LIBRO

Estadística inferencial

Licenciatura en Psicologia

Cuarto Cuatrimestre

Septiembre- Diciembre

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.

Estadística inferencial.

Objetivo de la materia:

Analizar y aplicar conceptos, técnicas de la estadística inferencial en la solución de problemas en áreas de la Ingeniería. Así como la toma de decisiones, con base en los elementos teóricos adquiridos, que permitan reducir la incertidumbre.

INDICE

UNIDAD 4

PRUEBAS DE HIPÓTESIS CON DOS MUESTRAS Y VARIAS MUESTRAS DE DATOS NUMÉRICOS.

4.2 Distribuciones normal y t de Student.

4.3 Pruebas de significancia.

4.4 Comparación de dos muestras independientes: Pruebas t para las diferencias entre dos medias.

4.5 Prueba de Fisher para varianzas y de igualdad de las varianzas de dos poblaciones normales.

UNIDAD I

INTRODUCCIÓN A LA ESTADÍSTICA INFERENCIAL

1.1 Breve historia de la estadística.

1.1 BREVE HISTORIA DE LA ESTADISTICA

La palabra Estadística procede del vocablo "Estado", pues era función principal de los Gobiernos de los Estados establecer registros de población, nacimientos, defunciones, impuestos, cosechas... La necesidad de poseer datos cifrados sobre la población y sus condiciones materiales de existencia han debido hacerse sentir desde que se establecieron sociedades humanas organizadas.

Es difícil conocer los orígenes de la Estadística. Desde los comienzos de la civilización han existido formas sencillas de estadística, pues ya se utilizaban representaciones gráficas y otros símbolos en pieles, rocas, palos de madera y paredes de cuevas para contar el número de personas, animales o ciertas cosas.

Su origen empieza posiblemente en la isla de Cerdeña, donde existen monumentos prehistóricos pertenecientes a los Nuragas, las primeros habitantes de la isla; estos monumentos constan de bloques de basalto superpuestos sin mortero y en cuyas paredes de encontraban grabados toscos signos que han sido interpretados con mucha verosimilidad como muescas que servían para llevar la cuenta del ganado y la caza.

Hacia el año 3.000 a.C. los babilonios usaban ya pequeñas tablillas de arcilla para recopilar datos en tablas sobre la producción agrícola y los géneros vendidos o cambiados mediante trueque.

Los egipcios ya analizaban los datos de la población y la renta del país mucho antes de construir las pirámides. En los antiguos monumentos egipcios se encontraron interesantes documentos en que demuestran la sabia organización y administración de este pueblo; ellos llevaban cuenta de los movimientos poblacionales y continuamente hacían censos. Tal era su dedicación por llevar siempre una relación de todo que hasta tenían a la diosa Safnkit, diosa de los libros y las cuentas. Todo esto era hecho bajo la dirección del Faraón y fue a partir del año 3050 a.C.

En la Biblia observamos en uno de los libros del Pentateuco, bajo el nombre de Números, el censo que realizó Moisés después de la salida de Egipto. Textualmente dice: "Censo de las tribus: El día primero del segundo año después de la salida de Egipto, hablό Yavpe a Moisés en el desierto de Sinaí en el tabernáculo de la reunión, diciendo: "Haz un censo general de toda la asamblea de los hijos de Israel, por familias y por linajes, describiendo por cabezas los nombres de todos los varones aptos para el servicio de armas en Israel. En el libro bíblico Crónicas describe el bienestar material de las diversas tribus judías.

En China existían los censos chinos ordenados por el emperador Tao hacia el año 2.200 a.C.

Posteriormente, hacia el año 500 a.C., se realizaron censos en Roma para conocer la población existente en aquel momento. Se erigió la figura del censor, cuya misión consistía en controlar el número de habitantes y su distribución por los distintos territorios.

En la Edad Media, en el año 762, Carlomagno ordenó la creación de un registro de todas sus propiedades, así como de los bienes de la iglesia.

Después de la conquista normanda de Inglaterra en 1.066, el rey Guillermo I, el Conquistador, elaboró un catastro que puede considerarse el primero de Europa.

Los Reyes Católicos ordenaron a Alonso de Quintanilla en 1.482 el recuento de fuegos (hogares) de las provincias de Castilla.

En 1.662 un mercader de lencería londinense, John Graunt, publicó un tratado con las observaciones políticas y naturales, donde Graunt pone de manifiesto las cifras brutas de nacimientos y defunciones ocurridas en Londres durante el periodo 1.604-1.661, así como las influencias que ejercían las causas naturales, sociales y políticas de dichos acontecimientos. Puede considerarse el primer trabajo estadístico serio sobre la población.

