|A1)\cdot P(A1)+P(B|A2)\cdot P(A2)+\cdots+P(B|An)\cdot P(An)$

Ejemplo 4.1 La tienda online favorita del 33 por ciento de los socios de un foro es Pccomponendas, un 8 por ciento prefiere Medianamart, el 2 por ciento prefiere Evoy y el resto prefieren comprar en Amazonas. La probabilidad de que el pedido se pierda y no llegue al destinatario, según la casa que lo envíe, es 0.8, 0.9, 0.7 y 0.6 respectivamente. Pepe, forero del 2003, ha pedido el último modelo de linterna. Si acaba de entrar al foro y el primer hilo que abre ya es para insultar, ¿crees que está cabreado porque no ha recibido el envío?

Solución

Definimos los sucesos: A1A1="pedido a Pccomponendas",

A2A2=" pedido a Medianamart",

A3A3=" pedido a Evoy" ,

A4A4=" pedido a Amazonas",

y BB="el pedido no llega a tiempo".

Se tiene que

P(A1)=0.33,P(A2)=0.08,P(A3)=0.02,P(A4)=0.57.P(A1)=0.33,P(A2)=0.08,P(A3)=0.02,P(A4)=0.57.

Como vemos, los sucesos A1,A2,A3A1,A2,A3 y A4A4 son incompatibles y sus probabilidades suman 1, por lo que cumplen las hipótesis del teorema de las probabilidades totales.

Nos dicen, además, que P(B|A1)=0.8,P(B|A2)=0.9,P(B|A3)=0.7,P(B|A4)=0.6.P(B|A1)=0.8,P(B|A2)=0.9,P(B|A3)=0.7,P(B|A4)=0.6.

Por el teorema de las probabilidades totales, la probabilidad de que el pedido no se ha recibido

es P(B)=P(B)=P(B|A1)·P(A1)+P(B|A2)·P(A2)+P(B|A3)·P(A3)+P(B|A4)·P(A4)=P(B|A1)·P(A1)+P(B|A2)·P(A2)+P(B|A3)·P(A3)+P(B|A4)·P(A4)==0.8·0.33+0.9·0.08+0.7·0.02+0.6·0.57=0.692.=0.8·0.33+0.9·0.08+0.7·0.02+0.6·0.57=0.692.Vemos que la probabilidad de que Pepe no haya recibido el envío es más alta qué la probabilidad de que sí lo haya recibido, por lo tanto comprendemos su enfado

1.5.- Teorema de Bayers.

El teorema de Bayes, nos permite calcular la probabilidad de que ocurra un evento, a partir de valores conocidos de otras probabilidades relacionadas al evento.

El teorema de Bayes lo encontramos de dos formas diferentes, en su forma simple y en su forma extendida.

Forma simple del teorema de Bayes:

Teorema de Bayes

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

Dónde:

- A y B son eventos, y además: $P(B) \neq 0$.
- P(A|B): es la probabilidad de que ocurra A, dado que ha ocurrido B.
- P(B|A): es la probabilidad de que ocurra B, dado que ha ocurrido A.
- P(A): es la probabilidad de que ocurra A.
- P(B): es la probabilidad de que ocurra B.

El teorema expresa la probabilidad de que ocurra el evento A, dado que ha ocurrido B, en función de la probabilidad de que ocurra B dado que ha ocurrido A, de la probabilidad de A y de la probabilidad de B.

En la práctica tiene muchísimas aplicaciones, por ejemplo, conociendo la probabilidad de que una persona tenga fiebre dado que tiene gripe, nos permite calcular la probabilidad de que una persona que tiene gripe, dado que tiene fiebre. Tiene, además, aplicaciones importantísimas en la detección del cáncer y otras enfermedades.

Ejemplo 1:

En la academia de MateMovil, la probabilidad de que a un alumno seleccionado al azar le guste el helado es del 60 %, mientras que la probabilidad de que a un alumno le guste la torta es del 36 %. Además, se sabe que la probabilidad de que a un alumno le guste la torta dado que le gusta el helado es del 40 %. Calcular la probabilidad de que a un alumno le guste el helado, dado que le gusta la torta.

Solución:

Primero definimos los 2 eventos con los que vamos a trabajar:

- h: que a un alumno le guste el helado.
- t: que a un alumno le guste la torta.

Tenemos los siguientes datos:

- P(h) = 0,6.
- P(t) = 0,36.
- P(t|h) = 0,4.

Nos piden calcular P(h|t).

Aplicamos el teorema de Bayes:

$$P(h|t) = \frac{P(h) \cdot P(t|h)}{P(t)}$$

$$P(h|t) = \frac{0{,}6 \cdot 0{,}4}{0{,}36} = \frac{0{,}24}{0{,}36} = \frac{\cancel{24}}{\cancel{36}} = \frac{\mathbf{2}}{\mathbf{3}} = \mathbf{0{,}6667} = \mathbf{66{,}67}\ \%$$

Entonces, la probabilidad de que un alumno le guste el helado dado que le gusta la torta es de **0,6667 o 66,67** %.

Forma extendida del teorema de Bayes:

Teorema de Bayes

Sean $A_1, A_2, A_3, \dots, A_n$ eventos mutuamente excluyentes y exhaustivos, con $P(A_i) \neq 0$ para cada A_i. Sea B cualquier evento con $P(B) \neq 0$, entonces:

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

Esta forma extendida es la que encontrarás en la mayoría de libros de estadística. Emplea las particiones del espacio maestral.

1.6.- Cálculo de probabilidades.

Con ellos vamos a dar una serie de conceptos para poder desarrollar este tema y los sucesivos. o Fenómeno determinístico. - Cuando al repetirlo bajo idénticas condiciones iniciales se obtienen siempre los mismos resultados. o Fenómeno aleatorio. - Cuando al repetirlo bajo idénticas condiciones iniciales no se obtienen siempre los mismos resultados. Ejemplo: cuando lanzamos una moneda al aire observando la sucesión de caras y cruces que presentan. o Experimento aleatorio. - Operación que repetimos bajo idénticas condiciones iniciales y no se obtienen siempre los mismos resultados. Ejemplo: lanzamiento de un dado observando la sucesión de números que se presentan {1, 2, 3, 4, 5,6}. o Suceso elemental. - Cada uno de los resultados posibles del experimento aleatorio; luego un suceso elemental consta de un solo elemento del espacio maestral (E). En el ejemplo del dado: {1}. Suceso A= (2, 3,4) Suceso elem B= 1 ental 2 3 4 1 o Espacio maestral. - Conjunto de todos los sucesos elementales del experimento aleatorio y lo designaremos como (E). Ejemplo del dado: {1,2,3,4,5,6} o Suceso. - Conjunto formado por uno o más sucesos elementales, es decir, un subconjunto de resultados elementales del experimento aleatorio. Ejemplo del dado: nos interesa saber si el resultado a sido un número impar A={1, 3,5}. o Suceso seguro. - Coincide con el suceso elemental, ya que al realizar el experimento aleatorio se obtendrá con seguridad uno de los posibles resultados o sucesos elementales, y por tanto ocurrirá (E). o Dos sucesos se dice que son iguales, cuando todo suceso elemental de uno está en el otro, y viceversa. o Suceso imposible. - Es el que no tiene ningún elemento del espacio maestral (E), y por tanto no ocurrirá nunca, y se representa como ∅. Ejemplo: En el lanzamiento del dado no puede darse el 7. Introducción al cálculo de probabilidades 5 o Suceso complementario a un suceso A: Es el suceso que se verifica si, como resultado del experimento aleatorio, no se verifica A. Se acostumbra a denotar con el símbolo Ā. o Sucesos incompatibles: Los sucesos A y B son incompatibles o mutuamente excluyentes si no pueden ocurrir simultáneamente. A = {a, b}, B = {d, e} E A B a b c d e o Si tenemos dos sucesos cualesquiera A, B: A está contenido en B, entonces B no está contenido en A, A ⊂ B ⇒ B ⊄ A o Si tenemos dos sucesos cualesquiera A, B: donde A está contenido en B y B está contenido en A, entonces A = B. A, B / A ⊂ B ⇒ B ⊂ A ⇒ A = B 2.2.- Operaciones con sucesos Al ser los sucesos aleatorios nada más que subconjuntos de un conjunto E (espacio muestral), podemos aplicarles las conocidas operaciones con conjuntos, como son la unión, intersección y diferencia: o Suceso contenido en otro.- Un suceso A se dice que está contenido o inducido en otro B si siempre que se verifica A se verifica B. Se representa A⊂B. I_Apuntes de Estadística II 6 Ejemplo: Considerando el experimento aleatorio del lanzamiento de un dado, si designamos por: A= que aparezca el 2 ó el 4 = { 4,2 } B= que aparezca un número par: { 6,4,2 } El suceso A B, pues los resultados o sucesos elementales 2 y 4 de A, pertenecen a B. Diremos también que A implica a B y lo denotaremos A⇒B. ⊂ o Igualdad de sucesos.- Dados dos sucesos A y B, diremos que son iguales, si siempre que ocurre el suceso A también ocurre el suceso B, y siempre que ocurre el suceso B ocurre el suceso A, y lo indicaremos por A = B. Es decir, si se verifica: ⎭ ⎬ ⎫ ⊂ ⊂ AB BA ⇔ A = B

1.7.- Probabilidad condicional.

Como la probabilidad está ligada a nuestra ignorancia sobre los resultados de la experiencia, el hecho de que ocurra un suceso, puede cambiar la probabilidad de los demás. El proceso de realizar la historia clínica, explorar y realizar pruebas complementarias ilustra este principio.

La probabilidad de que ocurra el suceso A si ha ocurrido el suceso B se denomina *probabilidad condicionada* y se define

$$p(A|B) = \frac{p(A \cap B)}{p(B)} \quad \text{si } p(B) \neq 0$$

Esta definición es consistente, es decir cumple los axiomas de probabilidad.

Cuando ocurre un suceso cambia el espacio muestral, por eso cambia la probabilidad. A veces es más fácil calcular la probabilidad condicionada teniendo en cuenta este cambio de espacio muestral.

Ejemplo 3:

Una mujer es portadora de la enfermedad de Duchenne ¿Cuál es la probabilidad de que su próximo hijo tenga la enfermedad?

Según las leyes de Mendel, todos los posibles genotipos de un hijo de una madre portadora (xX) y un padre normal (XY) son xX, xY, XX, XY y tienen la misma probabilidad. El espacio muestral es W = {xX, xY, XX, XY} el suceso A={hijo enfermo} corresponde al genotipo xY, por tanto, según la definición clásica de probabilidad
p(A) = 1/4 = 0,25

La mujer tiene el hijo y es varón ¿qué probabilidad hay de que tenga la enfermedad?

Se define el suceso B = {ser varón} = {xY, XY} la probabilidad pedida es p(A|B) y aplicando la definición anterior
p(B) = 0,5; A Ç B = {xY}; p(A ÇB) = 0,25; p(A|B) = 0,25/0,5 = 0,5

Si sabemos que es varón, el espacio muestral ha cambiado, ahora es B. Por lo tanto se puede calcular p(A|B) aplicando la definición clásica de probabilidad al nuevo espacio muestral
p(A|B) = 1/2 = 0,5

Ejemplo 4:

Se sabe que el 50% de la población fuma y que el 10% fuma y es hipertensa. ¿Cuál es la probabilidad de que un fumador sea hipertenso?

A = {ser hipertenso} B = {ser fumador}
A Ç B = {ser hipertenso y fumador}
p(A|B) = 0,10/0,50 = 0,20

Obsérvese que los coeficientes falso-positivo y falso-negativo de las pruebas diagnósticas son probabilidades condicionadas.

La fórmula anterior se puede poner p(A Ç B) = p(B) p(A|B) = p(A) p(B|A) llamada *regla de la multiplicación,* que se puede generalizar a más sucesos p(A_1 Ç A_2 Ç A_3) = p((A_1 Ç A_2) Ç A_3) = p(A_1 Ç A_2) p(A_3|A_1 Ç A_2) = p(A_1) p(A_2|A_1) p(A_3|A_1 Ç A_2)

En general p(A_1 Ç A_2 Ç A_3 ...) = p(A_1) p(A_2|A_1) p(A_3|A_1 Ç A_2) ... llamado *principio de las probabilidades compuestas* y especialmente útil para aquellas situaciones en que las probabilidades condicionadas son más fáciles de obtener que las probabilidades de las intersecciones.

Ejemplo 5:

Se sabe por estudios previos que el 0,1% de la población tiene problemas vasculares. Un estudio sobre individuos con problemas vasculares revela que el 20% de ellos son placas de ateroma. Si el 10% de los individuos con placas de ateroma están expuestos a muerte súbita por desprendimiento de trombos ¿qué probabilidad tiene un individuo cualquiera de estar expuesto a muerte súbita por desprendimiento de trombos de una placa de ateroma?

A_1 = {problemas vasculares}; A_2 = {placas de ateroma}; A_3 = {expuesto a muerte súbita por ....}
p(A_1) = 0,001; p(A_2|A_1) = 0,20; p(A_3|A_1 Ç A_2) = 0,1
p(A1 Ç A2 Ç A3) = 0,001 x 0,20 x 0,1 = 0,000002

Ejemplo 6:

Una urna contiene 10 bolas, de las cuales 3 son rojas, 5 verdes y 2 azules. Se extraen al azar 3 bolas. Calcular la probabilidad de que la primera sea azul, y las otras dos verdes.

Definimos A_1 = {la 1ª bola es azul}; A_2 = {la 2ª bola es verde}; A_3 = {la 3ª bola es verde} p(A_1) = 2/10 aplicando la definición clásica de probabilidad, puesto que hay 10 bolas y 2 son verdes.

$p(A_2|A_1)$ = 5/9; si la primera bola extraída es azul, en la urna quedan 9 bolas, 5 de ellas verdes.

$p(A_3|A_1$ Ç $A_2)$ = 4/8; si la primera bola extraída es azul y la segunda verde en la urna quedan 8 bolas, 4 de ellas verdes.

p(A1 Ç A2 Ç A3) = 2/10 x 5/9 x 4/8 = 1/18

1.8.- Eventos independientes y mutuamente excluyentes.

El Puesto de Helados "Freezy" no cree tener la información suficiente para decidir si debería añadir los sabores Pan Integral y Algodón de Azúcar a su menú. Por lo tanto, lleva a cabo otro sondeo entre sus clientes. Se entera de que la probabilidad de que a un cliente le gusten ambos sabores es 0,33 y la probabilidad de que a un cliente le guste el sabor Algodón de Azúcar es 0,8. ¿Cuál es la probabilidad de que a un cliente le guste el sabor Pan Integral?

Orientación

Ya hemos abordado la fórmula para las probabilidades condicionales: $P(A|B)=P(A\cap B)P(B)$. Estos eventos no son independientes; son condicionales porque el resultado del evento B afecta el resultado del evento A . Cuando los eventos son independientes $P(A|B)=P(A)$, que significa que no importa si el evento B ha ocurrido, el resultado del evento B no afecta el resultado del evento A . no afecta el resultado del evento $P(A|B)$ con $P(A\cap B)P(B)$ en el planteamiento anterior para obtener $P(A\cap B)P(B)=P(A)$. Finalmente, multiplica ambos lados por $P(B)$ para obtener $P(A\cap B)=P(A)\times P(B)$ para los eventos independientes A y B . Podemos usar esta regla para determinar si los eventos son independientes o para encontrar la intersección de eventos independientes conocidos.

También es posible que dos eventos no tengan intersección o $P(A\cap B)=0$. Cuando esto ocurre, decimos que los eventos (o conjuntos) son Mutuamente Excluyentes. Si uno ha ocurrido, entonces el otro no puede ocurrir. Algunos ejemplos de conjuntos **Mutuamente Excluyentes** son hombres y mujeres, estudiantes de último año y estudiantes de primer año. No es posible ser ambos. Es importante observar que los eventos mutuamente excluyentes no pueden ser independientes al menos que la probabilidad de uno de los eventos sea cero, ya que los eventos independientes $P(A\cap B)=P(A)\times P(B)$ y la única forma de que un producto pueda ser igual a cero es si uno de los factores es igual a cero.

Ejemplo A

Dados dos eventos, A y B , se dice $P(A)=0.3$, $P(B)=0.5$ y $P(A\cup B)=0.65$

a. Encuentra $P(A\cap B)$.

b. Responde, con razones, si los eventos son independientes.

c. Responde, con razones, si los eventos son mutuamente excluyentes.

Solución:

a. Ya que se ha informado si los eventos son independientes, no sabemos que $P(A\cap B)=P(A)\times P(B)$. Sin embargo, para todos los eventos, independientes u otros, es cierto que $P(A)+P(B)-P(A\cap B)=P(A\cup B)$ así que

$0.3+0.5-P(A\cap B)P(A\cap B)=0.65=0.15$

b. Para determinar si los eventos son independientes, probaremos la regla $P(A\cap B)=P(A)\times P(B)$.

$P(A)*P(B)=0.3*0.5=0.15=P(A\cap B)$.

Por esto, los eventos son independientes ya que el producto de sus probabilidades es igual a la probabilidad de su intersección.

c. Los eventos no son mutuamente excluyentes porque $P(A\cap B)=0.15\neq 0$.

Ejemplo B

Dado que A y B son eventos independientes, por lo que $P(A)=0.4$ y $P(A\cup B)=0.76$, Encuentra

a. $P(B)$

b. Probabilidad de A o B pero **no** de que ambos ocurran.

Solución:

a. Ya que sabemos que los dos eventos son independientes, sabemos que $P(A\cap B)=0.4P(B)$. Ahora, podemos usas la fórmula para la probabilidad de la unión de los dos conjuntos y sustituir este producto por la probabilidad de la intersección:

$0.4+P(B)-0.4P(B)0.6P(B)P(B)=0.76=0.36=0.6$

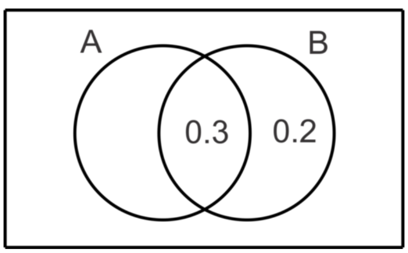
b. Para encontrar la probabilidad de que A o B ocurran, pero no ambos, necesitamos encontrar $P(A\cap B)$ y restarlo de $P(A\cup B)$.

$0.76-(0.4)(0.6)=0.76-0.24=0.52$

Ejemplo C

Los eventos A y B son independientes, por lo que $P(B\cap A')=0.2$ y $P(A\cap B)=0.3$. Encuentra $P(A\cup B)$.

Solución: Para este problema, un diagrama de Venn podría ser útil para ilustrar la información dada.



A partir del diagrama podemos ver que $P(B)=0.5$ y ya que sabemos que los eventos son independientes, sabemos que:

$P(A)*P(B)P(A)*0.5P(A)=P(A\cap B)=0.3=0.30.5=0.6$

Ahora, $P(A\cup B)=0.6+0.5-0.3=0.8$.

Revisión del Problema Introductorio Ya que sabemos que el gusto por el sabor Pan Integral (B) y el gusto por el sabor Algodón de Azúcar (A) son eventos independientes, también sabemos que $P(A\cap B)=0.33P(B)$. Ahora, podemos usar esta fórmula para la probabilidad de la intersección de los dos conjuntos para encontrar la información que estamos buscando.

$0.33=0.8P(B)P(B)=0.4125$

Por lo tanto, la probabilidad de que a un cliente le guste el sabor Algodón de Azúcar es 41,25%.

Práctica Guiada

1. Dados dos eventos, A y B , de tal forma que $P(A)=0.4$, $P(B)=0.5$ y $P(A\cup B)=0.75$

a. Encuentra $P(A\cap B)$.

b. Responde, con razones, si los eventos son independientes.

c. Encuentra $P(A|B)$.

2. Dado que A y B son eventos independientes, por lo que $P(A)=0.8$ y $P(A\cup B)=0.9$, Encuentra

a. $P(B)$

b. $P(B\cap A')$

3. Los eventos A y B son independientes, por lo que $P(A\cap B')=0.25$ y $P(A\cap B)=0.25$. Encuentra $P(A\cup B)$.

Respuestas

1. a. $0.4+0.5-P(A\cap B)P(A\cap B)=0.75=0.15$

b. Si los eventos son independientes, $P(A\cap B)=0.4\times 0.5=0.2$.

Ya que $0.2\neq 0.15$,los eventos no son independientes.

c. $P(A|B)=P(A\cap B)P(B)=0.150.5=0.3$.

2. a. $0.8+P(B)-0.8P(B)0.2P(B)P(B)=0.9=0.1=0.5$

b. $P(B\cap A')=P(B)-P(B\cap A)=0.5-0.8\times 0.5=0.1$

3. $P(A)=P(A\cap B')+P(A\cap B)=0.25+0.25=0.5$

$P(A\cap B)=0.5P(B)=0.25$ y, por lo tanto $P(B)=0.5$

$P(A\cup B)=P(A)+P(B)-P(A\cap B)=0.5+0.5-0.25=0.75$

1.9.- Teorema de la multiplicación.