Curiosamente, Graunt no conocía los trabajos de B. Pascal » (1.623-1.662) ni de C. Huygens (1.629-1.695) sobre estos mismos temas. Un poco más tarde, el astrónomo Edmund Halley (1.656- 1.742) presenta la primera tabla de mortalidad que se puede considerar como base de los estudios contemporáneos. En dicho trabajo se intenta establecer el precio de las anualidades a satisfacer a las compañías de seguros. Es decir, en Londres y en París se estaban construyendo, casi de manera simultánea, las dos disciplinas que actualmente llamamos estadística y probabilidad.

En el siglo XIX, la estadística entra en una nueva fase de su desarrollo con la generalización del método para estudiar fenómenos de las ciencias naturales y sociales. Galton » (1.822-1.911) y Pearson (1.857-1936) se pueden considerar como los padres de la estadística moderna, pues a ellos se debe el paso de la estadística deductiva a la estadística inductiva.

Los fundamentos de la estadística actual y muchos de los métodos de inferencia son debidos a R. A. Fisher. Se interesó primeramente por la eugenesia, lo que le conduce, siguiendo los pasos de Galton a la investigación estadística, sus trabajos culminan con la publicación de la obra Métodos estadísticos para investigaciones. En él aparece la metodología estadística tal y como hoy la conocemos.

A partir de mediados del siglo XX comienza lo que podemos denominar la estadística moderna, uno de los factores determinantes es la aparición y popularización de los computadores. El centro de gravedad de la metodología estadística se empieza a desplazar técnicas de computación intensiva aplicadas a grandes masas de datos, y se empieza a considerar el método estadístico como un proceso iterativo de búsqueda del modelo ideal

Las aplicaciones en este periodo de la Estadística a la Economía conducen a una disciplina con contenido propio: la Econometría. La investigación estadística en problemas militares durante la segunda guerra mundial y los nuevos métodos de programación matemática, dan lugar a la Investigación Operativa

1.2 Concepto de estadística.

La estadística se ocupa de la sistematización, recogida, ordenación y representación de los datos referentes a un fenómeno que presenta variabilidad o incertidumbre para su estudio metódico, con objeto de hacer previsiones sobre los mismos, tomar decisiones u obtener conclusiones. Teniendo en cuenta las funciones podemos considerar dos grandes áreas:

Estadística descriptiva: se organizan y resumen conjuntos de observaciones procedentes de una muestra o de la población total, en forma cuantitativa. Los procedimientos para una variable: índices de tendencia general, estadísticos de variabilidad y estadísticos de asimetría; y para dos variables: coeficientes de correlación y ecuaciones de regresión.

Estadística inferencial: se realizan inferencias acerca de una población basándose en los datos obtenidos a partir de una muestra. Los procedimientos: el cálculo de probabilidades.

Conceptos importantes: población es el conjunto de todos los elementos que cumplen una determinada característica objeto de estudio. Muestra es un subconjunto de una población.

Parámetro es una propiedad descriptiva (medida) de una población. Estadístico es una propiedad descriptiva (medida) de una muestra.

Las conclusiones obtenidas de una muestra sólo servirán para el total de una población si la muestra es representativa. Para asegurarnos que la muestra es representativa se utilizan métodos de muestreo probabilístico.

También existes las muestras no probabilísticas como por ejemplo la muestra de conveniencia o incidental.

1.3 Estadística descriptiva.

La estadística descriptiva es la rama de las Matemáticas que recolecta, representa y caracteriza un conjunto de datos (por ejemplo, edad de una población, altura de los estudiantes de una escuela, en los meses de verano, etc.) con el fin de describir apropiadamente las diversas características de ese conjunto.

La estadística descriptiva: se dedica a la descripción, visualización y resumen de datos originados a partir de los fenómenos de estudio. Los datos pueden ser resumidos numérica o gráficamente. Ejemplos básicos de parámetros estadísticos son: la media y la desviación estándar. Algunos ejemplos gráficos son: histograma, pirámide poblacional, gráfico circular, entre otros.

1.4 Estadística inferencial.

Se dedica a la generación de los modelos, inferencias y predicciones asociadas a los fenómenos en cuestión teniendo en cuenta la aleatoriedad de las observaciones. Se usa para modelar patrones en los datos y extraer inferencias acerca de la población bajo estudio. Estas inferencias pueden tomar la forma de respuestas a preguntas si/no (prueba de hipótesis), estimaciones de unas características numéricas (estimación), pronósticos de futuras observaciones, descripciones de asociación (correlación) o modelamiento de relaciones entre variables (análisis de regresión). Otras técnicas de modelamiento incluyen a nova, series de tiempo y minería de datos.

IMPORTANCIA DE LA ESTADISTICA INFERENCIAL

La Estadística Inferencial puede dar respuesta a muchas de las necesidades que la sociedad actual puede requerir. Su tarea fundamental es el análisis de los datos que se obtienen a partir de experimentos, con el objetivo de representar la realidad y conocerla. Permite la recolección de datos importantes para el estudio de situaciones que se presentan a diario y permite dar respuesta a los problemas de una forma útil y significativa.

La Estadística Inferencial se centra en tomar una pequeña muestra representativa de la población y a partir de ésta, infiere que el resto de la población tiene el mismo comportamiento.

En caso de que no sea factible realizar un estudio completo por cuestiones de tiempo, recursos o costo, se puede calcular un tamaño de muestra para medir solo algunos elementos de la población, posteriormente se infiere que el resto de la población se comporta igual que la muestra tomada

1.5 Breve introducción a la inferencia estadística.

El principal objetivo de la Estadística es inferir o estimar características de una población que no es completamente observable (o no interesa observarla en su totalidad) a través del análisis de una parte de ella a la que llamamos muestra. Las razones por las que generalmente se trabaja con muestras son principalmente:

- Económicas.

- Tiempo: si la población es muy grande llevaría tanto tiempo analizarla que incluso la característica de interés podría variar en ese período. Por ejemplo, la tasa de paro.

- Destrucción: la medición de cierta característica podría llevar a la destrucción del individuo. Por ejemplo, al estudiar la supervivencia de ciertos animales a un tratamiento.

Lo que se hace entonces es analizar la muestra y las conclusiones desde la muestra a la población. Ahora bien, para considerar válidas en la población las conclusiones obtenidas en la muestra, ésta ha de representar bien a la población (representativa). Por lo tanto, la selección de la muestra es de suma importancia, y para ello hay diversos métodos (métodos de muestreo). Cuando se intuye que la característica en estudio puede presentar valores homogéneos en la población, una forma de obtener una muestra representativa es eligiéndola al azar. A este método de selección de la muestra se le llama muestreo aleatorio simple y es el más sencillo.

La Inferencia Estadística se puede clasificar en inferencia paramétrica e inferencia no paramétrica.

La inferencia paramétrica tiene lugar cuando se conoce la distribución de la variable de estudio en la población, y el interés recae sobre los parámetros desconocidos de la misma.

La inferencia no paramétrica tiene lugar si no se conoce la distribución y sólo se suponen propiedades generales de la misma

1.6 Teoría de decisión en estadística.

Estudio formal sobre la toma de decisiones. Los estudios de casos reales, que se sirven de la inspección y los experimentos, se denominan teoría descriptiva de decisión; los estudios de la toma de decisiones racionales, que utilizan la lógica y la estadística, se llaman teoría preceptiva de decisión. Estos estudios se hacen más complicados cuando hay más de un individuo, cuando los resultados de diversas opciones no se conocen con exactitud y cuando las probabilidades de los distintos resultados son desconocidas. La teoría de decisión comparte características con la teoría de juegos, aunque en la teoría de decisión el „adversario' es la realidad en vez de otro jugador o jugadores.

Al hacer un análisis sobre esta teoría, y mirándola desde el punto de vista de un sistema, se puede decir que al tomar una decisión sobre un problema en particular, se debe tener en cuenta los puntos de dificultad que lo componen, para así empezar a estudiarlos uno a uno hasta obtener una solución que sea acorde a lo que se está esperando obtener de este, y si no, buscar otras soluciones que se acomoden a lo deseado.

La teoría de decisión, no solamente se puede ver desde el punto de vista de un sistema, sino en general, porque esta se utiliza a menudo para tomar decisiones de la vida cotidiana, ya que muchas personas piensan que la vida es como una de las teorías; La teoría del juego, que para poder empezarlo y entenderlo hay que saber jugarlo y para eso se deben conocer las reglas de este, para que no surjan equivocaciones al empezar la partida

1.7 Componentes de una investigación estadística.

El estudio estadístico de una situación con propósitos inferenciales se centra en dos conceptos fundamentales: población y muestra, los cuales serán definidos a continuación:

Población. Es el conjunto formado por todos los valores posibles que puede asumir, la variable objeto de estudio. Así por ejemplo, en un estudio sobre la preferencia de los votantes en una elección presidencial, la población consiste en todas las respuestas de los votantes registrados. Pero el término no sólo está asociado a la colección de seres humanos u organismos vivos; y tenemos así que, si se va a hacer una investigación de las ventas anuales de los supermercados, entonces las ventas anuales de todos los supermercados constituyen así mismo la población.

Es bueno tener en cuenta que el término población se interpreta de dos maneras cuando se hace un estudio estadístico, a saber:

1. La interpretación propia en el Análisis Estadístico, que corresponde a la que hemos presentado anteriormente.

2 Como el conjunto de objetos sobre los cuales actúa la variable considerada. Por tanto, no es extraño escuchar expresiones tales como, "se hizo un estudio de los niveles de ingreso de la población trabajadora colombiana", entendiéndose con ello que el elemento estadístico objeto de análisis fue el registro numérico de los ingresos.