Tomando como referencia la fórmula de probabilidad condicional,

$$p(A\mid E)=\frac{p(A\cap E)}{p(E)}$$

despejando,

p(AÇE) = p(E)p(A½E)   Teorema de la multiplicación para probabilidad condicional

donde:

p(AÇE) = probabilidad de que ocurran A y E

p(E) = probabilidad de que ocurra E

p(A½E) = probabilidad de que ocurra el evento A dado que el evento E ya ocurrió

Ejemplos:

1. En un lote de producción hay 25 productos, 5 de los cuales tienen defectos menores y 9 tienen defectos mayores, si se toman de este lote tres productos uno tras otro, determine la probabilidad de que: a. El primer producto no tenga defectos y que el segundo y tercero tengan defectos mayores, b. El primer producto tenga defectos menores, el segundo tenga defectos mayores y que el tercero no tenga defectos, c. El primer producto y el tercero no tengan defectos.

Solución:

a. Definiremos algunos eventos;

B_1 = evento de que el primer producto seleccionado no tenga defectos

DM_2 = evento de que el segundo producto seleccionado tenga defectos mayores

DM_3 = evento de que el tercer producto seleccionado tenga defectos mayores

p(B_1ÇDM_2ÇDM_3) = p(B_1)p(DM_2½B_1)p(DM_3½B_1ÇDM_2)

=(11/25)*(9/24)*(8/23)

$= 0.44*0.375*0.347826$

$= 0.05739$

b. Dm_1= evento de que el primer producto seleccionado tenga defectos menores

DM_2 = evento de que el segundo producto seleccionado tenga defectos mayores

B_3 = evento de que el tercer producto seleccionado no tenga defectos

$$P(Dm_1\text{Ç}DM_2\text{Ç}B_3) = p(Dm_1)p(DM_2\text{½}Dm_1)p(B_3\text{½}Dm_1\text{Ç}DM_2)$$

$$= (5/25)*(9/24)*(11/23)=$$

$$= 0.2*0.375*0.4782608= 0.03587$$

c. B_1 = evento de que el primer producto seleccionado no tenga defectos

B_2 = evento de que el segundo producto seleccionado no tenga defectos

Dm_2 = evento de que el segundo producto seleccionado tenga defectos menores

DM_2 = evento de que el segundo producto seleccionado tenga defectos mayores

B_3 = evento de que el tercer producto seleccionado no tenga defectos

En este caso como no se especifica de que tipo debe ser el segundo producto, se considera que este puede ser no defectuoso, con defectos menores o con defectos mayores; por lo tanto;

$$p(B_1\text{Ç}B_2\text{Ç}B_3) + p(B_1\text{Ç}Dm_2\text{Ç}B_3) + p(B_1\text{Ç}DM_2\text{Ç}B_3)$$

$$= p(B_1)p(B_2\text{½}B_1)p(B_3\text{½}B_1\text{Ç}B_2) + P(B_1)p(Dm_2\text{½}B_1)p(B_3\text{½}B_1\text{Ç}Dm_2) + p(B_1)p(DM_2\text{½}B_1)p(B_3\text{½}B_1\text{Ç}DM_2)$$

$$=(11/25)*(10/24)*(9/23) + (11/25)*(5/24)*(10/23) + (11/25)*(9/24)*(10/23)$$

=(0.44)(0.41666)(0.39130) + (0.44)(0.20833)(0.43478) + (0.44)(0.375)(0.43478)

= 0.07173 + 0.03985 + 0.07174

= 0.18332

2. Doce personas (6 mujeres, 4 hombres y dos niños) realizan un paseo en un pequeño autobús, al llegar a cierto lugar, bajan del autobús cuatro personas una tras otra, determine la probabilidad de que; a. La primera y segunda persona que bajen sean mujeres, el tercero sea un niño y por último baje un hombre, b. Que baje un niño, luego un hombre, luego otro niño y por último que baje una mujer, c. Que baje una mujer, luego un hombre, después otra mujer y por último otro hombre.

Solución:

a. M_1 = evento de que baje del autobús primero una mujer

M_2 = evento de que baje en segundo lugar una mujer

N_3 = evento de que baje en tercer lugar un niño

H_4 = evento de que baje en cuarto lugar un hombre

P(M_1ÇM_2ÇN_3ÇH_4) = p(M_1)p(M_2½M_1)p(N_3½M_1ÇM_2)p(H_4½M_1ÇM_2ÇN_3) =

= (6/12)*(5/11)*(2/10)*(4/9)

= 240/11,880 = 0.0202

b. N_1 = evento de que baje en primer lugar un niño

H_2 = evento de que baje en segundo lugar un hombre

N_3 = evento de que baje en tercer lugar un niño

M_4 = evento de que baje en cuarto lugar una mujer

$p(N_1\text{Ç}H_2\text{Ç}N_3\text{Ç}M_4) = p(N_1)p(H_2\text{½}N_1)p(N_3\text{½}N_1\text{Ç}H_2)p(M_4\text{½}N_1\text{Ç}H_2\text{Ç}N_3) =$

$= (2/12)*(4/11)*(1/10)*(6/9)$

$= 48/11,880$

$= 0.00404$

c. M_1 = evento de que baje en primer lugar una mujer

H_2 = evento de que baje en segundo lugar un hombre

M_3 = evento de que en tercer lugar baje una mujer

H_4 = evento de que en cuarto lugar baje un hombre

$p(M_1\text{Ç}H_2\text{Ç}M_3\text{Ç}H_4) = p(M_1)p(H_2\text{½}M_1)p(M_3\text{½}M_1\text{Ç}H_2)p(H_4\text{½}M_1\text{Ç}H_2\text{Ç}M_3)$

$= (6/12)*(4/11)*(5/10)*(3/9)$

$= 360/11,880$

$= 0.0303$

UNIDAD II

DISTRIBUCIÓN DE PROBABILIDAD

2.1.- Distribución de Bernoulli y distribución normal.

Utilice la distribución de Bernoulli cuando un proceso aleatorio tenga exactamente dos resultados: evento o no evento. Por ejemplo, en el campo de la calidad, un producto se puede clasificar como bueno o malo.

Las variables de Bernoulli pueden tomar dos valores numéricos, 0 y 1, donde 1 corresponde a un evento y 0 corresponde a un no evento. Una variable aleatoria X sigue una distribución de Bernoulli si P(X = 1) = p y P(X = 0) = 1 – p, donde p es la probabilidad de ocurrencia del evento.

La distribución de Bernoulli es una distribución discreta que está relacionada con muchas distribuciones, tales como la distribución binomial, geométrica y binomial negativa. La distribución de Bernoulli representa el resultado de 1 ensayo. Las secuencias de ensayos de Bernoulli independientes generan las demás distribuciones: la distribución binomial modela el número de éxitos en n ensayos, la distribución geométrica modela el número de fallas antes del primer éxito y la distribución binomial negativa modela el número de fallas antes del éxito xésimo.

Esta gráfica muestra una distribución binomial que tiene 1 ensayo y una probabilidad de evento de 0.15. Una distribución binomial con 1 ensayo es igual que una distribución de Bernoulli.

En teoría de probabilidad y estadística, la distribución de Bernoulli (o distribución dicotómica), nombrada así por el matemático suizo Jacob Bernoulli, es una distribución de probabilidad discreta, que toma valor 1 para la probabilidad de éxito ({\displaystyle p}p) y valor 0 para la probabilidad de fracaso ({\displaystyle q=1-p}{\displaystyle q=1-p}).

Si {\displaystyle X}X es una variable aleatoria que mide el "número de éxitos", y se realiza un único experimento con dos posibles resultados (éxito o fracaso), se dice que la variable aleatoria {\displaystyle X\,}{\displaystyle X\,} se distribuye como una Bernoulli de parámetro {\displaystyle p\,}{\displaystyle p\,}.

{\displaystyle X\sim Be(p)\,}{\displaystyle X\sim Be(p)\,}

Su función de probabilidad viene definida por:

{\displaystyle f(x)=p^{x}(1-p)^{1-x}\,\qquad {\text{ con }}\,x=\{0,1\}\,}{\displaystyle f(x)=p^{x}(1-p)^{1-x}\,\qquad {\text{ con }}\,x=\{0,1\}\,}

La fórmula será:

{\displaystyle f\left(x;p\right)=\left\{{\begin{matrix}1-p&{\mbox{si }}x=0,\\p&{\mbox{si }}x=1,\\0&{\mbox{en cualquier otro caso}}\end{matrix}}\right.}{\displaystyle f\left(x;p\right)=\left\{{\begin{matrix}1-p&{\mbox{si }}x=0,\\p&{\mbox{si }}x=1,\\0&{\mbox{en cualquier otro caso}}\end{matrix}}\right.}

Un experimento al cual se aplica la distribución de Bernoulli se conoce como Ensayo de Bernoulli o simplemente ensayo, y la serie de esos experimentos como ensayos repetidos.

Índice

Propiedades características

Valor esperado:

{\displaystyle E\left[X\right]=p=u}{\displaystyle E\left[X\right]=p=u}

Varianza:

{\displaystyle var\left[X\right]=p\left(1-p\right)=pq}{\displaystyle var\left[X\right]=p\left(1-p\right)=pq}

Función generatriz de momentos:

{\displaystyle \left(q+pe^{t}\right)}{\displaystyle \left(q+pe^{t}\right)}

Función característica:

{\displaystyle \left(q+pe^{it}\right)}{\displaystyle \left(q+pe^{it}\right)}

Moda:

0 si q > p (hay más fracasos que éxitos)

1 si q < p (hay más éxitos que fracasos)

0 y 1 si q = p (los dos valores, pues hay igual número de fracasos que de éxitos)

Asimetría (Sesgo):

{\displaystyle \gamma _{1}={\frac {q-p}{\sqrt {qp}}}}{\displaystyle \gamma _{1}={\frac {q-p}{\sqrt {qp}}}}

Curtosis:

{\displaystyle \gamma _{2}={\frac {6p^{2}-6p+1}{p(1-p)}}}{\displaystyle \gamma _{2}={\frac {6p^{2}-6p+1}{p(1-p)}}}

La Curtosis tiende a infinito para valores de {\displaystyle p}p cercanos a 0 ó a 1, pero para {\displaystyle p={\frac {1}{2}}}{\displaystyle p={\frac {1}{2}}} la distribución de Bernoulli tiene un valor de curtosis menor que el de cualquier otra distribución, igual a -2.

Caracterización por la binomial:

{\displaystyle X\sim Be(p)\Longleftrightarrow X\sim B(1,p)\,}{\displaystyle X\sim Be(p)\Longleftrightarrow X\sim B(1,p)\,} ; donde {\displaystyle B(1,p)\,}{\displaystyle B(1,p)\,} es una distribución binomial.

Distribuciones Relacionadas

Si {\displaystyle X_{1},X_{2},X_{3},\dots ,X_{n}}{\displaystyle X_{1},X_{2},X_{3},\dots ,X_{n}} son {\displaystyle n} n variables aleatorias identicamente distribuidas con la distribución de Bernoulli con la misma probabilidad de éxito {\displaystyle p}p en todas, entonces la variable aleatoria {\displaystyle X=X_{1}+X_{2}+\dots +X_{n}}{\displaystyle X=X_{1}+X_{2}+\dots +X_{n}} presenta una distribución binomial de probabilidad.

{\displaystyle X\sim B(n,p)}{\displaystyle X\sim B(n,p)}

Ejemplo

"Lanzar una moneda, probabilidad de conseguir que salga cruz".

Se trata de un solo experimento, con dos resultados posibles: el éxito (p) se considerará sacar cruz. Valdrá 0,5. El fracaso (q) que saliera cara, que vale (1 - p) = 1 - 0,5 = 0,5.

La variable aleatoria X medirá "número de cruces que salen en un lanzamiento", y sólo existirán dos resultados posibles: 0 (ninguna cruz, es decir, salir cara) y 1 (una cruz).

Por tanto, la v.a. X se distribuirá como una Bernoulli, ya que cumple todos los requisitos.

{\displaystyle X\sim Be(0.5)}{\displaystyle X\sim Be(0.5)}

{\displaystyle P(X=0)=f(0)=0,5^{0}0,5^{1}=0,5}{\displaystyle P(X=0)=f(0)=0,5^{0}0,5^{1}=0,5}

{\displaystyle P(X=1)=f(1)=0,5^{1}0,5^{0}=0,5}{\displaystyle P(X=1)=f(1)=0,5^{1}0,5^{0}=0,5}

Ejemplo:

"Lanzar un dado y salir un 6".

Cuando lanzamos un dado tenemos 6 posibles resultados:

{\displaystyle \Omega =\{1,2,3,4,5,6\}}{\displaystyle \Omega =\{1,2,3,4,5,6\}}

Estamos realizando un único experimento (lanzar el dado una sola vez).

Se considera éxito sacar un 6, por tanto, la probabilidad según la Regla de Laplace (casos favorables dividido entre casos posibles) será 1/6.

{\displaystyle p=1/6}{\displaystyle p=1/6}

Se considera éxito sacar un 6, por tanto, se considera fracaso sacar cualquier otro resultado.

{\displaystyle q=1-p=1-1/6=5/6}{\displaystyle q=1-p=1-1/6=5/6}

La variable aleatoria X medirá "número de veces que sale un 6", y solo existen dos valores posibles, 0 (que no salga 6) y 1 (que salga un 6).

Por tanto, la variable aleatoria X se distribuye como una Bernoulli de parámetro {\displaystyle p}p = 1/6

{\displaystyle X\sim Be(1/6)}{\displaystyle X\sim Be(1/6)}

La probabilidad de que obtengamos un 6 viene definida como la probabilidad de que X sea igual a 1.

{\displaystyle P(X=1)=f(1)=(1/6)^{1}*(5/6)^{0}=1/6=0.1667}{\displaystyle P(X=1)=f(1)=(1/6)^{1}*(5/6)^{0}=1/6=0.1667}

La probabilidad de que NO obtengamos un 6 viene definida como la probabilidad de que X sea igual a 0.

P(X=0)=f(0)=(1/6)^{0}*(5/6)^{1}=5/6=0.8333{

P(X=0)=f(0)=(1/6)^{0}*(5/6)^{1}=5/6=0.8333}

2.2.- Relación entre la distribución de poisson.

Características:

En este tipo de experimentos los éxitos buscados son expresados por unidad de área, tiempo, pieza, etc, etc,:

- # de defectos de una tela por m2

- # de aviones que aterrizan en un aeropuerto por día, hora, minuto, etc, etc.

- # de bacterias por cm2 de cultivo

- # de llamadas telefónicas a un conmutador por hora, minuto, etc, etc.

- # de llegadas de embarcaciones a un puerto por día, mes, etc, etc.

Para determinar la probabilidad de que ocurran x éxitos por unidad de tiempo, área, o producto, la fórmula a utilizar sería:

donde:

p(x, l) = probabilidad de que ocurran x éxitos, cuando el número promedio de ocurrencia de ellos es

l = media o promedio de éxitos por unidad de tiempo, área o producto

e = 2.718

x = variable que nos denota el número de éxitos que se desea que ocurra

Hay que hacer notar que en esta distribución el número de éxitos que ocurren por unidad de tiempo, área o producto es totalmente al azar y que cada intervalo de tiempo es independiente de otro intervalo dado, así como cada área es independiente de otra área dada y cada producto es independiente de otro producto dado.

Ejemplos:

Si un banco recibe en promedio 6 cheques sin fondo por día, ¿cuáles son las probabilidades de que reciba, a) cuatro cheques sin fondo en un día dado, b) 10 cheques sin fondos en cualquiera de dos días consecutivos?

Solución:

a) x = variable que nos define el número de cheques sin fondo que llegan al banco en un día cualquiera = 0, 1, 2, 3, ..., etc, etc.

l = 6 cheques sin fondo por día

e = 2

b)

x= variable que nos define el número de cheques sin fondo que llegan al banco en dos días consecutivos = 0, 1, 2, 3, ..., etc., etc.

l = 6 x 2 = 12 cheques sin fondo en promedio que llegan al banco en dos días consecutivos

Nota: l siempre debe de estar en función de x siempre o dicho de otra forma, debe "hablar" de lo mismo que x.

En la inspección de hojalata producida por un proceso electrolítico continuo, se identifican 0.2 imperfecciones en promedio por minuto. Determine las probabilidades de identificar a) una imperfección en 3 minutos, b) al menos dos imperfecciones en 5 minutos, c) cuando más una imperfección en 15 minutos.

Solución:

a) x = variable que nos define el número de imperfecciones en la hojalata por cada 3 minutos = 0, 1, 2, 3, ..., etc., etc.

l = 0.2 x 3 =0.6 imperfecciones en promedio por cada 3 minutos en la hojalata

b) x = variable que nos define el número de imperfecciones en la hojalata por cada 5 minutos = 0, 1, 2, 3, ..., etc., etc.

l = 0.2 x 5 =1 imperfección en promedio por cada 5 minutos en la hojalata

=1-(0.367918+0.367918) = 0.26416

c) x = variable que nos define el número de imperfecciones en la hojalata por cada 15 minutos = 0, 1, 2, 3, ..., etc., etc.

l = 0.2 x 15 = 3 imperfecciones en promedio por cada 15 minutos en la hojalata

= 0.0498026 + 0.149408 = 0.1992106

2.3.- El teorema del límite central.

En el resultado anterior, veíamos que la suma de variables aleatorias normales es otra variable aleatoria normal. Sin embargo, la normalidad de una suma de variables no se limita solo a las variables normales. El teorema central del límite es un resultado matemático que garantiza que, si sumamos variables cualesquiera (no necesariamente normales), la variable suma también seguirá una distribución normal (esto siempre que se cumplan algunas condiciones básicas).

Así, cuando un dato o resultado es la suma de contribuciones independientes, de igual magnitud y "con un tamaño típico", este resultado corresponderá a una distribución Gaussiana siempre que el número de contribuciones (el número de sumandos) sea un número considerable (no pequeño).

Con un tamaño típico se quiere garantizar que las contribuciones tienen que "estar controladas", esto es, las contribuciones extremas tienen que estar controladas por una probabilidad muy pequeña (En jerga matemática las contribuciones tiene que tener varianza finita).

Este teorema asegura, de manera esquemática, que, cuando sumamos un número grande de variables, la variable resultante sigue una distribución normal.

De manera general, si $X1,X2,...,XnX1,X2,...,Xn$ son variables de media o esperanza $\mu i=E(Xi)\mu i=E(Xi)$ y varianza $\sigma 2i=Var(Xi)$, $i=1,...,n,\sigma i2=Var(Xi)$, $i=1,...,n$, se verifica que la variable suma $Y=X1+X2+...+XnY=X1+X2+...+Xn$ (si nn es un número tendiendo a infinito) se puede aproximar por una variable normal, de media la suma de las medias y varianza la suma de varianzas (desviación típica = raiz de la suma de varianzas), es decir $Y=X1+X2+...+Xn \approx N\left(n\sum i=1\mu i,\square\square\ \square\sqrt{\boxed{n\sum i=1\sigma 2i}}\right)$. $Y=X1+X2+...+Xn \approx N(\sum i=1n\mu i,\sum i=1n\sigma i2)$.

En el caso de sumar variables aleatorias normales, la aproximación anterior no es tal, sino que es una distribución exacta, como hemos visto anteriormente.

Si, en vez de sumar variables, realizamos la media aritmética de las mismas, también podemos utilizar el teorema central del límite (puesto que la media aritmética es sumar y luego dividir por una constante).

Este teorema (del que damos únicamente una idea general, sin establecer las hipótesis matemáticas reales) establece la importancia de la distribución normal. Su resultado es que, cuando se suma un número grande de variables aleatorias, la variable resultante es

una variable con distribución aproximadamente igual a la distribución normal. Incluso, el término número grande(porque matemáticamente el teorema se establece cuando nn tiende a infinito) no lo es tanto, porque, en la práctica, con tener que nn sea un número mayor o igual a 3030, la aproximación ya proporciona buenas resultados.
Además, el teorema es cierto independientemente de la distribución que sigan las variables que se sumen (no importa si son exponenciales, binomiales, etc.). Lo único que se necesita es saber su media y su varianza.

El consumo de petroleo (gas, electricidad...) de una ciudad es la suma de los consumos individuales de las familias o particulares. Por ello la distribución de esta variable (consumo) va a seguir una distribución normal.

El tiempo de realización de un proyecto complejo (como construir una casa, un submarino, un avión, una red de carreteras, un oleoducto…) es la suma de los tiempos de las distintas tareas que componen el proyecto. A pesar de que habrá tareas que tendrán un tiempo fijo, la mayoría serán variables con diferente tiempo medio y diferente variación. Pero la suma de los tiempos seguirá una distribución normal, y se podrán calcular probabilidades de finalización en un tiempo determinado (y a su vez el coste de este tiempo).