Muestra. Es cualquier subconjunto de la población, escogido al seguir ciertos criterios de selección. La muestra es el elemento básico sobre el cual se fundamenta la posterior inferencia acerca de la población de donde se ha tomado. Por ello, su escogencia y selección debe hacerse siguiendo ciertos procedimientos que son ampliamente tratados en la parte de la estadística llamada Teoría de muestreo. El concepto de muestra tiene también las dos connotaciones que hemos señalado para la población. Las características de una población se resumen para su estudio generalmente irá mediante lo que se denominan parámetros; éstos a su vez se toman o consideran como valores verdaderos de la característica estudiada. Por ejemplo, la proporción de todos los clientes que declaran cierta preferencia por una marca particular de un producto dado, es un parámetro de la población de todos los clientes; es la verdadera proporción de la población. Igualmente, la media aritmética de las cuentas corrientes de los clientes de un banco determinado constituye un parámetro de la población de las cuentas de los clientes de ese banco.

Cuando la característica de la población estudiada se reduce a una muestra el resumen de esa característica se hace mediante una está (medida) o estadígrafo. Así por ejemplo. si se toman 100 de todos los posibles clientes y se les entrevista hará ver si están a favor de una marca particular de un producto, estos 100 clientes la constituyen una muestra.. Si hay 70 clientes que prefieren dicha marca entonces la proporción maestral será 0.70 y constituirá un estadígrafo; de igual manera si se escogen 1,000 cuentas del total de las cuentas comentes; las 1,000 observaciones conforman una muestra y el promedio aritmético de estas cuentas un estimador. La inferencia estadística se orienta a sacar conclusiones acerca del parámetro o parámetros poblacionales con base en el valor de un estimador obtenido a partir de los datos muéstrales extraídos de esa población. Para llegar a ese objetivo a través de un proceso racional y eficaz, se aconseja que se tengan en cuenta los siguientes pasos:

1. Formulación del problema. En este punto se debe especificar de manera clara la pregunta que se debe responder y la población de datos asociada a la pregunta. Los conceptos deben ser precisos y deben ponerse limitaciones adecuadas al problema motivadas por el tiempo, dinero disponible y la habilidad de los Investigadores. Algunos conceptos como, artículo defectuoso, económico, salario, pueden variar en cada caso y para cada problema debemos coincidir con las ideas señaladas en el estudio.

2. Diseño del experimento. Este aspecto es de gran importancia, puesto que la recolección de datos requiere dinero y tiempo. Es siempre nuestro deseo obtener máxima Información con el mínimo costo (dinero y tiempo) posible. Incluir excesiva Información en la muestra es a menudo costoso y antieconómico. Incluir poca también es poco satisfactorio. Esto implica, entre otras cosas, que debemos determinar el tamaño de la muestra o la cantidad o tipo de datos que nos permita resolver el problema de la manera más eficiente.

3. Recolección de datos. Esta parte, por lo general, es la que exige más tiempo en la Investigación. Esta recolección debe ajustarse a reglas estrictas ya que de los datos esperamos extraer la Información deseada.

4. Tabulación y descripción de los resultados. En esta etapa, los datos muéstrales se exponen de manera clara y se ilustran con representaciones tabulares y gráficas (diagramas. histogramas, etc.); además se calculan las medidas estadísticas apropiadas al proceso inferencial que haya sido escogido.

5. Inferencia estadística y conclusiones. Este último paso constituye tal vez la contribución más importante de la estadística al proceso inferencial. Aquí se fija el nivel de confiabilidad para la inferencia; esto es debido a que las conclusiones derivadas de inferencias estadísticas jamás se pueden tomar con un 100% de certeza, pero sí se les puede asociar un nivel de confiabilidad; en términos de probabilidad denominados nivel de confianza y nivel de significancia. El proceso Inferencial nos llevará a una conclusión estadística que servirá de orientación a quien o quienes deban tomar la decisión (administrativa o clínica) sobre el tema objeto de estudio.

1.8 Recolección de datos.

La recolección de datos se refiere al uso de una gran diversidad de técnicas y herramientas que pueden ser utilizadas por el analista para desarrollar los sistemas de información, los cuales pueden ser la entrevistas, la encuesta, el cuestionario, la observación, el diagrama de flujo y el diccionario de datos.

Para el caso de la materia de control estadístico de la calidad la recolección de datos se realiza mediante la utilización de hojas de verificación o comprobación, estos son formatos especialmente constituidos para colectar datos fácilmente, en la que todos los artículos o factores necesarios son previamente establecidos y en la que los registros de pruebas, resultados de inspección o resultados de operaciones son fácilmente descritos con marcas utilizadas para verificar.

EJEMPLO:

Imagen

Hoja de verificación		
Fecha: 12-02-2012	Fabrica: Estación de Servicio "Virgen del Valle"	Inspector: Grupo de Trabajo
Tipo de defectos: varios		
Tipo de defectos	**Verificación**	**subtotal**
El acondicionamiento de los surtidores.	IIIII IIIII IIIII IIIII IIIII IIIII IIII	40
Las altas temperaturas producidas por la máquina.	IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII II	50
Fallas en los componentes de los surtidores.	IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII	60
Falta de materia prima.	IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII	120
Los operarios no respetan su hora de descanso.	IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII	78
La Estructura.	IIIII IIIII IIIII IIIII I	25
Tiempo de ocio por parte de los operarios al manejar los surtidores.	IIIII IIII	10
Otros.	IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIII	70
		Total = 453

1.9 Estadística paramétrica (población y muestra aleatoria).

La estadística paramétrica es una rama de la estadística inferencial que comprende los procedimientos estadísticos y de decisión que están basados en distribuciones conocidas. Estas son determinadas usando un número finito de parámetros. Esto es, por ejemplo, si conocemos que la altura de las personas sigue una distribución normal, pero desconocemos cuál es la media y la desviación de dicha normal. La media y la desviación típica de la distribución normal son los dos parámetros que queremos estimar. Cuando desconocemos totalmente qué distribución siguen nuestros datos entonces deberemos aplicar primero un test no paramétrico, que nos ayude a conocer primero la distribución.

La mayoría de procedimientos paramétricos requiere conocer la forma de distribución para las mediciones resultantes de la población estudiada. Para la inferencia paramétrica es requerida como mínimo una escala de intervalo, esto quiere decir que nuestros datos deben tener un orden y una numeración del intervalo. Es decir nuestros datos pueden estar categorizados en: menores de 20 años, de 20 a 40 años, de 40 a 60, de 60 a 80, etc, ya que hay números con los cuales realizar cálculos estadísticos. Sin embargo, datos categorizados en: niños, jóvenes, adultos y ancianos no pueden ser interpretados mediante la estadística paramétrica ya que no se puede hallar un parámetro numérico (como por ejemplo la media de edad) cuando los datos no son numéricos

UNIDAD II

INFERENCIA ESTADÍSTICA: ESTIMACIÓN.

“El conjunto de métodos estadísticos que permiten deducir (inferir) como se distribuye la población en estudio o las relaciones estocásticas entre varias variables de interés a partir de la información que proporciona una muestra”.

Para que un método de inferencia estadística proporcione buenos resultados debe de:

Basarse en una técnica estadístico-matemática adecuada al problema y suficientemente validada.

Utilizar una muestra que realmente sea representativa de la población y de un tamaño suficiente.

Conceptos básicos que se utilizarán en este texto son los siguientes:

Población: es un conjunto homogéneo de individuos sobre los que se estudia una o varias características que son, de alguna forma, observables.

Muestra: es un subconjunto de la población. El número de elementos de la muestra se denomina tamaño muestral.

Muestreo aleatorio simple: es aquel en el que todos los individuos de la población tienen la misma probabilidad de ser elegidos.

Muestra aleatoria simple, de una variable aleatoria *X,* con distribución *F,* de tamaño *n,* es un conjunto de *n* variables aleatorias $X_1, X_2, \ldots, X_n$, independientes e igualmente distribuidas (i.i.d.) con distribución *F*.

Espacio muestral: es el conjunto de muestras posibles que pueden obtenerse al seleccionar una muestra aleatoria, de tamaño *n,* de una cierta población.

Parámetro: es cualquier característica medible de la función de distribución de la variable en estudio (media, varianza,..).

Estadístico: es una función de la muestra $T(X_1, X_2, \ldots, X_n)$. Por tanto, es una variable aleatoria que tiene una función de distribución que se denomina distribución en el muestreo de *T*. Los estadísticos independientes del parámetro a estimar se denominan estimadores.

Propiedades de los estimadores.

Sea $\hat{\theta}_n = \hat{\theta}_n(X_1, X_2, \ldots, X_n)$ un estimador del parámetro θ. Propiedades del estimador son las siguientes

1. Estimador centrado o insesgado, tiene sesgo cero,

$$Sesgo\left(\hat{\theta}_n\right) = E\left(\hat{\theta}_n\right) - \theta.$$

2. Estimador asintóticamente centrado o insesgado, verifica

$$\lim_{n\to\infty} Sesgo\left(\hat{\theta}_n\right) = 0 \Leftrightarrow \lim_{n\to\infty} E\left(\hat{\theta}_n\right) = \theta.$$

3. Error Cuadrático Medio de $\hat{\theta}_n$, es

$$ECM\left(\hat{\theta}_n\right) = E\left(\hat{\theta}_n - \theta\right)^2 = \left(Sesgo\left(\hat{\theta}_n\right)\right)^2 + Var\left(\hat{\theta}_n\right).$$