La proporción de una característica A en una muestra sigue una distribución normal. Comprobémoslo.

La proporción muestral de una característica A es el número de veces que dicha característica AA aparece en una muestra. Por ejemplo, si AA representa tener una enfermedad cualquiera, p=P(A)p=P(A) es la probabilidad de que una persona tenga la enfermedad.
Si se seleccionan, de manera independiente, nn personas, tenemos una muestra de nn individuos de esa población, y la proporción muestral es:^p=número de individuos en la muestra con esa enfermedadnp^=número de individuos en la muestra con esa enfermedadnEn vez de tener una enfermedad, AA puede representar estar de acuerdo o

no con algo, tener trabajo o no, etc (cualquier cosa que admita solo 2 posibilidades complementarias).

Cada vez que consideramos una persona, podemos considerar la variable de Bernoulli XX=tiene la enfermedad (o característica) A. Esta variable toma los valores 11 con probabilidad pp y 00 con probabilidad 1−p1−p.

De esta manera, la proporción muestral que acabamos de definir se puede considerar como^p=X1+X2+...Xnn,p^=X1+X2+...Xnn,donde X1X1 es la variable XX en el individuo 11,…, XnXn es la variable XX en el individuo nn, es decir vale 11 o 00 en cada individuo, según tenga la característica AA o no la tenga.

De manera que, si nn es grande, por el teorema central del límite, la variable suma X1X1+X2X2+…+XnXn se aproximará mediante una distribución normal, de media la suma de las medias (cada variable de Bernoulli tiene de media pp) y de desviación típica la raiz cuadrada de la suma de varianzas (y cada variable de Bernoulli tiene de varianza p·(1−p)p·(1−p)). En consecuencia, la variable suma YY verificará:Y=X1+X2+...+Xn ≈ N(μ,σ)Y=X1+X2+...+Xn ≈ N(μ,σ)donde μ=p+p+...+p=npμ=p+p+...+p=np y σ=$\sqrt{\boxed{p·(1−p)+...+p·(1−p)}}$=$\sqrt{\boxed{np(1−p)}}$σ=p·(1−p)+...+p·(1−p)=np(1−p).

Supongamos ahora que lanzamos una moneda. La variable XX que vale 11 si sale cara y 00 si sale cruz es una variable de Bernoulli. Si lanzamos una moneda, por ejemplo, 200200 veces, la variable que mide el número de caras que salen es una suma de 200200 variables (XiXi cuenta 11 o 00 si sale cara en el lanzamiento ii).

Como vemos, se parece mucho a la campana de Gauss, con media 100100 y desviación típica $\sqrt{\boxed{200·0.5·0.5}}$=7.071.200·0.5·0.5=7.071.

Consideremos de nuevo una proporción. Según acabamos de ver, la proporción muestral es^p=X1+X2+...Xnn,p^=X1+X2+...Xnn,y, como la suma de arriba es aproximadamente una distribución normal, de parámetros media npnp y varianza np(1−p)np(1−p), la proporción muestral también sigue aproximadamente una distribución normal.^p=X1+X2+...Xnn≈N$\left(\left(npn,\sqrt{\boxed{np(1−p)n2}}\right)\right)$=N$\left(\left(p,\sqrt{\boxed{p(1−p)n}}\right)\right)$p^=X1+X2+...Xnn≈N(npn,np(1−p)n2)=N(p,p(1−p)n)

Supongamos que entrevistamos a 5050 rusos. Parece ser que la probabilidad de que un ruso no crea que el hombre llegó a la luna es 0.570.57. fuente

La proporción de gente en la muestra que dirá "NO", cuando le preguntemos si cree que el hombre llegó a la luna, será un número más o menos cercano a 0.570.57 (quizá no demasiado, puesto que 5050 son pocas personas).

Podemos simular en R el proceso de preguntar a 5050 personas hoy, mañana, pasado y así, por ejemplo, durante 2020 días:

```
y=rbinom(20,50,0.57)
y/50  # proporción en cada muestra de 50 personas
##  [1] 0.32 0.60 0.66 0.56 0.60 0.54 0.54 0.54 0.56 0.56
## [11] 0.56 0.56 0.60 0.62 0.66 0.58 0.58 0.46 0.54 0.64
```

Si en vez de repetirlo 2020 días, lo hacemos 10001000, tenemos mil valores de la proporción muestral, y, como antes, podemos dibujar su histograma:

```
y=rbinom(1000,50,0.57)/50
hist(y, col="lightblue")
abline(v=0.57, col="red")
```

Vemos, por lo tanto, que la proporción muestral toma valores cuya distribución es aproximadamente normal, de media la verdadera proporción pp y desviación típica $\sqrt{\boxed{p(1-p)/n}}$p(1−p)/n, en este caso $\sqrt{\boxed{0.57(1-0.57)/1000}}$=0.015.0.57(1−0.57)/1000=0.015.

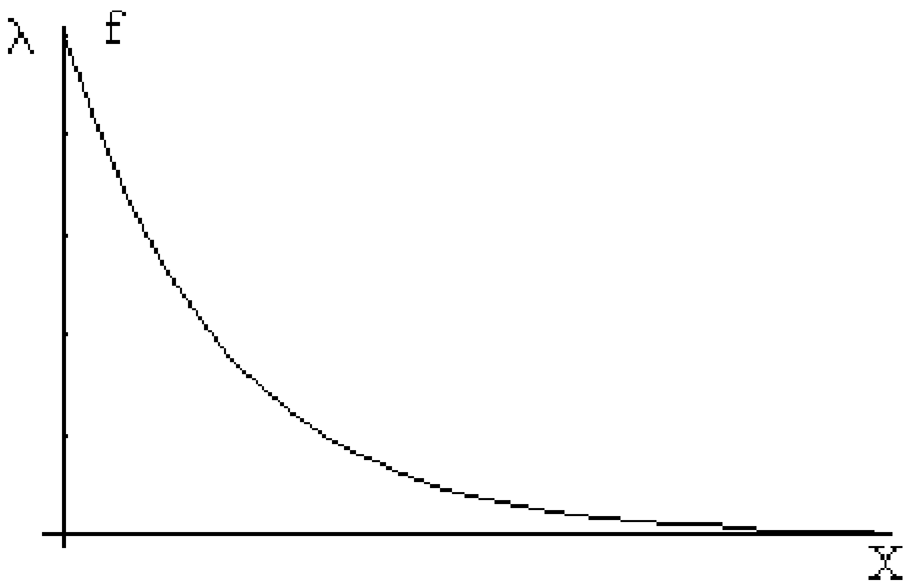
Son, precisamente, este tipo de resultados un fuerte apoyo de la teoría del muestreo, en la cual se fundamentan las encuestas de opinión o electorales. Si las muestras de la población en la que se realiza un sondeo son relativamente grandes, se puede precisar con bastante fiabilidad la opinión de una población a través de la muestra, y la variabilidad existente en esta "opinión muestral".

es tal que su función de densidad es

$$\boxed{f(x) = \lambda e^{-\lambda x} \text{si } 0 < x}$$

se dice que sigue una **distribución exponencial** de parámetro λ, $X \rightsquigarrow \mathbf{Exp}\,(\lambda)$.

Figura: Funci�n de densidad, *f*, de una $\mathbf{Exp}\,(\lambda)$.



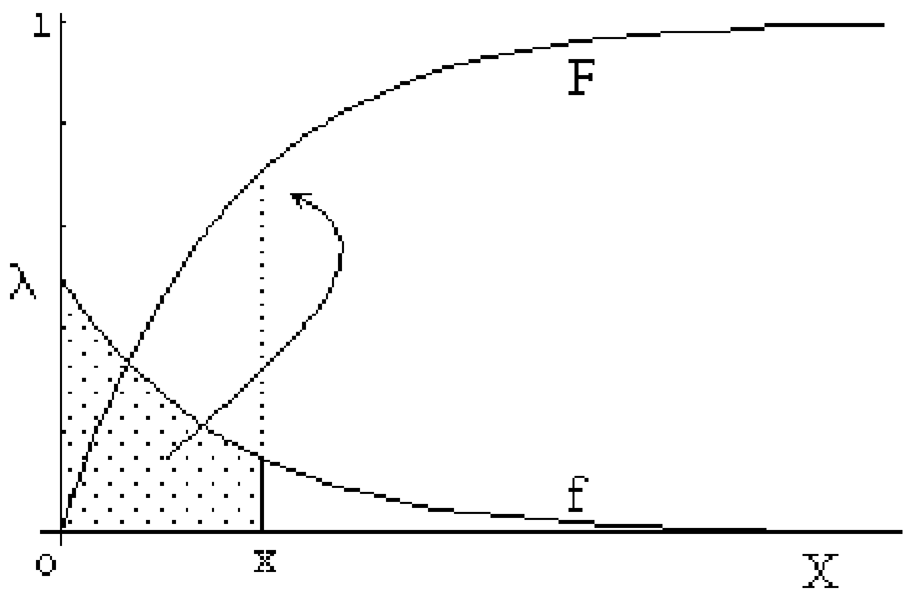
Un cálculo inmediato nos dice que si *x*>0,

$$\int_0^x \lambda e^{-\lambda t}\, dt = -e^{-\lambda t}\Big]_0^x = 1 - e^{-\lambda x}$$

luego la función de distribución es:

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{si } 0 < x \\ 0 & \text{en otro caso.} \end{cases}$$

Figura: Función de distribución



Para calcular el valor esperado y la varianza de la distribución exponencial, obtenemos en primer lugar la función característica

$$\phi_X(t) = \int_0^{+\infty} e^{itx}\lambda e^{-\lambda x}\,dx = \int_0^{+\infty} \lambda e^{(it-\lambda)x}\,dx = \frac{\lambda}{it-\lambda}e^{(it-\lambda)x}\Big]_0^{+\infty} = -\frac{\lambda}{it-\lambda}$$

para después, derivando por primera vez

$$\begin{aligned} \phi'_X(t) &= \frac{\lambda i}{(it-\lambda)^2} \\ &\Downarrow \\ \mathbf{E}[X] &= \frac{\phi'_X(0)}{i} = \frac{1}{\lambda} \end{aligned}$$

y derivando por segunda vez,

$$\begin{aligned} \phi''_X(t) &= \frac{-2\lambda i^2}{(it-\lambda)^3} \\ &\Downarrow \\ \mathbf{E}\left[X^2\right] &= \frac{\phi''_X(0)}{i^2} = \frac{-2\lambda}{-\lambda^3} = \frac{2}{\lambda^2} \end{aligned}$$

Entonces la varianza vale

$$\mathbf{Var}[X] = \mathbf{E}\left[X^2\right] - \mathbf{E}[X]^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

Ejemplo

Se ha comprobado que el tiempo de vida de cierto tipo de marcapasos sigue una distribución exponencial con media de 16 años. ¿Cuál es la probabilidad de que a una persona a la que se le ha implantado este marcapasos se le deba reimplantar otro antes de 20 años? Si el marcapasos lleva funcionando correctamente 5 años en un paciente, cuál es la probabilidad de que haya que cambiarlo antes de 25% años?

Solución: Sea *T* la variable aleatoria que mide la duración de un marcapasos en una persona. Tenemos que

$$T \rightsquigarrow \mathbf{Exp}\left(\lambda = \frac{1}{16}\right) \iff f(t) = \lambda e^{-\lambda t} \text{si} \quad \forall t \geq 0$$

$$\iff F(t) = 1 - e^{-\lambda t}$$

Entonces

$$\mathcal{P}[T \leq 20] = \int_0^{20} f(t)\,dt = F(20) = 1 - e^{-\frac{20}{16}} = 0,7135$$

En segundo lugar

$$
\begin{aligned}
\mathcal{P}[T \leq 25_{|T\geq 5}] &= \frac{\mathcal{P}[5 \leq T \leq 25]}{\mathcal{P}[T \geq 5]} = \frac{0,522}{0,7316} = 0,7135 \\
\mathcal{P}[5 \leq T \leq 25] &= \int_5^{25} f(t)\, dt = F(25) - F(5) = 1 - e^{-\frac{25}{16}} - 1 + e^{-\frac{5}{16}} = 0,522 \\
\mathcal{P}[T \geq 5] &= \int_5^{+\infty} f(t)\, dt = F(+\infty) - F(5) = 1 - 1 + e^{-\frac{5}{16}} = 0,7316
\end{aligned}
$$

Luego como era de esperar, por ser propio a un mecanismo exponencial,

$$\mathcal{P}[T \leq 25_{|T\geq 5}] = \mathcal{P}[T \leq 20]$$

2.6.- Estadística, teoría de muestreo.

Muestreo no probabilístico

Es aquel para el que no se puede calcular la probabilidad de extracción de una determinada muestra ya que no todos los sujetos tienen la misma probabilidad de ser elegidos. Por tal motivo, se busca seleccionar a individuos que tienen un conocimiento profundo del tema bajo estudio y se considera que la información aportada por esas personas es vital para la toma de decisiones.

Muestreo por cuotas

Es la técnica más difundida sobre todo en estudios de mercado y sondeos de opinión. En primer lugar es necesario dividir la población de referencia en varios estratos definidos por algunas variables de distribución conocida (como el género o la edad). Posteriormente se calcula el peso proporcional de cada estrato, es decir, la parte proporcional de población que representan. Finalmente se multiplica cada peso por el tamaño de "n" de la muestra para determinar la cuota precisa en cada estrato. Se diferencia del muestreo estratificado en que una vez determinada la cuota, el investigador es libre de elegir a los sujetos de la muestra dentro de cada estrato.

Muestreo de bola de nieve

Indicado para estudios de poblaciones clandestinas, minoritarias o muy dispersas pero en contacto entre sí. Consiste en identificar sujetos que se incluirán en la muestra a partir de los propios entrevistados. Partiendo de una pequeña cantidad de individuos que cumplen los requisitos necesarios, servirán como localizadores de otros con características análogas.

Muestreo subjetivo por decisión razonada

En este caso las unidades de la muestra se eligen en función de algunas de sus características de manera racional y no casual. Una variante de esta técnica es el [[muestreo compensado]] o [[muestreo equilibrado]], en el que se seleccionan las unidades de tal forma que la media de la muestra para determinadas variables se acerque a la media de la población, la cual funciona sobre la base de referencias o por recomendación, después se reconoce por medio de la estadística.

Muestreo aleatorio simple

Todos aquellos métodos para los que se puede calcular la probabilidad de extracción de cualquiera de las muestras posibles. Este conjunto de técnicas de muestreo es el más aconsejable, aunque en ocasiones no es posible optar por él.

2.6.- población o muestra

¿Qué es una población?

La población en investigación es un conjunto completo de elementos que poseen un parámetro común entre sí.

Es importante mencionar que todos somos conscientes de lo que la palabra "población" significa en nuestra vida cotidiana. A menudo se utiliza para describir la población humana o el número total de personas que viven en un área geográfica de algún país o estado.

La población en investigación no tiene que ser necesariamente humana. Puede ser cualquier colección de datos que posea un parámetro común, como por ejemplo el número total de tiendas de mascotas en una ciudad.

Aprende cómo medir una población con este artículo que tenemos para ti.

¿Qué es una muestra?

Una muestra es la parte más pequeña del total, es decir, un subconjunto de toda la población. Cuando se realizan encuestas, la muestra son los miembros de la población que son invitados a participar en la encuesta.

Dicho de manera sencilla, una muestra es un subgrupo o subconjunto dentro de la población, que puede ser estudiado para investigar las características o el comportamiento de los datos de población.

Las muestras de datos se crean utilizando varios métodos de investigación como el muestreo probabilístico y el muestreo no probabilístico. Los métodos de muestreo varían según los tipos de investigación y la calidad de la información requerida.

Por ejemplo: A una empresa de alimentos para gatos le gustaría conocer todas las tiendas de mascotas donde puede vender. La compañía tiene datos de población sobre el número total de tiendas de mascotas en una ciudad específica.

Este fabricante de alimentos para mascotas ahora puede crear una muestra de investigación en línea seleccionando únicamente las tiendas de mascotas que venden alimentos para gatos.

Los datos pueden ser estudiados para varias características y los resultados pueden ser mostrados en estadísticas e informes para una mejor comprensión del negocio.,

Utilizando los datos de la muestra, la empresa puede descubrir formas de hacer crecer su negocio hasta llegar a la población total de tiendas de mascotas.

EJERCICIO 2: (APLICACIÓN EN CIENCIAS DE LA INGENIERÍA)

En una fábrica de artículos electrónicos generalmente el 10% de los artículos presenta algún defecto de fabricación. Para analizar la calidad del producto se desea estimar la proporción de artículos electrónicos defectuosos de un lote de 2000 artículos listo para ser embarcado.

Pregunta: ¿Cuántos artículos deben ser elegidos del lote si se desea una confianza de 95% y un error de estimación no mayor a 0.05?

Esquema de solución

Paso 1: Leer cuidadosamente el enunciado del problema.

Paso 2: Identificar la característica en estudio y los parámetros involucrados.

Sea A = {Artículos electrónicos de la fábrica con algún defecto}.

En este caso el parámetro de interés es $p = P(A) =$ proporción de artículos electrónicos defectuosos en la fábrica.

Paso 3: Estimar los parámetros. En este caso se tiene, del enunciado del problema, que

$\hat{p} = 0.1$ (10%) es un estimador de p

Paso 4: Leer la pregunta y revisar cual de los conceptos se debe usar para obtener lo pedido. Para responder la pregunta se debe usar un diseño MAS para determinar el tamaño de muestra mínimo para estimar una proporción poblacional.

Paso 5: Determinación del tamaño de muestra.

$$n \geq n_0 = \left(\frac{z_{0.975}}{d_0}\right)^2 \hat{p}(1-\hat{p}) = \left(\frac{1.96}{0.05}\right)^2 0.1*0.9 = 138.2976 \approx 139$$

$$\frac{n_0}{N} = \frac{139}{2000} = 0.695 > 0.05 \Rightarrow n \geq n' = \frac{n_0}{1+\frac{n_0}{N}} = \frac{139}{1+0.695} = 82.0059 \approx 83$$

Paso 6: Redactar una respuesta.

Para estimar la proporción de artículos electrónicos en el lote usando un diseño MAS con una confianza de 95% y un error de estimación no mayor a 0.05 se debe elegir a lo menos 83 artículos del lote en la muestra

2.6.3.- Frecuencias, histogramas, parámetros

Tipos de frecuencia estadística

En estadística, podemos identificar 4 tipos de frecuencias: absoluta, relativa, absoluta acumulada y relativa acumulada.

Frecuencia absoluta

- Artículo principal: Frecuencia absoluta.

Se le llama frecuencia absoluta al número de veces que se repite una variable en un experimento. Esta se representa con f_i ó n_i.

Frecuencia relativa

Representa la cantidad de veces que se repite una observación, expresada como proporción de la muestra. Es decir, es el resultado de dividir el valor de la frecuencia absoluta por el tamaño de la muestra estadística.

Esta se representa con f_i y se define como $f = n/N$, siendo n el número de veces que se repite la respuesta y N el tamaño de la muestra. Su valor se expresa como porcentaje.

Frecuencia absoluta acumulada

La frecuencia acumulada es aquella que se obtiene al sumar todas las frecuencias absolutas inferiores o iguales al valor en cuestión. Se representa con N_i.

Frecuencia relativa acumulada

En esta se tiene en cuenta la sumatoria de todas las frecuencias relativas inferiores o iguales al valor en cuestión. Se representa con F_i ó H_i.

Ejemplo de frecuencia estadística

Supongamos que se realiza una investigación sobre 15 personas, para determinar la cantidad de mascotas que estas tienen.

Las respuestas analizadas son: 1, 2, 2, 3, 1, 2, 0, 1, 3, 4, 0, 2, 1, 2, 4. Por lo que la tabla de frecuencia estadística quedaría de la siguiente manera:

Mascotas	Frecuencia absoluta	Frecuencia relativa	Frecuencia absoluta acumulada	Frecuencia relativa acumulada
0	2	2/15 = 0,13	2	2/15
1	4	4/15 = 0,26	2 + 4 = 6	6/15

Mascotas	Frecuencia absoluta	Frecuencia relativa	Frecuencia absoluta acumulada	Frecuencia relativa acumulada
2	5	5/15 = 0,33	2 + 4 + 5 = 11	11/15
3	2	2/15 = 0,13	2 + 4 + 5 + 2 = 13	13/15
4	2	2/15 = 0,13	2 + 4 + 5 + 2 = 15	15/15
Σ	15	1		

2.6.4.- Media, mediana y moda. Cuarteles, rango

Duración	f_i	F_i
350 – 399	4	4
400 – 449	6	10
450 – 499	9	19
500 – 549	20	39
550 – 599	31	**70**
600 – 649	**80**	150
650 – 699	42	192
700 – 749	10	202
750 – 799	8	210
800 – 849	2	212
Total	212	

$$a = \frac{límite\ \ sup - límite\ \ inf}{\#\ de\ Intervalos}$$

$$a = \frac{849-350}{10} \rightarrow a = 49.9$$

$$Redondeamos\ \ al\ \ entero\ \ más\ \ cercano$$

$$a = 50$$

$$F_5 = 70\ \ frecuencia\ \ acumulada\ \ anterior$$

$$f_6 = 80\ \ frecuencia\ \ absoluta\ \ del\ \ intervalo\ \ 6$$

$$L_6 = 600\ \ límite\ \ real\ \ inferior\ \ de\ \ la\ \ mediana$$

Finalmente, calculamos la mediana reemplazando cada uno de los datos hallados:

$$Me = 600 + \left(\frac{106-70}{80} \right) \cdot 50$$

$$Me = 622.5$$

$i=4$ intervalo con mayor frecuencia absoluta

$$a=\frac{\text{lím superior}-\text{lím inferior}}{\#\text{ de intervalos}}$$

$$a=\frac{849-350}{10}\rightarrow a=49.9$$

Redondeamos al entero más cercano

$a=50$

$f_{4+1}=f_5=31$ frecuencia absoluta del intervalo 5

$f_{4-1}=f_3=9$ frecuencia absoluta del intervalo 3

$L_4=500$ límite inferior del intervalo 4

$$Mo=500+\left(\frac{31}{9+31}\right)\cdot 50$$

$Mo=538.75$

3.9 DATOS BIVARIADOS

Es muy frecuente que investigadores se interesen en más de sólo una variable que se pueda medir durante su investigación. Por ejemplo, una compañía aseguradora de autos podría estar interesada en el número de vehículos propiedad de un tenedor de pólizas, así como en el número de quienes conducen un vehículo en la familia. Un economista podría necesitar medir la cantidad gastada por semana en comestibles en una familia, y también el número de personas de esa familia. Un agente de ventas de bienes raíces podría medir el precio de venta de una propiedad residencial y la superficie en pies cuadrados de la sala.

Cuando dos variables se miden en una sola unidad experimental, los datos resultantes se denominan datos bivariados. ¿Cómo se deben presentar estos datos? No sólo son importantes ambas variables cuando se estudian por separado, sino que el experimentador también puede explorar la relación entre las dos variables. Los métodos para graficar datos bivariados, ya sean cualitativos o cuantitativos, permiten estudiar las dos variables juntas. Al igual que con datos univariados, se usan diferentes gráficas según el tipo de variables que se midan.

3.10 GRÁFICAS PARA VARIABLES CUALITATIVAS

Cuando al menos una de las dos variables es cualitativa, se pueden usar gráficas de pastel, ya sean sencillas o más elaboradas, gráficas de líneas y gráficas de barras para presentar y describir los datos. A veces habrá una variable cualitativa y una cuantitativa que se han medido en dos diferentes poblaciones o grupos. En este caso, se pueden usar dos gráficas de pastel lado a lado o una gráfica de barras en la que las barras para las dos poblaciones se colocan una al lado de la otra. Otra opción es usar una gráfica de barras apiladas, en la que las barras para cada categoría se ponen una sobre la otra.

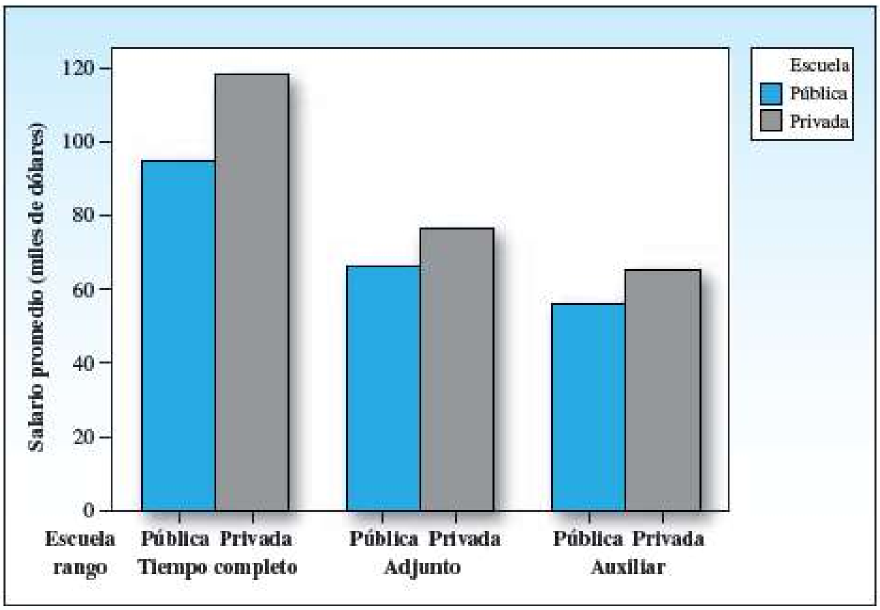
¿A los profesores de universidades privadas se les paga más que a los de universidades públicas? Los datos de la tabla 3.1 fueron recolectados de una muestra de 400 profesores universitarios cuyo rango, tipo de universidad y salario se registraron.1 El número en cada celda es el salario promedio (en miles de dólares) para todos los profesores que cayeron en esa categoría. Use una gráfica para contestar la pregunta planteada para esta muestra.

Salarios de profesores por rango y tipo de universidad

	De tiempo completo	Profesor adjunto	Profesor auxiliar
Pública	94.8	65.9	56.4
Privada	118.1	76.0	65.1

Fuente: Digest of Educational Statistics

Solución. Para presentar los salarios promedio de estos 400 profesores, usted puede usar una gráfica de barras lado a lado, como se muestra en la figura 3.1. La altura de las barras es el salario promedio, donde cada par de barras a lo largo del eje horizontal representa un rango profesional diferente. Los salarios son considerablemente más altos para profesores de tiempo completo en universidades privadas, pero hay menos diferencias sorprendentes en los dos rangos inferiores.



3.11 GRÁFICAS DE DISPERSIÓN PARA DOS VARIABLES CUANTITATIVAS

Cuando las dos variables que hayan de presentarse en una gráfica son cuantitativas, una de ellas se grafica a lo largo del eje horizontal y la otra a lo largo del eje vertical. Es frecuente que a la primera variable se le denomine x y, a la otra, y, de modo que la gráfica toma la forma de una gráfica en los ejes (x, y), que es más conocida. Cada par de valores de datos se grafica como punto en esta gráfica de dos dimensiones, llamada gráfica de dispersión. Es la extensión en dos dimensiones de la gráfica de puntos que usamos para graficar una variable cuantitativa en la sección 1.4.

Se puede describir la relación entre dos variables, x y y, usando los patrones que se muestran en la gráfica de dispersión.

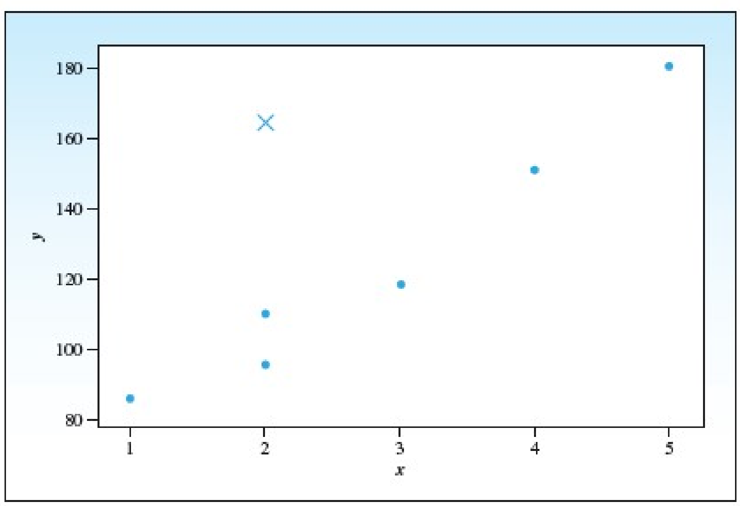
• ¿Qué tipo de modelo se muestra? ¿Hay una tendencia constante hacia arriba o hacia abajo que siga un modelo en línea recta? ¿Hay un modelo curvado? ¿No hay modelo en absoluto, sino sólo una dispersión aleatoria de puntos?

• ¿Qué tan fuerte es el modelo? ¿Todos los puntos siguen exactamente el modelo, o la relación es sólo débilmente visible?

• ¿Hay algunas observaciones poco comunes? Un resultado atípico es un punto que está lejos del conglomerado de los puntos restantes. ¿Los puntos se apiñan en grupos? Si es así, ¿hay una explicación para las agrupaciones observadas?

El número x de miembros de una familia, así como la cantidad y gastada por semana en comestibles, se miden para seis familias de una localidad. Trace una gráfica de dispersión de estos seis puntos de datos.

Solución. Marque el eje horizontal x y el eje vertical y. Grafique los puntos usando las coordenadas (x, y) por cada uno de los seis pares. La gráfica de dispersión de la figura 3.4 muestra los seis pares marcados como puntos. Se puede ver un modelo incluso con sólo seis pares de datos. El costo semanal de alimentos aumenta con el número de miembros de la familia en una relación aparente de línea recta. Supongamos que se encuentra que una séptima familia con dos miembros gastó $165 en alimentos. Esta observación se muestra como una X en la figura 3.4. No se ajusta al modelo lineal de las otras seis observaciones y está clasificada como resultado atípico. Posiblemente estas dos personas ¡tuvieron una fiesta en la semana de la encuesta!



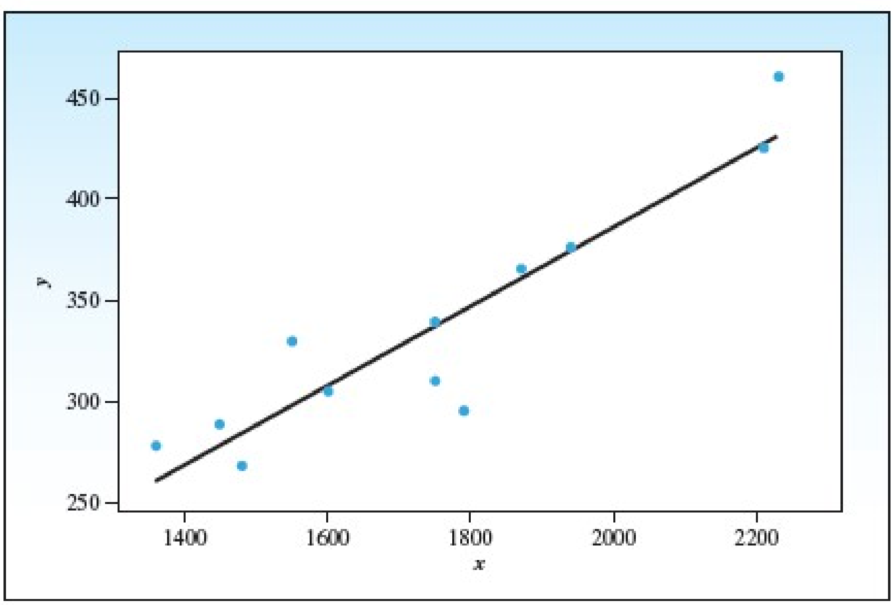
3.12 MEDIDAS NUMÉRICAS PARA DATOS CUANTITATIVOS BIVARIADOS

Una tasa constante de aumento o disminución es quizá el modelo más común que se encuentra en gráficas de dispersión bivariadas. La gráfica de dispersión de la figura 3.4 exhibe este modelo lineal, es decir, una recta con los puntos de datos arriba y debajo de la recta y a no más de una distancia fija desde la recta. Cuando éste es el caso, decimos que las dos variables exhiben una relación lineal.

Ejemplo. Los datos de la tabla 3.5 son la superficie del área de descanso (en pies cuadrados), x, y el precio de venta, y, de 12 residencias. La gráfica de dispersión del MINITAB de la figura 3.7 muestra un modelo lineal en los datos.

Área de descanso y precio de venta de 12 propiedades

Residencia	x (pies cuadrados)	y (en miles)
1	1360	$278.5
2	1940	375.7
3	1750	339.5
4	1550	329.8
5	1790	295.6
6	1750	310.3
7	2230	460.5
8	1600	305.2
9	1450	288.6
10	1870	365.7
11	2210	425.3
12	1480	268.8



Para los datos del ejemplo 3.5, se podría describir individualmente cada variable, x y

y, usando medidas descriptivas como lo son las medias

_

x y

_

y o las desviaciones estándar (sx y sy). No obstante, estas medidas no describen la relación entre x y y para una residencia en particular, es decir, la forma en que el tamaño del espacio de descanso afecta el precio de venta de la casa. Una medida sencilla que sirve a este propósito se denomina coeficiente de correlación, denotado por r, y se define como

r sxy

 s

xsy

UNIDAD III

DE LA ESTIMACIÓN

3.1.- Estimación insesgada y estimación ineficiente y viceversa.

Características estimadores

1) Sesgo. Se dice que un estimador es **insesgado** si la Media de la distribución del estimador es igual al parámetro.

Estimadores insesgados son la Media muestral (estimador de la Media de la población) y la Varianza (estimador de la Varianza de la población):

$$\boxed{\begin{array}{l} \overline{X} \Rightarrow \mu \\ \tilde{s}_X^2 \Rightarrow \sigma^2 \end{array}}$$

Ejemplo

En una población de 500 puntuaciones cuya Media (m) es igual a 5.09 han hecho un muestreo aleatorio (número de muestras= 10000, tamaño de las muestras= 100) y hallan que la Media de las Medias muestrales es igual a **5.09,** (la media poblacional y la media de las medias muestrales coinciden). En cambio, la Mediana de la población es igual a **5** y la Media de las Medianas es igual a **5.1** esto es, hay diferencia ya que la Mediana es un estimador sesgado.

La Varianza es un estimador sesgado. Ejemplo: La Media de las Varianzas obtenidas con la Varianza

$$\boxed{s^2 = \frac{\sum \left(X - \overline{X}(\ )\right)^2}{n}}$$

en un muestreo de 1000 muestras (n=25) en que la Varianza de la población es igual a **9.56** ha resultado igual a **9.12**, esto es, no coinciden. En cambio, al utilizar la Cuasivarianza

$$\tilde{s}^2 = \frac{\sum (X - \overline{X})^2}{n-1}$$

la Media de las Varianzas muestrales es igual a 9.5, esto es, coincide con la Varianza de la población ya que la Cuasivarianza es un estimador insesgado.

2) Consistencia. Un estimador es consistente si aproxima el valor del parámetro cuanto mayor es **n** (tamaño de la muestra).

Algunos estimadores consistentes son:

$$\overline{X} \Rightarrow \mu$$
$$\tilde{s}_X^2 \Rightarrow \sigma^2$$

Ejemplo

En una población de 500 puntuaciones cuya Media (m) es igual a **4.9** han hecho tres muestreos aleatorios (número de muestras= 100) con los siguientes resultados:

n	Media de las Medias muestrales
5	4.6
25	4.8
100	4.9

vemos que el muestreo en que n=100 la Media de las Medias muestrales toma el mismo valor que la Media de la población.

3) Eficiencia. Diremos que un estimador es más eficiente que otro si la Varianza de la distribución muestral del estimador es menor a la del otro estimador. Cuanto menor es la eficiencia, menor es la confianza de que el estadístico obtenido en la muestra aproxime al parámetro poblacional.

Ejemplo

La Varianza de la distribución muestral de la Media en un muestreo aleatorio (número de muestras: 1000, n=25) ha resultado igual a **0.4.** La Varianza de la distribución de Medianas ha resultado, en el mismo muestreo, igual a **1.12**, (este resultado muestra que la Media es un estimador más eficiente que la Mediana).

3.2.- Confiabilidad. Estimación por intervalo de confianza para parámetros de la población

Definiciones

El campo de la inferencia estadística trata básicamente de predicciones y generalizaciones. Por ejemplo, se puede afirmar, basándose en opiniones recogidas por medio de una encuesta, que en las próximas elecciones presidenciales el candidato de gobierno obtendrá 60% de los votos. Para hacer esta afirmación fue necesario determinar el porcentaje de votos favorables en una muestra seleccionada de la población. Al porcentaje obtenido de esta forma se le llama “estadístico” o “estadígrafo”.

A partir de este valor se puede estimar el porcentaje real de votos que dicho candidato obtendrá el día de la elección. Este porcentaje real proviene del universo de votantes o población y se le llama parámetro. En general, cualquiera de las medidas de resumen -promedio, desviación estándar, porcentaje, tasa- se considera “estadístico” si proviene de la muestra, y “parámetro” si proviene del universo o población.

La inferencia estadística entrega las herramientas para realizar afirmaciones acerca de un parámetro de la población, basándose en el valor del respectivo estadístico proveniente de una muestra. Para realizar estas inferencias es necesario conocer, previamente, la distribución de probabilidad del estadístico. A la distribución de probabilidad de un estadístico se le llama “distribución muestral”.

Podemos tomar una muestra, calcular en ella un estadístico (promedio o porcentaje, por ejemplo) y luego hacer afirmaciones respecto del correspondiente parámetro. Esto se conoce con el nombre de **estimación de parámetros**, y se puede hacer de dos formas:

- **Estimación puntual**: consiste en asumir que el parámetro tiene el mismo valor que el estadístico en la muestra.
- **Estimación por intervalos**: se asigna al parámetro un conjunto de posibles valores que están comprendidos en un intervalo asociado a una cierta probabilidad de ocurrencia. También se llaman "intervalos de confianza" debido a que la probabilidad asociada a ellos es la confianza de los mismos. Así, diremos que un intervalo de 99% de confianza es más confiable que uno de 95%. También se define la confianza de la estimación como la probabilidad de acertar con el intervalo.

La estimación que tiene valor estadístico para promedio o media y para el porcentaje de la población es esta última, que explicaremos a continuación.

Estimación de la media de la población

Explicaremos este punto con el siguiente ejemplo: queremos estimar el número de hijos promedio que tienen las mujeres de una población determinada. Con este objeto se seleccionó, por muestreo aleatorio simple, una muestra de 20 mujeres a quienes se entrevistó, obteniendo como resultado un promedio de 3,2 hijos y una desviación estándar de 0,8. Con estos resultados podríamos hacer una estimación puntual y decir que la población de interés tiene en promedio 3,2 hijos. Pero esta estimación tiene el inconveniente de que se desconoce el error que se está cometiendo.

Si a esta estimación le asignamos un error, que llamaremos E, podríamos decir que el promedio de hijos de la población está ubicado dentro de un intervalo de estimación que tiene como límite inferior 3,2 - E y como límite superior 3,2 + E. De este modo, le asignamos al resultado un intervalo de estimación. Si además le damos a este intervalo una probabilidad de ocurrencia de los valores comprendidos en él, habremos construido un intervalo de confianza para el promedio de hijos de nuestra población de mujeres.

Entonces, generalizando lo que se explicó para la variable "promedio de hijos", podemos decir que:

Un intervalo de confianza para estimar el promedio de la población está constituido por los siguientes elementos: el promedio de la muestra y el error de estimación.

El elemento esencial en la construcción del intervalo de estimación es el error.

¿Cómo se obtiene el error en la construcción de un intervalo para el promedio?

Desarrollando la fórmula siguiente:

$$E = t \times \frac{S}{\sqrt{n}}$$

Está compuesta por la desviación estándar de la muestra (S), el tamaño de la muestra (n) y, aquí aparece un elemento nuevo, **t** –*Student*-, que corresponde a una distribución de probabilidad muy similar a la distribución normal.

En la tabla de **t** los valores se buscan en función de dos cosas:

- la probabilidad que hemos elegido para nuestro intervalo, y
- los "grados de libertad" que se calculan restando 1 al tamaño de la muestra (n).

En nuestro ejemplo elegimos una confianza de 95% que, asociada a los 19 grados de libertad (n-1), nos conduce a un valor de tabla **t** de *Student* igual a 2,093. Ya veremos en forma detallada el uso práctico de la tabla **t**, recordemos por ahora el valor de "**t**" encontrado porque lo utilizaremos para la construcción del intervalo.

Volviendo a la fórmula para calcular el error, vemos entonces que el error está compuesto por tres elementos:

- El valor **t** que se obtiene de la tabla t de *Student*.
- La desviación estándar de la muestra.

- El tamaño de la muestra.

Volviendo a nuestro ejemplo, calculemos el error. Recordemos que deseamos conocer el número promedio de hijos que tienen las mujeres en esa población y que, estudiando una muestra de 20 mujeres, el resultado fue un promedio de 3,2 hijos y una desviación estándar de 0,8.

¿Cuáles son, entonces, los elementos que nos permitirán calcular el error de nuestra estimación?