4. Estimador consistente en media cuadrática, verifica

$$\lim_{n\to\infty} ECM\left(\hat{\theta}_n\right) = 0.$$

por tanto

5. La precisión o eficacia del estimador $\hat{\theta}_n$ es

$$efic\left(\hat{\theta}_n\right) = \frac{1}{ECM\left(\hat{\theta}_n\right)}.$$

Si el estimador es insesgado

$$efic\left(\hat{\theta}_n\right) = \frac{1}{Var\left(\hat{\theta}_n\right)}.$$

6. Estimador de la media poblacional, se utiliza la media muestral definida por

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i. \tag{1.1}$$

7. Si X sigue una distribución $N\left(\mu, \sigma^2\right)$, se verifica que

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Leftrightarrow \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N\left(0,1\right) \tag{1.2}$$

8. Estimador de la varianza poblacional, se utiliza la cuasi varianza maestral definida por

$$\hat{S}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2. \tag{1.3}$$

Si X sigue una distribución $N\left(\mu, \sigma^2\right)$, se verifica que

$$\frac{(n-1)\,\hat{S}^2}{\sigma^2} \sim \chi^2_{n-1}. \tag{1.4}$$

9. Dado que normalmente la varianza poblacional se desconoce y es necesario estimarla, es de interés el siguiente resultado

$$\frac{\bar{X} - \mu}{\hat{S}/\sqrt{n}} \sim t_{n-1} \tag{1.5}$$

2.2.- muestreo

El muestreo probabilístico es un método de muestreo (muestreo se refiere al estudio o el análisis de grupos pequeños de una población) que utiliza formas de métodos de selección aleatoria.

El requisito más importante del muestreo probabilístico es que todos en una población tengan la misma oportunidad de ser seleccionados.

Por ejemplo, si tienes una población de 100 personas, cada persona tendría una probabilidad de 1 de 100 de ser seleccionado. El método de muestreo probabilístico te ofrece la mejor oportunidad de crear una muestra representativa de la población.

Este método utiliza la teoría estadística para seleccionar al azar un pequeño grupo de personas (muestra) de una gran población existente y luego predecir que todas las respuestas juntas coincidirán con la población en general.

Por ejemplo, es prácticamente imposible enviar una encuesta a cada una de las personas de todo un país para recabar información, pero lo que puedes hacer utilizar el método de muestreo de probabilidad para obtener datos que pueden ser muy buenos (incluso aunque se obtengan de una población más pequeña).

Tipos de muestreo probabilístico

El muestreo aleatorio simple, tal y como su nombre lo indica, es un método completamente aleatorio que se utiliza para seleccionar una muestra. Este método de muestreo es tan fácil como asignar números a los individuos (muestra) y luego elegir de manera aleatoria números entre los números a través de un proceso automatizado. Finalmente, los números que se eligen son los miembros que se incluyen la muestra.

Existen dos formas en que las muestras se eligen: A través de un sistema de lotería y uso de software de generación de números aleatorios. Esta técnica de muestreo funciona generalmente en grandes poblaciones y tiene tanto ventajas como desventajas.

Muestreo estratificado: este es un método en el cual una población grande se divide en dos grupos más pequeños, que generalmente no se superponen, sino que representan a toda la población en conjunto.

Durante el muestreo, estos grupos pueden organizarse y luego de estos se puede obtener una muestra de cada grupo por separado.

Algo común en este tipo de método es organizar o clasificar las muestras por sexo, edad, etnia, etc. Este método divide sujetos en grupos mutuamente exclusivos y luego utiliza un muestreo aleatorio simple para elegir miembros de los grupos.

Los miembros de cada uno de estos grupos deben ser distintos para que todos los miembros de todos los grupos tengan la misma oportunidad de ser seleccionados utilizando la probabilidad simple.

Muestreo por conglomerados: este es un método que selecciona de manera aleatoria a los participantes cuando están dispersos geográficamente.

Por ejemplo, tenemos a 1000 participantes de toda la población de México, supongamos que es probable que no sea posible obtener una lista completa de todos estos. Pero en cambio, lo que hace el investigador es seleccionar áreas de manera aleatoria (es decir, ciudades, comunidades, etc), y selecciona al azar dentro de esos límites.

El muestreo por conglomerados por lo general analiza a una población particular en la que la muestra consiste en varios elementos, por ejemplo, ciudad, familia, universidad, etc. Los conglomerados se seleccionan básicamente dividiendo la población mayor en varias secciones más pequeñas.

Muestreo sistemático: este se enfoca en elegir a cada "enésima" persona para que sea parte de la muestra. Por ejemplo, puedes elegir que cada quinta persona sea parte de la muestra, o que cada décima persona sea parte de ella.