El valor t que obtuvimos de la tabla t de Student	t = 2,093
La desviación estándar de la muestra	S = 0,8
El tamaño de la muestra	n = 20

Reemplazando esos valores en la fórmula obtendremos el error, que es:

$$E = t \times \frac{S}{\sqrt{n}} = 2{,}093 \times \frac{0{,}8}{\sqrt{20}} = 0{,}37$$

Intervalo de estimación

Construiremos ahora el intervalo de estimación, sumando y restando al promedio, el error. De esta manera el límite inferior será: promedio - E; y el límite superior: promedio + E.

Límite inferior (a) = 3,2 - 0,37 = 2,83
Límite superior (b) = 3,2 + 0,37 = 3,57

3.3.- Intervalo de confianza para medias y proporciones.

Intervalo de confianza para la media

Ya vimos que la distribución muestral de las medias corresponde a:

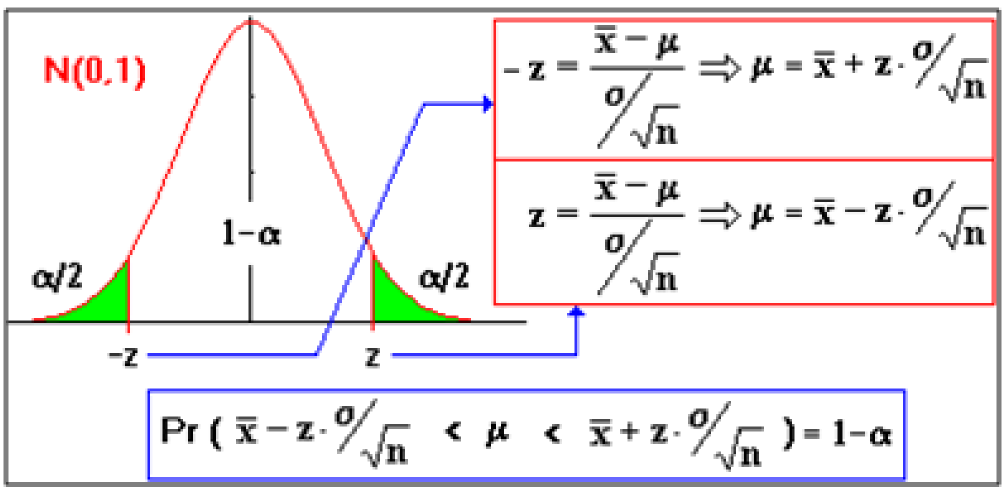
$$\bar{X} \rightarrow N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

Queremos **estimar** la media poblacional μ a partir de la media muestral $\bar{x}$, obteniendo para ello un intervalo de forma que tengamos una probabilidad alta (1 − α).100% de que la media poblacional esté en dicho intervalo.

Tipificando la expresión anterior:

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \rightarrow N(0,1)$$

Si fijamos una probabilidad α, podemos obtener -z y z que limitan un área de valor 1 - α. Deshaciendo la tipificación obtenemos el intervalo de confianza para la media:



En Resumen:

Intervalo de confianza para la media poblacional μ con un **nivel de confianza** de 1 − α es:

a) Varianza poblacional conocida:

$$\left(\bar{x} - z.\frac{\sigma}{\sqrt{n}}, \bar{x} + z.\frac{\sigma}{\sqrt{n}}\right)$$

b) Varianza poblacional desconocida y muestras grandes ($n \geq 30$):

$$\left(\bar{x} - z.\frac{s}{\sqrt{n-1}}, \bar{x} + z.\frac{s}{\sqrt{n-1}}\right)$$

Donde z, llamado **valor crítico**, es el valor que en la distribución N(0,1) deja a su derecha un área de $\frac{\alpha}{2}(cuartil\ 1 - \frac{\alpha}{2}$), s la desviación típica muestral y n el tamaño de la muestra.

Intervalo de confianza para la proporción

Sea p desconocida la proporción de elementos en la población pertenecientes a una categoría C, sacamos una muestra y se trata de obtener un intervalo de forma que tengamos una probabilidad alta (1-alfa).100% de que la proporción esté en ese intervalo.

Si se cumple una de las siguientes hipótesis, y que habrá de comprobarlas en todos los problemas son:

$n.\widehat{p} > 5 \quad n.(1 - \widehat{p}) > 5$

En estas condiciones se obtienen los siguientes intervalos según el tamaño de la muestra:

a) El tamaño de la muestra es mayor de 30 y menor o igual de 100.

$$\left(\widehat{p} - z.\sqrt{\frac{1}{4n}}, \widehat{p} + z.\sqrt{\frac{1}{4n}}\right)$$

b) El tamaño de la muestra es mayor de 100.

$$\left(\widehat{p} - z.\sqrt{\frac{\widehat{p}.(1-\widehat{p}}{n}}, \widehat{p} + z.\sqrt{\frac{\widehat{p}.(1-\widehat{p}}{n}}\right)$$

La determinación del tamaño muestral

Consiste en calcular el tamaño de la muestra necesario para que el error cometido al estimar el parámetro sea menor que una cierta cantidad, E, con un nivel de confianza de (1-alfa).100%.

a) Como para **la media** el error máximo cometido era de $z.\frac{\sigma}{\sqrt{n}}$ para garantizar que este valor sea menor o igual que un cierto error E, despejando n resulta:

$$n \geq \frac{z^2\sigma^2}{E^2}$$

b) Como para **la proporción** el error máximo cometido era de $z.\sqrt{\frac{\widehat{p}.(1-\widehat{p}}{n}}$ para garantizar que este valor sea menor o igual que un cierto error E, despejando n resulta:

$$n \geq \frac{z^2\widehat{p}.(1-\widehat{p}}{E^2}$$

3.4.- Intervalo para diferencias de medias y de proporciones.

DIFERENCIA DE PROPORCIONES $\pi_1 - \pi_2$

El estadístico de prueba que permite contrastar $H_0{:}\pi_1 = \pi_2$ frente a $H_1{:}\pi_1 \neq \pi_2$ a partir de dos muestras aleatorias e independientes es $Z=\frac{p_1 - p_2}{\sqrt{\frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2}}}$ siendo p la estimación de π obtenida del total de observaciones.

Si se consideran las proporciones como medias y se aplica la prueba t utilizada para comparar medias poblacionales los resultados no son fiables ya que la estimación del error típico que realiza el programa no coincide con la del estadístico de prueba. Para resolver el problema con el programa SPSS se deberá cruzar la variable analizada con la que define los grupos (obtener la tabla de contingencia) y realizar el contraste de independencia Chi-cuadrado.

El estadístico de prueba Chi-cuadrado se define: $\chi^2 = \sum_{j=1}^{2} \sum_{i=1}^{2} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$ y presenta una distribución Chi-cuadrado con (I-1)(J-1) grados de libertad. Las Eij se calculan multiplicando las frecuencias marginales y dividendo el producto por n. Estas Eij son estimaciones de las frecuencias absolutas que cabría esperar en cada casilla bajo el supuesto de que la proporción de éxitos es la obtenida a partir del total de observaciones muestrales sin considerar diferencias entre los dos grupos.

La secuencia es:

Analizar

Estadísticos Descriptivos

Tablas de contingencia

En el cuadro de diálogo se indica la variable que se quiere contrastar (filas), la variable que define los dos grupos (columnas) y se selecciona la opción Chi-cuadrado en Estadísticos.

EJEMPLO

Con referencia a la encuesta ***Enctrans.sav*** *se quiere comprobar si la proporción de alumnos con vehículo difiere significativamente entre los grupos definidos según el género.*

La hipótesis nula del contraste es $H_0: \pi_1 = \pi_2$; siendo π_1 la proporción poblacional de hombres con vehículo y π_2 la proporción poblacional de mujeres con vehículo.

Con la secuencia *Analizar > Estadísticos Descriptivos > Tablas de contingencia* se accede al cuadro de diálogo donde se indica que la variable a contrastar es Vehículo y que la variable de agrupación es el Género, y se selecciona la opción *Chi-cuadrado en Estadísticos*. Al aceptar se obtiene el siguiente cuadro de resultados.

Tabla de contingencia Tiene vehículo? * GENERO

			GENERO		Total
			hombre	mujer	
Tiene vehículo?	No	Recuento	33	42	75
		Frecuencia esperada	35,5	39,5	75,0
		% del total	28,9%	36,8%	65,8%
	Sí	Recuento	21	18	39
		Frecuencia esperada	18,5	20,5	39,0
		% del total	18,4%	15,8%	34,2%
Total		Recuento	54	60	114
		Frecuencia esperada	54,0	60,0	114,0
		% del total	47,4%	52,6%	100,0%

Pruebas de chi-cuadrado

	Valor	gl	Sig. asint. (bilateral)	Sig. exacta (bilateral)	Sig. exacta (unilateral)
Chi-cuadrado de Pearson	,998[b]	1	,318		
Corrección de continuidad[a]	,642	1	,423		
Razón de verosimilitud	,998	1	,318		
Estadístico exacto de Fisher				,331	,212
Asociación lineal por lineal	,989	1	,320		
N de casos válidos	114				

a. Calculado sólo para una tabla de 2x2.

b. 0 casillas (,0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 18,47.

Si es cierto que la proporción de propietarios de vehículo es la misma en los dos grupos, $\pi_1 = \pi_2 = \pi$, la estimación de π es la proporción de propietarios de vehículo para el total de alumnos de la muestra, es decir, 39/114=0,3421. La frecuencia esperada de hombres con vehículo se obtendrá multiplicando esta proporción por el total de hombres en la muestra, o sea, 0,3421 ·54=18,5; y de la misma forma se obtendrá la frecuencia esperada de mujeres con vehículo: 0,3421 ·60=20,5 (veáse que estas frecuencias esperadas coinciden con las que cabría esperar en el caso de que las variables Género y Vehículo fueran independientes).

El estadístico Chi-cuadrado toma el valor 0,998 y el nivel de significación crítico es 0,318, por lo tanto no se rechaza la hipótesis nula para los niveles de significación habituales y se puede aceptar que no hay diferencia entre la proporción de hombres y mujeres propietarios de vehículos.

3.4.1.- Curva de ajuste, regresión y correlación

Muy a menudo se encuentra en la práctica que existe una relación entre dos (o más) variables. Por ejemplo: los pesos de los hombres adultos dependen en cierto modo de sus alturas; las longitudes de las circunferencias y las áreas de los círculos dependen del radio, y la presión de una masa de gas depende de su temperatura y de su volumen.

Si todos los valores de las variables cumplen exactamente una relación exacta, entonces se dice que las variables están perfectamente correlacionadas o que hay una correlación perfecta entre ellas o, más sencillamente, que existe una función o una fórmula que las relaciona.

Así la longitud L de una circunferencia y su radio r están perfectamente correlacionados pues se verifica exactamente que :

L = 2p r

Por el contrario, si se lanzan simultáneamente dos dados unas cuantas veces, no existirá una relación entre los puntos que se obtengan en cada dado (salvo que los dados estén cargados), es decir no existirá correlación entre las puntuaciones de cada dado.

En otros casos, parece que existe *cierta correlación,* aunque ésta no sea perfecta. Por ejemplo, las variables altura y peso de los individuos parecen tener cierto grado de relación aunque no exista una fórmula que nos permita adivinar el peso de un individuo conocida su altura.

Entonces, surge la siguiente pregunta fundamental:

¿En qué medida están relacionados la altura y el peso de un individuo?

Si somos capaces de encontrar una forma de medir adecuadamente esa relación, entonces, por ejemplo, podemos decidir si la altura y el peso de un individuo están mas relacionados entre sí que la altura de ese individuo y la altura de su padre.

Cuando se busca una medida para medir esa relación se dice que se está buscando medir la correlación entre esas dos variables.

Por tanto, averiguar la correlación entre dos variables se refiere siempre a hallar una medida de la relación entre esas dos variables.

Cuando se trata de dos variables solamente, se habla de correlación simple y cuando se trata de más de dos variables se habla de correlación múltiple.

Aquí vamos a tratar solamente de la correlación simple.

Volviendo al ejemplo de la altura y el peso de un individuo, y aunque todos aceptemos que no existe una fórmula exacta que relacione esas dos variables, sí que parece bastante difícil (aunque no totalmente imposible, en principio) que una persona de 120 kilogramos de peso mida 80 centímetros de altura.

Es decir, surge otra pregunta fundamental:

¿Podríamos hallar una fórmula para estimar -siquiera aproximadamente- el peso de un individuo a partir de su altura (o al revés)?

Cuando se busca una fórmula de ese tipo se dice que se está buscando una regresión entre esas dos variables.

Por tanto, hallar una regresión entre dos variables se refiere siempre a hallar una fórmula o ecuación que represente la relación aproximada entre esas dos variables.

Y de la misma forma que antes, cuando se trata de dos variables solamente, se habla de regresión simple. Cuando se trata de más de dos variables se habla de regresión múltiple.

Aquí vamos a tratar solamente de la regresión simple.

Correlación y regresión (I).

Nube de puntos.

Para estudiar y medir la relación entre dos variables, el primer paso es recoger los datos que muestren los correspondientes valores de las variables consideradas.

Por ejemplo, si disponemos de los datos de la altura y del peso de 100 individuos, lo primero sería representar en un gráfico cartesiano los 100 puntos (x,y) donde x e y serían la altura y el peso respectivo de cada individuo.

El conjunto de puntos que así se obtiene se suele denominar diagrama de dispersión o mas sencillamente nube de puntos.

Por ejemplo, en la escena siguiente se puede contemplar una nube de puntos real, obtenida a partir de datos reales de los que luego hablaremos.

Propuesta de trabajo:

Observarás que la nube de puntos está constituida por doce elementos, denominados con las doce primeras letras del alfabeto.

Si haces "clic" sobre cada punto verás que aparecen las coordenadas del mismo.

Si te apetece, puedes jugar con los otros botones y no te preocupes si te equivocas, pues siempre puedes volver a la situación inicial pulsando el botón denominado inicio.

De hecho, necesariamente deberás manejar esos botones para ver el punto L, que no se encuentra en la imagen inicial.

Tarea A) Lo que te pedimos, en primer lugar, es que escribas en un papel las coordenadas de cada uno de los doce puntos.

La primera coordenada corresponde al dato estadístico del llamado Índice de Precios Industriales (IPRI), elaborado por el Instituto Nacional de Estadística (INE) y que es un dato que pretende medir la evolución de los precios de los productos industriales.

La segunda coordenada corresponde al llamado IBEX-35 que es el índice de referencia mas importante de la Bolsa española. Para facilitar su representación en la escena anterior, los valores reales se han dividido por 100.

En la Bolsa de Nueva York, la aparición del IPRI americano tiene gran influencia, cosa que no ocurre (hasta ahora) en la Bolsa española.

Precisamente, lo que tratamos de estudiar es qué correlación existe entre ambas variables *y si sería posible estimar el IBEX-35 a partir del IPRI.*

Los datos que se han usado en los doce puntos corresponden a los doce últimos datos mensuales de ambas variables, disponibles el 1 de enero del año 2000.

Como dato mensual del IBEX-35 se ha usado el de cierre del primer día hábil en la Bolsa de cada uno de los doce meses de 1999.

Respecto al IPRI, el dato usado es el más reciente en cada uno de esos primeros días hábiles, es decir el que hipotéticamente puede haber influido más. Concretamente, los 12 datos del IPRI usados son los correspondientes a noviembre y diciembre de 1998 y los diez primeros meses de 1999.

Tarea B) En segundo lugar, lo que te pedimos es que estudies detenidamente la nube de puntos anterior y que intentes seleccionar aquellos puntos que te parecen mas "raros" o que se alejan mas del resto.

Correlación y regresión (II).

Recta de ajuste.

Con el diagrama de dispersión o nube de puntos, es posible frecuentemente representar una curva que se aproxime a los datos.

Tal curva se llama curva de aproximación.

En la mayor parte de las nubes de puntos obtenidas a partir de casos reales es difícil imaginarse cuál sería la mejor curva de aproximación y, generalmente, hay que optar por una determinada (usando algunos criterios específicos) que se suele denominar curva de ajuste.

Nosotros vamos a usar como criterio el de la simplicidad y dado que la curva mas sencilla es la recta, vamos a optar por buscar una recta de ajuste que se ajuste adecuadamente a nuestra nube de puntos.

Desde luego, *la forma mas sencilla de obtener una recta de ajuste es dibujando una recta encima de la nube de puntos, tratando de que dicha recta se ajuste lo mejor posible a la nube de puntos.*

Observa la escena siguiente y comprueba que puedes obtener otra recta moviendo el punto P y modificando la pendiente m. Para mover el punto P puedes arrastrarlo directamente con tu ratón, o bien usar los controles situados debajo de la escena y para modificar la pendiente m puedes usar el control de la pendiente situado también debajo de la escena y no olvides que siempre puedes volver a la posición inicial pulsando el botón inicio.

<u>Propuesta de trabajo:</u>

Lo que intentamos es encontrar gráficamente una buena recta de ajuste.

Tarea C) Lo que te pedimos ahora es que escojas aquella recta que, en tu opinión, *se ajuste mejor a la nube de puntos.*

Para ello, haz todas las pruebas que consideres necesarias y, cuando hayas escogido *tu recta de ajuste*, apunta el valor de la pendiente y las coordenadas del punto usado, de modo que puedas comparar posteriormente tu elección con *la elección de los matemáticos.*

Correlación y regresión (III).

Recta de regresión por mínimos cuadrados.

Es fácilmente comprensible que los matemáticos hayan intentado encontrar un procedimiento común para seleccionar la misma recta de ajuste, de modo que todo el mundo esté de acuerdo y no haya que atenerse a opiniones subjetivas.

La recta de ajuste seleccionada por los matemáticos es la llamada Recta de regresión por mínimos cuadrados y que se obtiene seleccionando de entre todas las rectas de ajuste posibles, aquélla que hace mínimo la suma de los cuadrados de las distancias verticales de los puntos a la recta (ahora te explicamos todo esto).

Propuesta de trabajo:

A partir de la escena anterior, puedes hacerte una idea de cómo los matemáticos han seleccionado *su mejor recta de ajuste*.

Lo que han hecho es seleccionar aquella recta que hace mínimo el resultado de sumar el cuadrado de cada una de las longitudes representadas en color magenta.

Esto se ha hecho así por diversas razones pero es importante precisar que podía haberse hecho de otra forma y que, incluso, en determinados casos especiales se prefiere usar otro criterio para la selección de la mejor recta de ajuste, aunque el caso aquí desarrollado es, desde luego, el más frecuente.

Ese cuadrado (de cada una de los segmentos de color magenta) se puede imaginar como el área de cada uno de los cuadrados que podrían construirse sobre cada uno de esos segmentos de color magenta.

Tarea D) Lo que te pedimos ahora es que escojas aquella otra recta que, en tu opinión, *se ajuste mejor a la nube de puntos, usando el criterio de los matemáticos.*

Para ello, haz todas las pruebas que consideres necesarias y, cuando hayas escogido *tu recta de ajuste usando el criterio de los matemáticos,* apunta el valor de la pendiente y las coordenadas del punto usado, de modo que puedas comparar posteriormente tu elección con *el resultado correcto, calculado matemáticamente con toda exactitud.*

Correlación y regresión (IV).

Ecuación de la recta de regresión por mínimos cuadrados.

A continuación vamos a indicar cuál es la ecuación de la recta de regresión por mínimos cuadrados, sin entrar en las demostraciones matemáticas exactas.

Para calcular la ecuación de la recta de regresión por mínimos cuadrados se hace lo siguiente, que explicamos paso a paso, indicando también las fórmulas exactas.

- Se calcula la media aritmética (μ_x) y la desviación típica (σ_x)de los datos de la primera variable. En nuestro caso, se debe calcular la media aritmética y la desviación típica de las coordenadas "x" de los puntos de la nube de puntos.

- Se calcula la media aritmética de los datos de la segunda variable(μ_y). En nuestro caso, se debe calcular la media aritmética de las coordenadas "y" de los puntos de la nube de puntos.

- Se calcula la llamada covarianza entre las las dos variables(σ_{xy}). En nuestro caso, se debe calcular la covarianza entre las coordenadas "x" e "y" de los puntos de la nube de puntos.

Las fórmulas exactas para lo anterior son las que siguen.

Denotamos a los doce puntos de la nube de puntos de la siguiente forma:

$(x^1, y^1), (x^2, y^2), \ldots, (x^{12}, Y^{12})$

Entonces, se tiene:

$$\mu_x = \frac{x_1 + x_2 + \ldots. x_{12}}{12}$$

$$\mu_y = \frac{y_1 + y_2 + \ldots.. + y_{12}}{12}$$

$$\sigma_x = \sqrt{\frac{(x_1 - \mu_x)^2 + (x_2 - \mu_x)^2 + \ldots\ldots + (x_{12} - \mu_x)^2}{12}}$$

$$\sigma_{xy} = \frac{(x_1 - \mu_x)(y_1 - \mu_y)}{12} + \frac{(x_2 - \mu_x)(y_2 - \mu_y)}{12} + \ldots + \frac{(x_{12} - \mu_x)(y_{12} - \mu_y)}{12}$$

Hecho todo lo anterior, se obtiene la ecuación de la recta de regresión por mínimos cuadrados de la forma siguiente:

$$y - \mu_y = m\,(x - \mu_x)$$

donde la pendiente m es igual a:

$$\frac{\sigma_{xy}}{\sigma_x^2}$$

Es decir que la recta de regresión por mínimos cuadrados es la recta que pasa por el punto

(μ_x , μ_y) y que tiene por pendiente a: $\dfrac{\sigma_{xy}}{\sigma_x^2}$

Y si aplicamos las fórmulas anteriores a nuestra nube de puntos (y redondeando para usar sólo tres decimales), resulta la recta siguiente, que es la la recta de regresión por mínimos cuadrados de nuestra nube de puntos:

y = 15,283 + 0,711x

En la siguiente escena, hemos representado en color magenta a dicha recta de regresión por mínimos cuadrados de nuestra nube de puntos y se trata de que la compares con tus rectas de ajuste obtenidas anteriormente, en las Tareas C y D anteriores.

Propuesta de trabajo:

Usando los controles inferiores de la escena, representa en ella a tus rectas de ajuste (seleccionadas en las tareas C y D anteriores) y compáralas con la recta matemáticamente obtenida y que figura representada en color magenta.

Tarea E) Lo que te pedimos ahora es que escribas un breve informe, tratando de explicar tus resultados en comparación a la recta de regresión por mínimos cuadrados y, especialmente, tratando de indicar si la recta de ajuste de los matemáticos te parece una buena opción, es decir si se ajusta bien a la nube de puntos (en tu opinión) o no y las razones de todo ello.

Correlación y regresión (V).

Coeficiente de correlación lineal.

A continuación vamos a hablar de la medida usual de correlación lineal entre dos variables, el llamado Coeficiente de correlación lineal (ρ , léase "ro").

Hecho todo lo anterior, es muy fácil calcular dicho coeficiente, a partir de la fórmula siguiente:

$$\rho = \frac{\sigma_{xy}}{\sigma_x \, \sigma_y}$$

donde σ_y se define de forma análoga a σ_x, es decir:

$$\sigma_y = \sqrt{\frac{(y_1 - \mu_y)^2 + (y_2 - \mu_y)^2 + \ldots\ldots + (y_{12} - \mu_y)^2}{12}}$$

Lo que este coeficiente ρ mide es el grado de ajuste de la recta de regresión a una determinada nube de puntos.

Cuanto mayor sea este ajuste, más confiados debemos estar en que es correcto usar a una recta como modelo de nuestra nube de puntos, pero si el ajuste no se considera bueno, deberemos pensar que nuestra nube de puntos no se representa bien por una recta y habrá que buscar otros modelos (quizás nuestra nube de puntos se ajuste mejor a una parábola, por ejemplo).

Conviene también destacar que a partir de la definición anterior, se puede comprobar que la ecuación de la recta de regresión puede expresarse también como:

y - μ_y = m(x - μ_x)

donde la pendiente m es igual a:

$$\rho \frac{\sigma_y}{\sigma_x}$$

Es decir que la recta de regresión por mínimos cuadrados es la recta que pasa por el punto

(μ_x, μ_y) y que tiene por pendiente :

$$\rho \frac{\sigma_y}{\sigma_x}$$

Algunas consideraciones importantes sobre el coeficiente de correlación lineal:

- Es una cantidad sin dimensiones, es decir no depende de las unidades empleadas. Por ejemplo, si se está buscando hallar el coeficiente de correlación entre el peso y la altura de los niños en determinada ciudad, entonces el resultado será el mismo independientemente de si el peso de *todos* los niños se mide en Kilogramos o en gramos e independientemente de si la altura de *todos* los niños se mide en metros o centímetros.

- Se verifica siempre que:

$$-1 \leq \rho \leq 1$$

- Si el coeficiente de correlación es igual a 1, entonces hay una correlación lineal positiva perfecta, es decir que los datos se ajustan perfectamente a una recta de pendiente positiva, es decir una recta *que crece*, o sea que *cuando x aumenta, entonces también lo hace y*.

- Si el coeficiente de correlación es igual a -1, entonces hay una correlación lineal negativa perfecta, es decir que los datos se ajustan perfectamente a una recta de pendiente negativa, es decir una recta *que decrece*, o sea que *cuando x aumenta, entonces y disminuye*.

- En cualquier otro caso, para aceptar si hay una correlación lineal aceptable, no hay ninguna regla estricta. Normalmente, para aceptar la existencia de dicha correlación, el coeficiente debe ser mayor que 0,7 o menor que -0,7. En caso contrario, se suele rechazar la existencia de correlación lineal.

¿Qué puede deducirse si se rechaza la existencia de correlación lineal si, por ejemplo, se encuentra un coeficiente de correlación lineal de 0,3 entre dos variables?

- Lo único que puede deducirse es que los datos no se ajustan a una recta.

- Pero esto no significa que no haya relación entre ellos dado que podrían ajustarse a una parábola o a cualquier otra curva. Sólo se deduce que *no hay correlación lineal aunque pudiera haber una correlación no lineal.*

- Este es el gran inconveniente del coeficiente de correlación lineal: no sirve para decidir si hay o no una posible *relación* entre dos variables, sólo sirve para decidir si hay o no una posible *relación lineal* entre dos variables.

- Ello hace que, definitivamente, *la única manera de decidir inicialmente si debe sospecharse o no la existencia de relación entre dos variables es estudiar detenidamente el diagrama de dispersión correspondiente, o sea la nube de puntos.*

- Y, en su caso, sólo después habrá que decidir con que curva se intentan ajustar los datos.

3.4.2.- Pruebas de hipótesis para medias, proporciones, diferencia de medias

Las pruebas de proporciones son adecuadas cuando los datos que se están analizando constan de cuentas o frecuencias de elementos de dos o más clases. El objetivo de estas pruebas es evaluar las afirmaciones con respecto a una proporción (o Porcentaje) de población. Las pruebas se basan en la premisa de que una proporción muestral (es decir, x ocurrencias en n observaciones, o x/n) será igual a la proporción verdadera de la población si se toman márgenes o tolerancias para la variabilidad muestral. Las pruebas suelen enfocarse en la diferencia entre un número esperado de ocurrencias, suponiendo que una afirmación es verdadera, y el número observado realmente. La diferencia se compara con la variabilidad prescrita mediante una distribución de muestreo que tiene como base el supuesto de que H_0 es realmente verdadera.

En muchos aspectos, las pruebas de proporciones se parecen a las pruebas de medias, excepto que, en el caso de las primeras, los datos muestrales se consideran como cuentas en lugar de como mediciones. Por ejemplo, las pruebas para medias y proporciones se pueden utilizar para evaluar afirmaciones con respecto a:

1) Un parámetro de población único (prueba de una muestra)

2) La igualdad de parámetros de dos poblaciones (prueba de dos muestras), y

3) La igualdad de parámetros de más de dos poblaciones (prueba de k muestras). Además, para tamaños grandes de muestras, la distribución de muestreo adecuada para pruebas de proporciones de una y dos muestras es aproximadamente normal, justo como sucede en el caso de pruebas de medias de una y dos muestras.

Prueba de proporciones de una muestra

Cuando el objetivo del muestreo es evaluar la validez de una afirmación con respecto a la proporción de una población, es adecuado utilizar una prueba de una muestra. La metodología de prueba depende de si el número de observaciones de la muestra es grande o pequeño.

Como se habrá observado anteriormente, las pruebas de grandes muestras de medias y proporciones son bastante semejantes. De este modo, los valores estadísticos de prueba miden la desviación de un valor estadístico de muestra a partir de un valor propuesto. Y ambas pruebas se basan en la distribución normal estándar para valores críticos. Quizá la única diferencia real entre las ambas radica en la forma corno se obtiene la desviación estándar de la distribución de muestreo.

Esta prueba comprende el cálculo del valor estadístico de prueba Z.

$$Z_{prueba} = \frac{\frac{x}{n} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Donde:

$x = ocurrencias$

$n = observaciones$

$\frac{x}{n} = proporción\ de\ la\ muestra$

$p_0 = proporción\ propuesta$

$$\sqrt{\frac{p_0(1-p_0)}{n}} = desviación\ estándar\ de\ la\ proporción$$

Si se muestrea a partir de una población finita

$$\frac{n}{N} \cdot 100\% > 5\%$$

Se debe utilizar el factor finito de corrección

$$Z_{prueba} = \frac{\frac{x}{n} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}} \cdot \sqrt{\frac{N-n}{N-1}}}$$

Posteriormente este valor es comparado con el valor de Z, obtenido a partir de una tabla normal a un nivel de significación seleccionado.

Como ocurrió con la prueba de medias de una muestra, las pruebas de proporciones pueden ser de una o dos colas.

El tipo de prueba refleja H_1. Por ejemplo, hay tres posibilidades para H_1:

H_1: $p > p_0$ H_1: $p < p_0$ H_1: $p \neq p_0$

La hipótesis nula es: H_0: $p = p_0$

La primera alternativa establece una prueba de cola derecha, la segunda, izquierda y la tercera, una prueba de dos colas.

Ejemplo ilustrativo

En un estudio se afirma que 3 de 10 estudiantes universitarios trabajan. Pruebe esta aseveración, a un nivel de significación de 0,025, respecto a la alternativa de que la proporción real de los estudiantes universitarios trabajan es mayor de lo que se afirma, si una muestra aleatoria de 600 estudiantes universitarios revela que 200 de ellos trabajan. La muestra fue tomada de 10000 estudiantes**.**

Los datos son:

$$p_0 = \frac{3}{10} = 0{,}333$$

$\alpha = 0{,}025$
$n = 600$
$X = 200$
$N = 10000$

Las hipótesis son:

$H_0 : p = p_0$

$H_1 : p > p_0$

Con lectura en la tabla para un área de 0,025 le corresponde un valor $Z_{tabla} = 1{,}96$. Se toma en cuenta el valor positivo porque se trata de una prueba de hipótesis a cola derecha.

Como en los datos aparece el tamaño de la población, se debe verificar si el tamaño de la nuestra es mayor que el 5%. Se remplaza valores en la siguiente fórmula:

$$\frac{n}{N} \cdot 100\% > 5\%$$

$$\frac{600}{10000} \cdot 100\% = 6\%$$

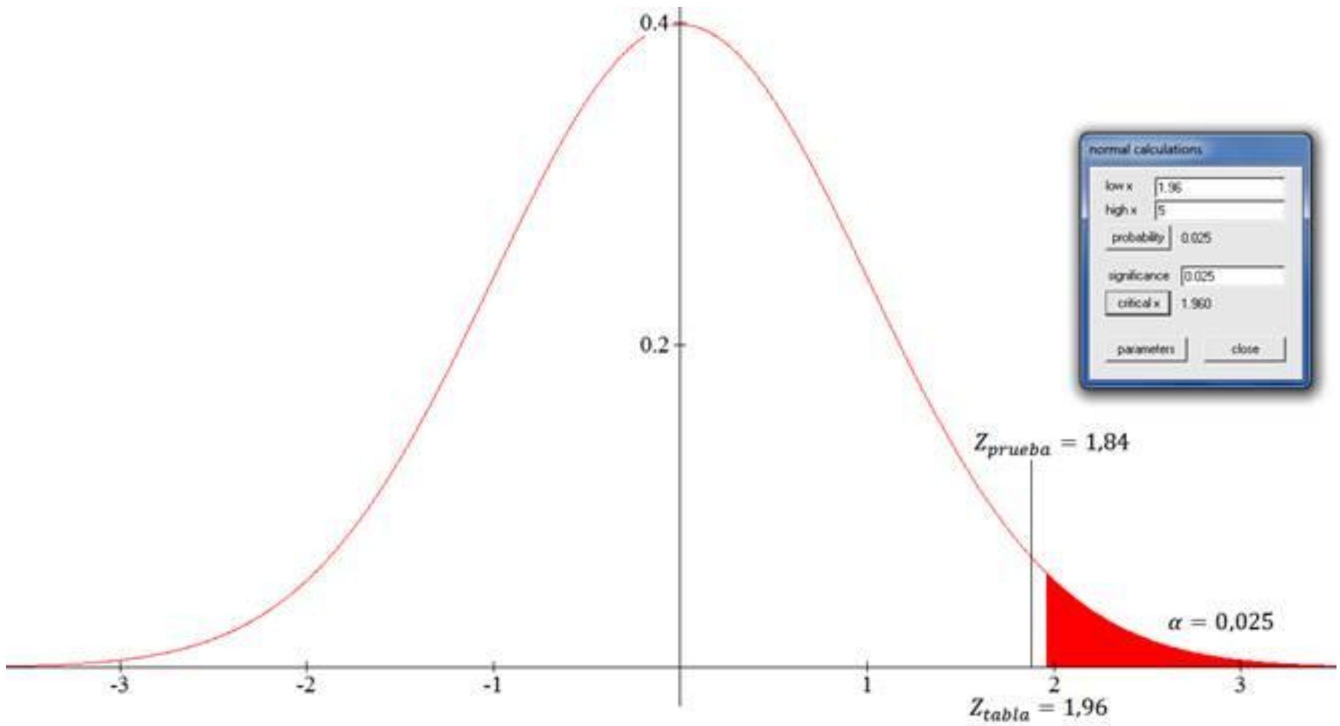
Por lo tanto se debe utilizar la fórmula con el factor finito de corrección.

$$Z_{prueba} = \frac{\frac{x}{n} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}} \cdot \sqrt{\frac{N-n}{N-1}}} = \frac{\frac{200}{600} - 0{,}333}{\sqrt{\frac{0{,}333(1-0{,}333)}{600}} \cdot \sqrt{\frac{10000-6000}{10000-1}}} = 1{,}84$$

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E	F	G	H	I
1	p_o	0,3	=3/10						
2	n	600							
3	x	200							
4	α	0,025							
5	N	10000	$\frac{n}{N} \cdot 100\% > 5\%$		6	=(B2/B5)*100			
6	$H_O: p = p_0$								
7	$H_1: p > p_0$								
8	z tabla	-1,96	=DISTR.NORM.ESTAND.INV(B4)						
9		1,96	=B8*-1						
10									
11	$z_{prueba} = \frac{\frac{x}{n} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}} \sqrt{\frac{N-n}{N-1}}}$			1,84	=(B3/B2-B1)/(RAIZ(B1*(1-B1)/B2)*RAIZ((B5-B2)/(B5-1)))				
12									
13									

El gráfico elaborado en Winstats y Paint se muestra a continuación:



Decisión:

H_0 es aceptada, ya que $Z_{prueba} = 1{,}84$ es menor que $Z_{tabla} = 1{,}96$, por lo tanto es cierto que 3 de 10 estudiantes universitarios trabajan.

Prueba de proporciones de dos muestras

El objetivo de una prueba de dos muestras es determinar si las dos muestras independientes fueron tomadas de dos poblaciones, las cuales presentan la misma proporción de elementos con determinada característica. La prueba se concentra en la diferencia relativa (diferencia dividida entre la desviación estándar de la distribución de muestreo) entre las dos proporciones muestrales. Diferencias pequeñas denotan únicamente la variación casual producto del muestreo (se acepta H0), en tanto que grandes diferencias significan lo contrario (se rechaza H0). El valor estadístico de prueba (diferencia relativa) es comparado con un valor tabular de la distribución normal, a fin de decidir si H0 es aceptada o rechazada. Una vez más, esta prueba se asemeja considerablemente a la prueba de medias de dos muestras.

La hipótesis nula en una prueba de dos muestras es

$$H_0: p_1 = p_2$$

Las hipótesis alternativas posibles son

$$H_1: p_1 \neq p_2 \quad H_1: p_1 > p_2 \quad H_1: p_1 < p_2$$

La estimación combinada de p se puede calcular de la siguiente manera:

$$p = \frac{x_1 + x_2}{n_1 + n_2}$$

Donde:
p = proporción muestral
x_1 = número de aciertos en la muestra 1
x_2 = número de aciertos en la muestra 2
n_1 = número de observaciones de la muestra 1
n_2 = número de observaciones de la muestra 2

Este valor de p se utiliza para calcular el valor estadístico de prueba

$$z_{prueba} = \frac{\frac{x_1}{n_1} - \frac{x_2}{n_2}}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Ejemplo ilustrativo

Se ponen a prueba la enseñanza de la Estadística empleando Excel y Winstats. Para determinar si los estudiantes difieren en términos de estar a favor de la nueva enseñanza se toma una muestra de 20 estudiantes de dos paralelos. De paralelo A 18 están a favor, en tanto que del paralelo B están a favor 14. ¿Es posible concluir con un nivel de significación de 0,05 que los estudiantes que están a favor de la nueva enseñanza de la Estadística es la misma en los dos paralelos?.

Los datos son:

$n_1 = 20$

$n_2 = 20$

$x_1 = 18$

$x_2 = 14$

$\alpha = 0{,}05$

Las hipótesis son

$H_0: p_1 = p_2$

$H_1: p_1 \neq p_2$

Como se trata de una prueba de hipótesis a dos colas se debe calcular

$$\frac{\alpha}{2} = \frac{0{,}05}{2} = 0{,}025$$

Con lectura en la tabla para un área de 0,025 le corresponde un valor $Z_{tabla} = \pm 1{,}96$.

Calculando la proporción muestral se obtiene:

$$p = \frac{x_1 + x_2}{n_1 + n_2} = \frac{18 + 14}{20 + 20} = 0{,}8$$

Calculando Z_{prueba} se obtiene:

$$Z_{prueba} = \frac{\frac{x_1}{n_1} - \frac{x_2}{n_2}}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{\frac{18}{20} - \frac{14}{20}}{\sqrt{0{,}8(1-0{,}8)\left(\frac{1}{20} + \frac{1}{20}\right)}} = 1{,}58$$

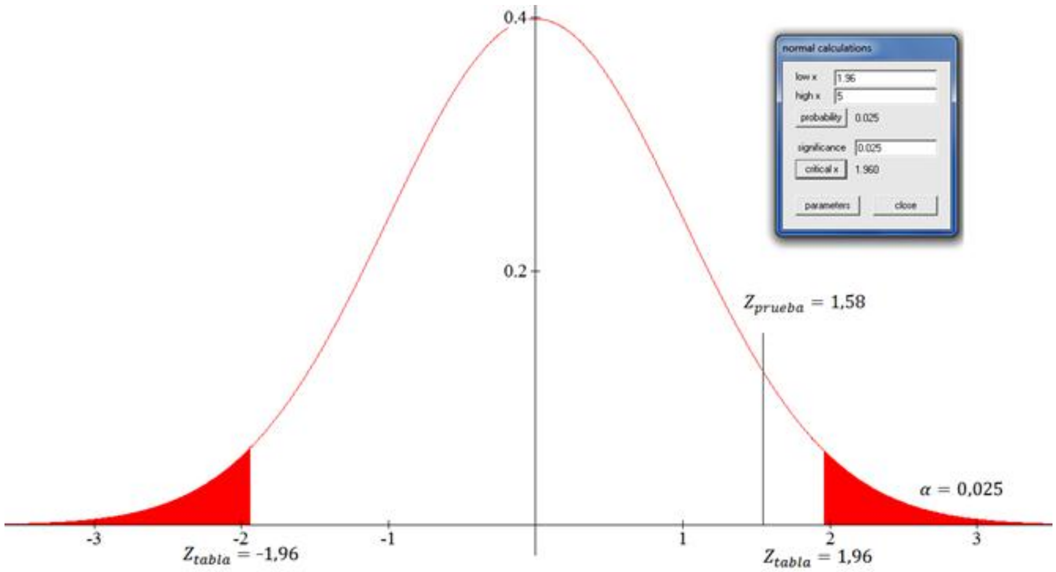
$$Z_{prueba} = \frac{\frac{18}{20} - \frac{14}{20}}{\sqrt{0{,}8(1-0{,}8)\left(\frac{1}{20} + \frac{1}{20}\right)}} = 1{,}58$$

$$Z_{prueba} = 1{,}58$$

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E	F	G
1	n_1	20					
2	n_2	20					
3	x_1	18					
4	x_2	14					
5	α	0,05					
6	$H_0: p_1 = p_2$						
7	$H_1: p_1 \neq p_2$						
8	$\frac{\alpha}{2}$	0,025					
9							
10	Z_{tabla}	-1,96	=DISTR.NORM.ESTAND.INV(B8)				
11		1,96	=B10*-1				
12	$p = \frac{x_1 + x_2}{n_1 + n_2}$	0,8	=(B3+B4)/(B1+B2)				
13							
14							
15	$Z_{prueba} = \frac{\frac{x_1}{n_1} - \frac{x_2}{n_2}}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$	1,58	=(B3/B1-B4/B2)/RCUAD(B12*(1-B12)*(1/B1+1/B2))				
16							
17							

El gráfico elaborado en Winstats y Paint se muestra a continuación:



Decisión:

H_0 es aceptada, ya que $Z_{prueba} = 1{,}58$ está en la zona de aceptación $Z_{tabla} = \pm 1{,}96$, entonces, la proporción de los estudiantes que están a favor de la nueva enseñanza de la Estadística es la misma en los dos paralelos.