El muestreo sistemático es una implementación extendida de la mismísima técnica de probabilidad en la que cual, cada miembro de un grupo es seleccionado en periodos regulares para formar una muestra. Cuando se utiliza este método de muestreo, existe una oportunidad igual para que cada miembro de una población sea seleccionado.

¿Cuáles son los pasos para llevar a cabo un muestreo probabilístico?

1.- Elige cuidadosamente tu población de interés: piensa detenidamente y elige entre la población de manera correcta. Las personas que crees que tienen opiniones que deban recopilarse son las que tienes que incluir en tu muestra.

2.- Determina un marco de muestra adecuado: tu marco debe incluir una muestra de tu población de interés y nadie del exterior. Esto es importante si quieres recopilar datos precisos y que te sirvan.

3.-Selecciona tu muestra y comienza tu encuesta: a veces puede ser difícil encontrar la muestra correcta y determinar el marco de muestra adecuado. Incluso cuando todos los factores están a nuestro favor, muchas veces pueden haber problemas imprevistos como el factor de costo, la calidad de los encuestados y la rapidez de estos en responder.

Obtener una muestra para responder a una verdadera encuesta de probabilidad puede ser difícil, pero no imposible.

En la mayoría de los casos, utilizar la técnica de muestreo probabilístico te ahorrará tiempo, dinero y mucha frustración. Probablemente no puedas enviar encuestas a todas las personas, pero siempre puedes darles a todos la oportunidad de participar, de esto es de lo que se trata la técnica de muestreo de probabilidad.

Toma en cuenta estas consideraciones para tener el mejor muestreo.

¿Cuándo utilizar el muestreo probabilístico?

1.- Cuando se tiene que reducir el sesgo en el muestreo: este método de muestreo se utiliza comúnmente cuando el sesgo debe ser mínimo.

La selección de la muestra determina en gran medida la calidad de la investigación. Y la forma en la que los investigadores seleccionan su muestra determina la calidad de sus hallazgos.

El muestreo probabilístico proporciona en gran medida calidad en los hallazgos del investigador, esto sucede porque se trata de investigar a una representación imparcial de la población. Esto es de especial importancia para eliminar el sesgo en tus encuestas.

2.- Cuando la población es diversa: cuando el tamaño de la población es grande y diversa, este método de muestreo es útil ya que ayuda a los investigadores a crear muestras que representan completamente a la población.

Supongamos que queremos saber cuántas personas prefieren el turismo médico antes de recibir un tratamiento en su propio país, este método de muestreo puede ayudarle al investigador a recoger muestras de diversos estratos socioeconómicos, antecedentes, etc., para representar a la población general.

Conoce más de la importancia de una muestra representativa para una investigación eficaz.

3.- Para crear una muestra precisa: el muestreo probabilístico ayuda a los investigadores a crear una muestra precisa de su población. Los investigadores pueden utilizar este método para crear un tamaño de muestra preciso que les pueda ayudar a obtener datos bien definidos.

Ventajas del muestreo probabilístico

1.- Es rentable: este proceso es rentable y efectivo en relación al tiempo y costo.

2.- Es simple y fácil: el muestreo de probabilidad es un método fácil ya que no implica un proceso complicado. Es rápido y ahorra tiempo.

3.- No es técnico: este método de muestreo no requiere ningún conocimiento técnico debido a la simplicidad con la que puede realizarse. Este método no requiere ningún tipo de conocimiento complejo y por suerte, no es nada largo.

Formula del Muestreo Probabilístico

Existe una gran cantidad de fórmulas para realizar un muestreo probabilístico, una de las más comunes por su sencillez es la del muestreo estatificado, sin embargo te recomendamos leer e investigar los diversos métodos de muestre que hemos mencionado anteriormente.

Ejemplo muestreo estratificado

Expresamos en una tabla todos los datos

	Hombres	Mujeres	Niños	TOTAL
Población	**700**	**800**	**500**	**2000**
Muestra	**x**	**y**	**z**	**80**

Expresamos la proporcionalidad:

$$\frac{700}{x} = \frac{800}{y} = \frac{500}{z} = \frac{2000}{80}$$

Para calcular cualquiera de las incógnitas, buscamos una proporción donde conozcamos 3 de los 4 datos:

$$\frac{700}{x} = \frac{2000}{80} \Rightarrow x = \frac{700 \cdot 80}{2000} = 28$$