Prueba de proporciones de k muestras

La finalidad de una prueba de k muestras es evaluar la aseveración que establece que todas las k muestras independientes provienen de poblaciones que presentan la misma proporción de algún elemento. De acuerdo con esto, las hipótesis nula y alternativa son

H_0:Todas las proporciones de la población son iguales.

H_1: No todas las proporciones de la población son iguales.

La estimación combinada de la proporción muestral “p” se calcula de la siguiente manera:

$$p = \frac{x_1 + x_2 + x_3 + \cdots x_n}{n_1 + n_2 + n_3 + \cdots n_n}$$

En una muestra se puede dar un conjunto de sucesos, los cuales ocurren con frecuencias observadas "o"(las que se observa directamente) y frecuencias esperadas o teóricas "e" (las que se calculan de acuerdo a las leyes de probabilidad).

La frecuencia esperada "e" se calcula así: $e = p \cdot o_{total}$

p = proporción muestral

o_{total} = frecuencia total observada

El estadístico de prueba es

$$\chi^2_{prueba} = \frac{(o_1 - e_1)^2}{e_1} + \frac{(o_2 - e_2)^2}{e_2} + \frac{(o_3 - e_3)^2}{e_3} + \cdots \frac{(o_n - e_n)^2}{e_n}$$

$$\chi^2_{prueba} = \sum \frac{(o_i - e_i)^2}{e_i}$$

Donde:
χ es la letra griega ji
χ^2 se lee ji cuadrado

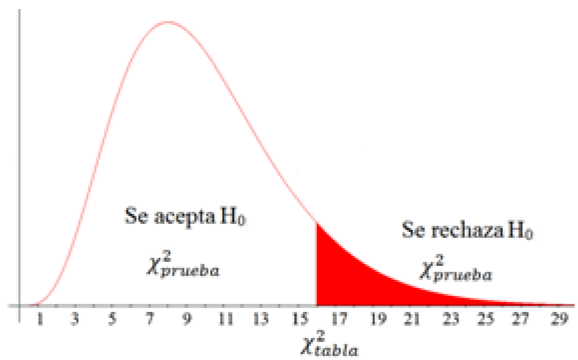
Por lo tanto el valor estadístico de prueba para este caso es la prueba *ji cuadrado* o conocida también como *chi cuadrado*

Como sucede con las distribuciones t y F, la distribución ji cuadrado tiene una forma que depende del número de grados de libertad asociados a un determinado problema.

Para obtener un valor crítico (valor que deja un determinado porcentaje de área en la cola) a partir de una tabla de ji cuadrado, se debe seleccionar un nivel de significación y determinar los grados de libertad para el problema que se esté resolviendo.

Los grados de libertad son una función del número de casillas en una tabla de $2 \cdot k$. Es decir, los grados de libertad reflejan el tamaño de la tabla. Los grados de libertad de la columna son el número de filas (categorías) menos 1, o bien, $r - 1$ .Los grados de libertad de cada fila es igual al número de columnas (muestras) menos 1, o bien, $k - 1$. El efecto neto es que el número de grados de libertad para la tabla es el producto de (número de filas -1) por (número de columnas -1), o bien, $(r-1)(k-1)$.Por lo tanto con 2 filas y 4 columnas, los grados de libertad son $(2-1)(4-1) = 3$

La prueba ji cuadrado requiere la comparación del χ^2_{prueba} con el χ^2_{tabla}. Si el valor estadístico de prueba es menor que el valor tabular, la hipótesis nula es aceptada, caso contrario, H_0 es rechazada.



Nota: Un valor estadístico de χ^2_{prueba}menor que el valor crítico χ^2_{tabla} o igual a él se considera como prueba de la variación casual en donde H_0 es aceptada.

Ejemplos ilustrativos:

1) El siguiente valor $3 \cdot 4$ representa el tamaño de una tabla $r \cdot k$.

Determine el número de grados de libertad y obtenga el valores crítico en el niveles 0,05 se significación.

Solución:

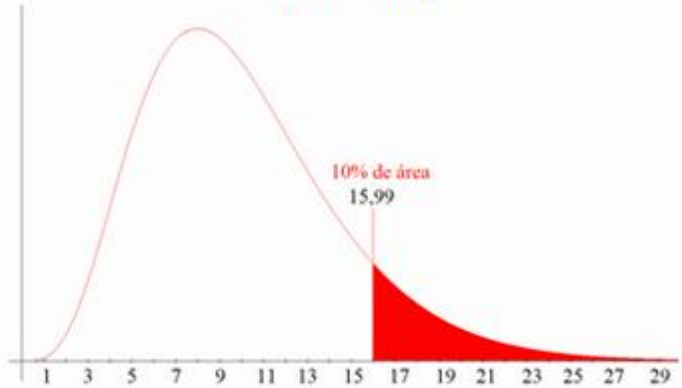
Los grados de libertad se calculan aplicando la fórmula:

$$Grados\ de\ libertad = (r-1)(k-1)$$

$$Grados\ de\ libertad = (3-1)(4-1) = 12$$

TABLA Nº 6

DISTRIBUCIÓN χ^2



Ejemplo:

Para 10 grados de libertad

$$P(\chi^2 > 15,99) = 0,10 = 10\%$$

	0,995	0,990	0,975	0,950	0,900	0,750	0,500	0,250	0,100	**0,050**	0,025	0,010	0,005
1	0,000	0,000	0,001	0,004	0,016	0,102	0,455	1,323	2,706	3,841	5,024	6,635	7,879
2	0,010	0,020	0,051	0,103	0,211	0,575	1,386	2,773	4,605	5,991	7,378	9,210	10,597
3	0,072	0,115	0,216	0,352	0,584	1,213	2,366	4,108	6,251	7,815	9,348	11,345	12,838
4	0,207	0,297	0,484	0,711	1,064	1,923	3,357	5,385	7,779	9,488	11,143	13,277	14,860
5	0,412	0,554	0,831	1,145	1,610	2,675	4,351	6,626	9,236	11,070	12,833	15,086	16,750
6	0,676	0,872	1,237	1,635	2,204	3,455	5,348	7,841	10,645	12,592	14,449	16,812	18,548
7	0,989	1,239	1,690	2,167	2,833	4,255	6,346	9,037	12,017	14,067	16,013	18,475	20,278
8	1,344	1,646	2,180	2,733	3,490	5,071	7,344	10,219	13,362	15,507	17,535	20,090	21,955
9	1,735	2,088	2,700	3,325	4,168	5,899	8,343	11,389	14,684	16,919	19,023	21,666	23,589
10	2,156	2,558	3,247	3,940	4,865	6,737	9,342	12,549	15,987	18,307	20,483	23,209	25,188
11	2,603	3,053	3,816	4,575	5,578	7,584	10,341	13,701	17,275	19,675	21,920	24,725	26,757
12	3,074	3,571	4,404	5,226	6,304	8,438	11,340	14,845	18,549	**21,026**	23,337	26,217	28,300

Con lectura en la tabla con 12 grados de libertad y 0,05 de área se obtiene $\chi^2_{tabla} = 21,026$

3.4.4.- Errores Tipo I y tipo II de significancia, prueba de una y dos colas

¿Qué son los errores de tipo I y tipo II?

Ninguna prueba de hipótesis es 100% cierta. Puesto que la prueba se basa en probabilidades, siempre existe la posibilidad de llegar a una conclusión incorrecta. Cuando usted realiza una prueba de hipótesis, puede cometer dos tipos de error: tipo I y tipo II. Los riesgos de estos dos errores están inversamente relacionados y se determinan según el nivel de significancia y la potencia de la prueba. Por lo tanto, usted debe determinar qué error tiene consecuencias más graves para su situación antes de definir los riesgos.

Error de tipo I

Si usted rechaza la hipótesis nula cuando es verdadera, comete un error de tipo I. La probabilidad de cometer un error de tipo I es α, que es el nivel de significancia que usted establece para su prueba de hipótesis. Un α de 0.05 indica que usted está dispuesto a aceptar una probabilidad de 5% de estar equivocado al rechazar la hipótesis nula. Para reducir este riesgo, debe utilizar un valor menor para α. Sin embargo, usar un valor menor para alfa significa que usted tendrá menos probabilidad de detectar una diferencia si esta realmente existe.

Error de tipo II

Cuando la hipótesis nula es falsa y usted no la rechaza, comete un error de tipo II. La probabilidad de cometer un error de tipo II es β, que depende de la potencia de la prueba. Puede reducir el riesgo de cometer un error de tipo II al asegurarse de que la prueba tenga suficiente potencia. Para ello, asegúrese de que el tamaño de la muestra sea lo suficientemente grande como para detectar una diferencia práctica cuando esta realmente exista.

La probabilidad de rechazar la hipótesis nula cuando es falsa es igual a $1-\beta$. Este valor es la potencia de la prueba.

	Verdad acerca de la población	
Decisión basada en la muestra	H_0 es verdadera	H_0 es falsa
No rechazar H_0	Decisión correcta (probabilidad = 1 - α)	**Error tipo II -** no rechazar H_0 cuando es falsa (probabilidad = β)
Rechazar H_0	**Error tipo I -** rechazar H_0 cuando es verdadera (probabilidad = α)	Decisión correcta (probabilidad = 1 - β)

Ejemplo de error de tipo I y tipo II

Para entender la interrelación entre los errores de tipo I y tipo II, y para determinar cuál error tiene consecuencias más graves para su situación, considere el siguiente ejemplo.

Un investigador médico desea comparar la efectividad de dos medicamentos. Las hipótesis nula y alternativa son:

- Hipótesis nula (H_0): $\mu_1 = \mu_2$

  Los dos medicamentos tienen la misma eficacia.

- Hipótesis alternativa (H_1): $\mu_1 \neq \mu_2$

  Los dos medicamentos no tienen la misma eficacia.

Un error de tipo I se produce si el investigador rechaza la hipótesis nula y concluye que los dos medicamentos son diferentes cuando, en realidad, no lo son. Si los medicamentos tienen la misma eficacia, el investigador podría considerar que este error no es muy grave, porque de todos modos los pacientes se beneficiarían con el mismo nivel de eficacia independientemente del medicamento que tomen. Sin embargo, si se produce un error de tipo II, el investigador no rechaza la hipótesis nula cuando debe rechazarla. Es decir, el investigador concluye que los medicamentos son iguales cuando en realidad son diferentes. Este error puede poner en riesgo la vida de los pacientes si

UNIDAD IV

PROBABILIDAD Y DISTRIBUCIONES DE PROBABILIDAD

4.1 EL PAPEL DE LA PROBABILIDAD EN ESTADÍSTICA

La probabilidad y la estadística están relacionadas en una forma importante. La probabilidad se emplea como herramienta; permite que usted evalúe la confiabilidad de sus conclusiones acerca de la población cuando tenga sólo información muestral. Considere estas situaciones:

• Cuando lance al aire una sola moneda, verá cara (H) o cruz (T). Si lanza la moneda varias veces al aire, va a generar un número infinitamente grande de caras o cruces, es decir, toda la población. ¿Qué aspecto tiene esta población? Si la moneda es imparcial, entonces la población debe contener 50% de H y 50% de T. Ahora lance al aire la moneda una vez más. ¿Cuál es la probabilidad de que resulte una cara? Casi todos dirían que la "probabilidad" es 1/2.

• Ahora suponga que no está usted seguro de que la moneda sea imparcial, esto es, no sabe con certeza si la composición de la población es 50-50 y decide hacer un experimento sencillo. Lanza al aire la moneda n   10 veces y observa 10 caras consecutivas. ¿Puede concluir que la moneda es imparcial? Es probable que no, porque si así fuera, observar 10 caras en fi la sería muy improbable; esto es, la "probabilidad" sería muy pequeña. Es más probable que la moneda esté "cargada".

Al igual que en el ejemplo de lanzar al aire una moneda, los expertos en estadística usan la probabilidad en dos formas. Cuando la población es conocida, se usa la probabilidad para describir la probabilidad de observar un resultado muestral en particular. Cuando la población es desconocida y sólo se dispone de una muestra de esa población, la probabilidad se usa para hacer enunciados acerca de la composición de la población, es decir, hacer inferencias estadísticas.

En los capítulos 4-7 usted verá numerosas formas diferentes para calcular probabilidades. Supondrá que la población es conocida y calculará la probabilidad de observar varios resultados muestrales. Una vez que empiece a usar la probabilidad para inferencia estadística en el capítulo 8, la población será desconocida y usted usará su conocimiento de probabilidad para hacer inferencias confiables a partir de información muestral. Empecemos con algunos ejemplos sencillos para ayudarle a captar conceptos básicos de probabilidad.

4.2 EVENTOS Y EL ESPACIO MUESTRAL

Se obtienen datos al observar ya sea eventos no controlados en la naturaleza o situaciones controladas en un laboratorio. Usamos el término experimento para describir cualquiera de los dos métodos de recolección de datos.

Definición Un experimento es el proceso mediante el cual se obtiene una observación (o medición).

La observación o medición generada por un experimento puede o no producir un valor numérico. A continuación, veamos algunos ejemplos de experimentos:

• Registrar la calificación de un examen

• Medir la cantidad de lluvia diaria

• Entrevistar a un dueño de casa para obtener su opinión sobre un reglamento para distribuir por zonas un área verde

• Probar una tarjeta de circuito impreso para determinar si es un producto defectuoso o aceptable.

• Lanzar al aire una moneda y observar el lado que aparece.

Cuando se realiza un experimento, lo que observamos es un resultado llamado evento simple, con frecuencia denotado por la mayúscula E con un subíndice.

Definición Un evento simple es el resultado que se observa en una sola repetición del experimento.

Experimento: Lance un dado y observe el número que aparece en la cara superior. Haga una lista de los eventos sencillos del experimento.

Solución Cuando el dado se lanza una vez, hay seis posibles resultados. Hay los eventos sencillos citados a continuación:

Evento E1: observar un 1 Evento E4: observar un 4

Evento E2: observar un 2 Evento E5: observar un 5

Evento E3: observar un 3 Evento E6: observar un 6

Ahora podemos defi nir un evento como un conjunto de eventos sencillos, a menudo denotado por una letra mayúscula.

Definición Un evento es un conjunto de eventos sencillos.

Podemos definir los eventos A y B para el experimento de lanzar al aire un dado:

A: observar un número impar

B: observar un número menor a 4

Como el evento A se presenta si la cara superior es 1, 3 o 5, es un conjunto de tres eventos sencillos y escribimos A {E1, E3, E5}. Del mismo modo, el evento B ocurre si la cara superior es 1, 2 o 3 y está definido como una serie o conjunto de estos tres eventos sencillos: B {E1, E2, E3}.

A veces, cuando ocurre un evento, significa que no puede ocurrir otro. Definición Dos eventos son mutuamente excluyentes si, cuando ocurre un evento, los otros no pueden ocurrir y viceversa.

En el experimento de lanzar al aire un dado, los eventos A y B no son mutuamente excluyentes, porque tienen dos resultados en común, si el número de la cara superior del dado es 1 o 3. Ambos eventos, A y B, ocurrirán si se observa E1 o E3 cuando se realiza el experimento. En contraste, los seis eventos simples E1, E2, . . . , E6 forman un conjunto de todos los resultados mutuamente excluyentes del experimento. Cuando el experimento se realiza una vez, puede ocurrir uno y sólo uno de estos eventos sencillos.

Definición El conjunto de todos los eventos sencillos se denomina espacio muestral, S.

4.3 CÁLCULO DE PROBABILIDADES CON EL USO DE EVENTOS SENCILLOS

La probabilidad de un evento A es una medida de nuestra creencia de que el evento A ocurrirá. Una manera práctica de interpretar esta medida es con el concepto de frecuencia relativa. Recuerde del capítulo 1 que si un experimento se realiza n veces, entonces la frecuencia relativa de un suceso particular, por ejemplo A, es

Frecuencia relativa _ F_re_c_un_e_ n_c_i_a

donde la frecuencia es el número de veces que ocurrió el evento A. Si hacemos que el número n de repeticiones del experimento se haga cada vez más grande (n ▯), en última instancia se genera toda la población. En ésta, la frecuencia relativa del evento A se define como la probabilidad del evento A; esto es,

P(A) lim

n ▯

_F_re_c_un_e_ n_c_i_a

Como P(A) se comporta como una frecuencia relativa, P(A) debe ser una proporción que se encuentre entre 0 y 1; P(A) 0 si el evento A nunca ocurre, y P(A) 1 si el evento A siempre ocurre. Cuanto más cercano sea P(A) a 1, es más probable que A ocurra.

Por ejemplo, si se lanza al aire un dado balanceado de seis caras un número de veces infinito, se esperaría que la frecuencia relativa para cualesquiera de los seis valores, x 1, 2, 3, 4, 5, 6, fuera 1/6. Sobra decir que sería muy lento, si no imposible, repetir un experimento un número infi nito de veces. Por esta razón, hay métodos alternativos para calcular probabilidades que hacen uso del concepto de frecuencia relativa.

Una consecuencia importante de la definición de frecuencia relativa de una probabilidad involucra a eventos sencillos. Como los eventos sencillos son mutuamente excluyentes, sus probabilidades deben satisfacer dos condiciones.

REQUISITOS PARA PROBABILIDADES DE UN EVENTO SIMPLE

• Cada probabilidad debe estar entre 0 y 1.

• La suma de las probabilidades de todos los eventos sencillos en S igual a 1.

Cuando es posible escribir los eventos sencillos asociados con un experimento y determinar sus probabilidades respectivas, podemos hallar la probabilidad de un evento A si sumamos las probabilidades de todos los eventos sencillos contenidos en el evento A.

Definición La probabilidad de un evento A es igual a la suma de las probabilidades de los eventos sencillos contenidos en A.

4.4 REGLAS ÚTILES DE CONTEO

Suponga que un experimento comprende un gran número N de eventos simples y que usted sabe que todos esos eventos son igualmente probables. Entonces cada evento simple tiene una probabilidad 1/N y la probabilidad de un evento A se puede calcular como

P(A)

_n_A

N

donde nA es el número de eventos simples que resultan en el evento A. En esta sección, presentamos tres reglas sencillas que se pueden usar para contar ya sea N, el número de eventos simples del espacio muestral, o nA, el número de eventos simples del evento A.

Una vez que haya obtenido estas cuentas, puede hallar P(A) sin en realidad hacer una lista de todos los eventos simples.

LA REGLA mn

Considere un experimento que se realiza en dos etapas. Si la primera etapa se puede efectuar en m formas y, para cada una de éstas, la segunda etapa se puede lograr en n formas, entonces hay mn formas para efectuar el experimento.

Por ejemplo, supongamos que usted puede ordenar un auto en uno de tres estilos y en uno de cuatro colores de pintura. Para averiguar cuántas opciones hay disponibles, puede considerar primero escoger uno de los m 3 estilos y luego seleccionar uno de los n 4 colores de pintura. Con el uso de la Regla mn, como se muestra en la fi gura 4.5, tiene mn (3)(4) 12 posibles opciones.

4.5 RELACIONES DE EVENTO Y REGLAS DE PROBABILIDAD

Hay veces en que el evento de interés se puede formar como una combinación de algunos otros eventos. Sean A y B dos eventos defi nidos en el espacio muestral S. Aquí hay tres relaciones importantes entre eventos.

Definición La unión de los eventos A y B, denotada por A B, es el evento en que ocurren A o B o ambos.

Definición La intersección de eventos A y B, denotada por A B, es el evento en que ocurren A y B.†

Definición El complemento de un evento A, denotado por Ac, es el evento en que A no ocurre.

Las figuras 4.8, 4.9 y 4.10 muestran representaciones del diagrama de Venn de A B, A B y Ac, respectivamente. Cualquier evento simple en el área sombreada es un posible resultado que aparece en el evento apropiado. Una forma de hallar las probabilidades de la unión, la intersección o el complemento es sumar las probabilidades de todos los eventos simples asociados.

CONSEJO
Intersección ⇔ "ambos . . . y" o sólo "y".
Unión ⇔ "uno de dos . . . o ambos" o sólo "o".

FIGURA 4.8
Diagrama de Venn de $A \cup B$

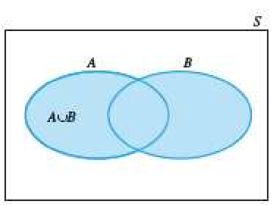


FIGURA 4.9
Diagrama de Venn $A \cap B$

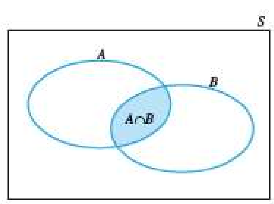
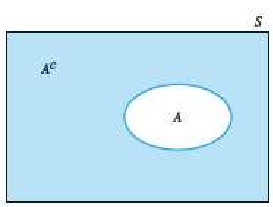


FIGURA 4.10
El complemento de un evento



4.6 INDEPENDENCIA, PROBABILIDAD CONDICIONAL Y LA REGLA DE LA MULTIPLICACIÓN

Hay una regla de la probabilidad que se puede usar para calcular la probabilidad de la intersección de varios eventos, pero esta regla depende del importante concepto estadístico de eventos independientes o dependientes.

Definición Se dice que dos eventos, A y B, son independientes si y sólo si la probabilidad

del evento B no está influenciada o cambiada por el suceso del evento A, o viceversa.

Daltonismo Suponga que un observador ve el género de una persona y si ésta no distingue los colores rojo y verde. ¿Cambia la probabilidad de que una persona sea daltónica, dependiendo de si es hombre o no? Defina dos eventos:

A: la persona es hombre

B: la persona es daltónica

En este caso, como el daltonismo es una característica relacionada con el sexo masculino, la probabilidad de que un hombre sea daltónico será mayor que la probabilidad de que una persona escogida de la población general sea daltónica. La probabilidad del evento B, que una persona sea daltónica, depende de si ha ocurrido o no ha ocurrido el evento A, que la persona sea hombre. Decimos que A y B son eventos dependientes.

Tirar dados. Por el contrario, considere tirar un solo dado dos veces y defina dos eventos:

A: observar un 2 en el primer tiro

B: observar un 2 en el segundo tiro

Si el dado es imparcial, la probabilidad del evento A es P(A)   1/6. Considere la probabilidad

del evento B. Ya sea que el evento A haya ocurrido o no haya ocurrido, la probabilidad de observar un 2 en el segundo tiro todavía es 1/6. Podríamos escribir:

P(B dado que A ocurrió)   1/6

P(B dado que A no ocurrió)   1/6

Como la probabilidad del evento B no ha cambiado por el suceso del evento A, decimos que A y B son eventos independientes.

La probabilidad de un evento A, dado que el evento B ha ocurrido, se denomina probabilidad

condicional de A, dado que B ha ocurrido, denotada por P(A   B). La barra vertical se lee “dada” y los eventos que aparecen a la derecha de la barra son aquellos que se sabe han ocurrido. Usaremos estas probabilidades para calcular la probabilidad de que A y B ocurran cuando se realice el experimento.

REGLA GENERAL DE LA MULTIPLICACIÓN

La probabilidad de que A y B ocurran cuando el experimento se realiza es

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

o

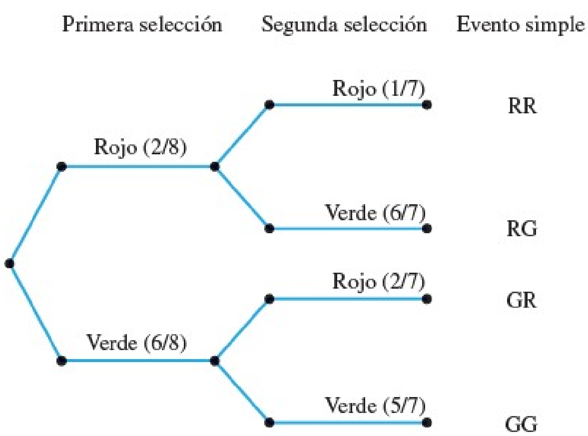
$$P(A \cap B) = P(B)P(A|B)$$

En un experimento de preferencia de color, ocho juguetes se ponen en un recipiente. Los juguetes son idénticos excepto por el color, dos son rojos y seis son verdes. Se pide a un niño que escoja dos juguetes al azar. ¿Cuál es la probabilidad de que el niño escoja los dos juguetes rojos?

Solución Se puede visualizar el experimento usando un diagrama de árbol como se muestra en la figura. Defina los eventos siguientes:

R: se escoge juguete rojo

G: se escoge juguete verde



El evento A (ambos juguetes son rojos) se puede construir como la intersección de dos eventos:

A (R en la primera selección) $\cap$ (R en la segunda selección)

Como sólo hay dos juguetes rojos en el recipiente, la probabilidad de escoger el rojo en la primera selección es 2/8. No obstante, una vez que haya sido escogido este juguete rojo, la probabilidad del rojo en la segunda selección depende del resultado de la primera selección (véase la fi gura 4.13). Si la primera selección fue un juguete rojo, la probabilidad de escoger un segundo juguete rojo es sólo 1/7 porque hay sólo un juguete rojo entre los siete restantes.

Si la primera selección fue verde, la probabilidad de escoger rojo en la segunda selección es 2/7 porque hay dos juguetes rojos entre los siete restantes.

Usando esta información y la Regla de la multiplicación, se puede hallar la probabilidad del evento A.

P(A)   P(R en la primera selección □ R en la segunda selección)

P(R en la primera selección) P(R en la segunda selección)   R en la primera)

2

___

56

1

___

28

A veces es necesario usar la Regla de la multiplicación en una forma ligeramente diferente,

de modo que se puede calcular la probabilidad condicional, P(A□ B). Sólo reacomode los términos de la Regla de la multiplicación.

PROBABILIDADES CONDICIONALES

La probabilidad condicional del evento A, dado que el evento B ha ocurrido, es

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{si} \quad P(B) \neq 0$$

La probabilidad condicional del evento B, dado que el evento A ha ocurrido, es

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \quad \text{si} \quad P(A) \neq 0$$

Daltonismo, continúa Suponga que en la población general, hay 51% de hombres y 49% de mujeres, y que las proporciones de hombres y mujeres daltónicos se muestran en la siguiente tabla de probabilidad:

Hombres (B) Mujeres (BC) Total

Daltónico (A) .04 .002 .042

No daltónico (AC) .47 .488 .958

Total .51 .49 1.00

Si una persona se escoge al azar de entre esta población y se encuentra que es hombre (evento B), ¿cuál es la probabilidad de que el hombre sea daltónico (evento A)? Si sabemos que el evento B ha ocurrido, debemos restringir nuestra atención a sólo 51% de la población que es de hombres. La probabilidad de ser daltónico, dado que la persona es hombre, es 4% de 51%, o sea

P(A B)

P(A B)

____

P(B)

_.0_4_

.51

.078

¿Cuál es la probabilidad de ser daltónico, dado que la persona es mujer? Ahora estamos restringidos a sólo el 49% de la población que es de mujeres y

P(A BC)

P(A BC)

____

P(BC)

_.0_0_2_

.49

.004

Observe que la probabilidad del evento A cambió, dependiendo de si el evento B ocurrió. Esto indica que estos dos eventos son dependientes.

Cuando dos eventos son independientes, es decir, si la probabilidad del evento B es igual, ya sea que el evento A haya o no haya ocurrido, entonces el evento A no afecta al

evento B y entonces

P(B A) P(B)

Ahora se puede simplificar la Regla de la multiplicación.

LA REGLA DE LA MULTIPLICACIÓN PARA EVENTOS INDEPENDIENTES

Si dos eventos A y B son independientes, la probabilidad de que ocurran A y B es

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

Del mismo modo, si A, B y C son eventos mutuamente independientes (todos los pares de eventos son independientes), entonces la probabilidad de que A, B y C ocurran es

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

Tiros de monedas en juegos de fútbol Un equipo de fútbol interviene en dos periodos de tiempo extra durante un juego determinado, de modo que hay tres tiros de monedas al aire. Si la moneda es imparcial, ¿cuál es la probabilidad de que pierdan los tres tiros?

Solución Si la moneda es imparcial, el evento se puede describir en tres pasos:

A: perder el primer tiro

B: perder el segundo tiro

C: perder el tercer tiro

Como los tiros son independientes y como P(gana) = P(pierde) = .5 para cualquiera de los tres tiros,

$P(A \cap B \cap C) = P(A)P(B)P(C) = (.5)(.5)(.5) = .125$

¿Cómo se puede verificar si los dos eventos son independientes o dependientes? La solución más fácil es redefinir el concepto de independencia en un modo más formal.

VERIFICACIÓN DE INDEPENDENCIA

Se dice que dos eventos A y B son **independientes** si y sólo si

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

o bien,

$$P(B|A) = P(B)$$

De otro modo, se dice que los eventos son **dependientes**.

Tire al aire dos monedas y observe el resultado. Defina estos eventos:

A: cara en la primera moneda

B: cruz en la segunda moneda

¿Los eventos A y B son independientes?

Solución De los ejemplos previos, sabemos que S {HH, HT, TH, TT}. Utilice estos cuatro eventos simples para hallar

$P(A)$ $\frac{1}{2}$, $P(B)$ $\frac{1}{2}$ y $P(A \cap B)$ $\frac{1}{4}$.

Como $P(A)P(B)$ $\frac{1}{4}$ y $P(A \cap B)$ $\frac{1}{4}$, tenemos $P(A)P(B)$ $P(A \cap B)$ y los dos eventos deben ser independientes.

4.7 REGLA DE BAYES (OPCIONAL)

Daltonismo Reconsideremos el experimento referente a daltonismo visto en la sección 4.6. Observe que los dos eventos

B: la persona seleccionada es un hombre

BC: la persona seleccionada es una mujer

tomados juntos conforman el espacio muestral S, formado de hombres y mujeres. Como los daltónicos pueden ser hombres o mujeres, el evento A, que es que una persona sea daltónica, está formado de los eventos simples que estén en A y además en B y de los eventos simples que estén en A y además en BC. Como estas dos intersecciones son mutuamente excluyentes, se puede escribir el evento A como

$$A = (A \cap B) \cup (A \cap BC)$$

y

$$P(A) = P(A \cap B) + P(A \cap BC)$$

$$= .04 + .002 = .042$$

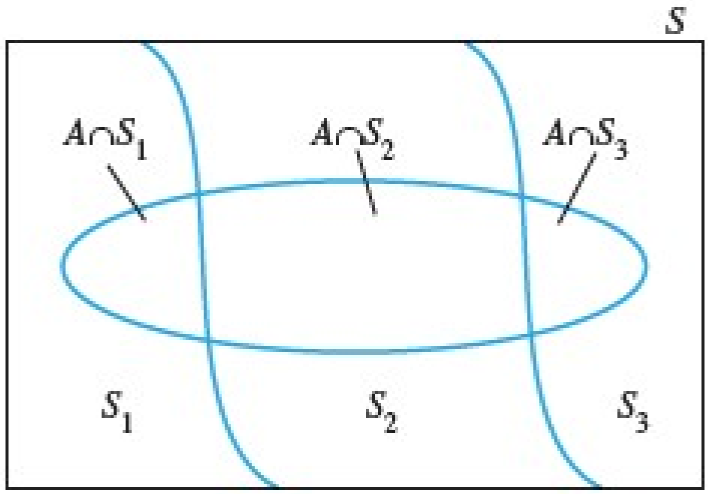
Suponga ahora que el espacio muestral se puede dividir en k subpoblaciones, S1, S2, S3, . . . , Sk, que, al igual que en el ejemplo de daltonismo, son mutuamente excluyentes y exhaustivos; esto es, tomados juntos conforman todo el espacio muestral. De un modo semejante, se puede expresar un evento A como

$$A = (A \cap S1) \cup (A \cap S2) \cup (A \cap S3) \cup \cdots \cup (A \cap Sk)$$

Entonces

$$P(A) = P(A \cap S1) + P(A \cap S2) + P(A \cap S3) + \cdots + P(A \cap Sk)$$

Esto está ilustrado para k = 3 en la fi gura 4.14.



Se puede avanzar un paso más y usar la Regla de la multiplicación para escribir P(A Si) como P(Si)P(A Si), para i 1, 2, . . . , k. El resultado se conoce como la Ley de probabilidad total.

LEY DE PROBABILIDAD TOTAL

Dado un conjunto de eventos $S_1, S_2, S_3, \ldots, S_k$ que son mutuamente excluyentes y exhaustivos y un evento A, la probabilidad del evento A se puede expresar como

$$P(A) = P(S_1)P(A|S_1) + P(S_2)P(A|S_2) + P(S_3)P(A|S_3) + \cdots + P(S_k)P(A|S_k)$$

Los zapatos tenis ya no son sólo para jóvenes. De hecho, casi todos los adultos tienen varios pares de ellos. La tabla 4.6 da la fracción de adultos estadounidenses de 20 años de edad o más que tienen cinco o más pares de zapatos tenis en buen estado, junto con la fracción de adultos estadounidenses de 20 años o más en cada uno de los cinco grupos de edad. 5 Use la Ley de probabilidad total para determinar la probabilidad incondicional de un adulto de 20 años de edad o más que tenga cinco o más pares de zapatos tenis en buen estado.

Tabla de probabilidad

	Grupos y edades				
	G_1 20–24	G_2 25–34	G_3 35–49	G_4 50–64	G_5 ≥ 65
Fracción con ≥ 5 pares	.26	.20	.13	.18	.14
Fracción de adultos de 20 años o más	.09	.20	.31	.23	.17

Solución Sea A el evento de que una persona seleccionada al azar de entre la población de adultos estadounidenses de 18 años de edad y más tenga cinco o más pares de zapatos tenis en buen estado. Con $G_1, G_2, \ldots, G_5$ represente el evento de que la persona seleccionada pertenezca a cada uno de los cinco grupos de edades, respectivamente. Como los cinco grupos son *exhaustivos*, se puede escribir el evento A como

$$A = (A \cap G_1) \cup (A \cap G_2) \cup (A \cap G_3) \cup (A \cap G_4) \cup (A \cap G_5)$$

Usando la Ley de probabilidad total, se puede hallar la probabilidad de A como

$$\begin{aligned} P(A) &= P(A \cap G_1) + P(A \cap G_2) + P(A \cap G_3) + P(A \cap G_4) + P(A \cap G_5) \\ &= P(G_1)P(A|G_1) + P(G_2)P(A|G_2) + P(G_3)P(A|G_3) \\ &\quad + P(G_4)P(A|G_4) + P(G_5)P(A|G_5) \end{aligned}$$

De las probabilidades de la tabla 4.6,

$$\begin{aligned} P(A) &= (.09)(.26) + (.20)(.20) + (.31)(.13) + (.23)(.18) + (.17)(.14) \\ &= .0234 + .0400 + .0403 + .0414 + .0238 = .1689 \end{aligned}$$

La probabilidad incondicional de que una persona, seleccionada de entre la población de adultos estadounidenses de 20 años de edad y más, tenga al menos cinco pares de zapatos tenis en buen estado es de alrededor de .17. Observe que la Ley de probabilidad total es un promedio ponderado de las probabilidades dentro de cada grupo, con pesos .09, .20, .31, .23 y .17, que refleja los tamaños relativos de los grupos.

4.8 VARIABLES ALEATORIAS DISCRETAS Y SUS DISTRIBUCIONES DE PROBABILIDAD

Las variables se definieron como características que cambian o varían con el tiempo y/o para diferentes personas u objetos bajo consideración. Las variables cuantitativas generan datos numéricos, en tanto que las variables cualitativas generan datos categóricos. No obstante, incluso las variables cualitativas pueden generar datos numéricos si las categorías son codificadas numéricamente para formar una escala. Por ejemplo, si se lanza al aire una sola moneda, el resultado cualitativo podría registrarse como “0” si es cara o como “1” si es cruz.

4.9 VARIABLES ALEATORIAS

Una variable x valuada numéricamente varía o cambia, dependiendo del resultado particular del experimento que se mida. Por ejemplo, suponga que se tira un dado y se mide x, el número observado en la cara superior. La variable x puede tomar cualquiera de seis valores: 1, 2, 3, 4, 5, 6, dependiendo del resultado aleatorio del experimento. Por esta razón, la variable x se conoce como variable aleatoria.

Definición Una variable x es variable aleatoria si el valor que toma, correspondiente al resultado de un experimento, es una probabilidad o evento aleatorio.

Se pueden considerar numerosos ejemplos de variables aleatorias:

• x   Número de defectos en una pieza de mueble seleccionada al azar

• x   Califi cación de examen de aptitud escolar (SAT) para un solicitando universitario seleccionado al azar

• x   Número de llamadas telefónicas recibidas por una línea directa de intervención en crisis durante un periodo seleccionado al azar

Al igual que en el capítulo 1, las variables aleatorias cuantitativas se clasifican ya sea como discretas o como continuas, de acuerdo con los valores que x pueda tomar. Es importante distinguir entre variables aleatorias discretas y continuas, porque se usan técnicas diferentes para describir sus distribuciones. Nos concentramos en variables aleatorias discretas en el resto de este capítulo

4.10 DISTRIBUCIONES DE PROBABILIDAD

Usted aprendió a construir la distribución de frecuencia relativa para un conjunto de mediciones numéricas en una variable x. La distribución dio esta información acerca de x:

• ¿Qué valores de x se presentaron?

• ¿Con qué frecuencia se presentó cada valor de x?

Usted también aprendió a usar la media y desviación estándar para medir el centro y variabilidad de este conjunto de datos.

En este capítulo, definimos la probabilidad como el valor limitando de la frecuencia relativa cuando el experimento se repite una y otra vez. Ahora definimos la distribución de probabilidad para una variable aleatoria x como la distribución de frecuencia relativa construida para toda la población de mediciones.

Definición La distribución de probabilidad para una variable aleatoria discreta es una fórmula, tabla o gráfi ca que da los posibles valores de x, y la probabilidad p(x) asociada con cada valor de x.

Los valores de x representan eventos numéricos mutuamente excluyentes. Sumar p(x) sobre todos los valores de x es equivalente a sumar las probabilidades de todos los eventos simples y por tanto es igual a 1.

REQUISITOS PARA UNA DISTRIBUCIÓN DE PROBABILIDAD DISCRETA

- $0 \leq p(x) \leq 1$
- $\Sigma\, p(x) = 1$

4.11 LA MEDIA Y DESVIACIÓN ESTÁNDAR PARA UNA VARIABLE ALEATORIA DISCRETA

La distribución de probabilidad para una variable aleatoria discreta luce muy semejante a la distribución de frecuencia relativa vista en el capítulo 1. La diferencia es que la distribución de frecuencia relativa describe una muestra de n mediciones, en tanto que la distribución de probabilidad se construye como un modelo para toda la población de mediciones. Así como la media x y la desviación estándar s midieron el centro y dispersión de los datos muestrales, usted puede calcular medidas similares para describir el centro y dispersión de la población.

La media poblacional, que mide el valor promedio de x en la población, también se denomina valor esperado de la variable aleatoria x. Es el valor que se esperaría observar en promedio si el experimento se repite una y otra vez. La fórmula para calcular la media poblacional es más fácil de entender por ejemplo. Lance otra vez al aire esas dos monedas imparciales, y sea x el número de caras observado. Construimos esta distribución de probabilidad para x:

x	0	1	2
p(x)	¼	1/2	1/4

Suponga que el experimento se repite un gran número de veces, por ejemplo n 4 000 000 de veces. Intuitivamente, se esperaría observar alrededor de un millón de ceros, dos millones de números 1 y un millón de números dos. Entonces el valor promedio de x sería igual a

$$\frac{\text{Suma de mediciones}}{n} = \frac{1\,000\,000(0) + 2\,000\,000(1) + 1\,000\,000(2)}{4\,000\,000}$$

$$= \left(\frac{1}{4}\right)(0) + \left(\frac{1}{2}\right)(1) + \left(\frac{1}{4}\right)(2)$$

Observe que el primer término de esta suma es (0)p(0), el segundo es igual a (1)p(1) y el tercero es (2)p(2). El valor promedio de x, entonces, es

$$\Sigma xp(x) = 0 + \frac{1}{2} + \frac{2}{4} = 1$$

Este resultado da alguna justificación intuitiva para la definición del valor esperado de una variable aleatoria x discreta.

Definición Sea x una variable aleatoria discreta con distribución de probabilidad p(x). La media o valor esperado de x está dada como

m   E(x)   S xp(x)

donde los elementos se suman sobre todos los valores de la variable aleatoria x.

Podríamos usar un argumento similar para justificar las fórmulas para la varianza poblacional s2 y la desviación estándar de la población s. Estas medidas numéricas describen la dispersión o variabilidad de la variable aleatoria usando el "promedio" o "valor esperado" del cuadrado de las desviaciones de los valores x desde su media m.

Definición Sea x una variable aleatoria discreta con distribución de probabilidad p(x) y media m. La varianza de x es

s2   E[(x m)2]   S(x m)2p(x)

donde la sumatoria es sobre todos los valores de la variable aleatoria x.†

Definición La desviación estándar s de una variable aleatoria x es igual a la raíz cuadrada positiva de su varianza.