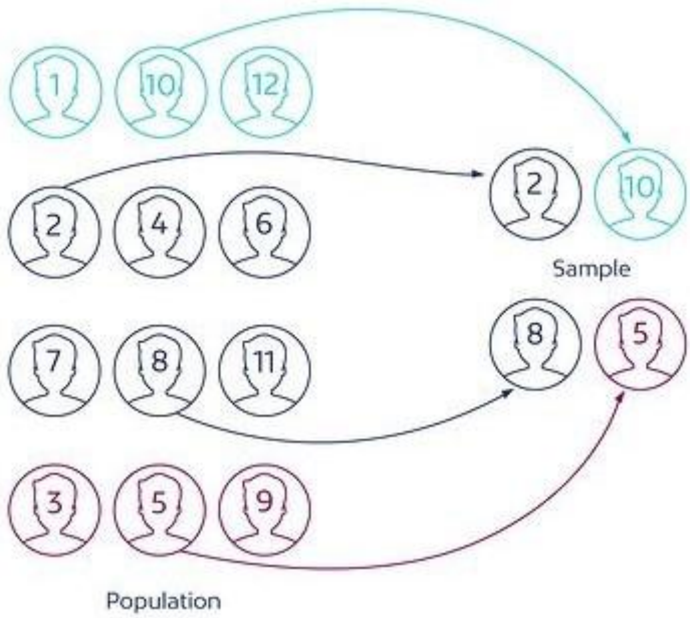
$$\frac{800}{y} = \frac{2000}{80} \Rightarrow y = \frac{800 \cdot 80}{2000} = 32$$

$$\frac{500}{z} = \frac{2000}{80} \Rightarrow x = \frac{500 \cdot 80}{2000} = 20$$

Muestreo probabilístico: muestreo estratificado

Vimos en un post anterior la definición, ventajas e inconvenientes del muestreo aleatorio simple. Veamos ahora el muestreo estratificado.

Esta técnica, perteneciente a la familia de muestreos probabilísticos, consiste en dividir toda la población objeto de estudio en diferentes subgrupos o estratos disjuntos, de manera que un individuo sólo puede pertenecer a un estrato. Una vez definidos los estratos, para crear la muestra se seleccionan individuos empleando una técnica de muestreo cualquiera a cada uno de los estratos por separado. Si por ejemplo empleamos muestreo aleatorio simple en cada estrato, hablaremos de **muestreo aleatorio estratificado** (M.A.E. en adelante). Del mismo modo, podríamos usar otras técnicas de muestreo en cada estrato (muestreo sistemático, aleatorio con reposición, etc.).



Los estratos suelen ser grupos homogéneos de individuos, que a su vez son heterogéneos entre diferentes grupos. Por ejemplo, si en un estudio esperamos encontrar un comportamiento muy diferente entre hombres y mujeres, puede ser conveniente definir dos estratos, uno por cada sexo. Si la selección de estos estratos es correcta:

1. Los hombres deberían comportarse de forma parecida entre ellos.
2. Las mujeres deberían comportarse de forma muy similar entre ellas.
3. Hombres y mujeres deberían mostrar comportamientos dispares entre sí.

Si la anterior condición se cumple (estratos homogéneos internamente, heterogéneos entre sí) el uso del muestreo aleatorio estratificado reduce el error muestral, mejorando la precisión de nuestros resultados al realizar un estudio sobre la muestra.

Es relativamente habitual definir estratos de acuerdo a algunas variables características de la población como son la edad, sexo, clase social o región geográfica. Estas variables permiten dividir fácilmente la muestra en grupos mutuamente excluyentes y con bastante frecuencia, permiten discriminar comportamientos diferentes dentro de la población.

Tipos de muestreo estratificado

Dependiendo del tamaño que asignamos a los estratos, hablaremos de diferentes tipos de muestreo estratificado. También se acostumbra a hablar de diferentes formas de "afijación" de la muestra en estratos.

1. Muestreo estratificado proporcionado

Cuando seleccionamos una característica de los individuos para definir los estratos, suele ocurrir que el tamaño de las subpoblaciones resultantes en el universo son diferentes. Por ejemplo, si queremos estudiar el tanto por ciento de la población que fuma en México y pensamos que la edad puede ser un buen criterio para estratificar (es decir, pensamos que existen diferencias importantes en el hábito de fumar dependiendo de la edad), podemos definir 3 estratos: menores de 20 años, de 20 a 44 años y mayores de 44 años. Es de esperar que al dividir toda la población mexicana en estos 3 estratos no resulten grupos de igual tamaño. Efectivamente, si miramos datos oficiales, obtenemos:

* Estrato 1 - Población Mexicana menor de 19 años: 42,4 millones (41,0%)

* Estrato 2 - Población Mexicana de 20 a 44 años: 37,6 millones (36,3%)

* Estrato 3 - Población Mexicana mayor de 44 años: 23,5 millones (22,7%)

Si usamos **muestreo estratificado proporcionado**, la muestra deberá tener estratos que guarden las mismas proporciones observadas en la población. Si en este ejemplo queremos crear una muestra de 1.000 individuos, los estratos tendrán que tener un tamaño como sigue:

Estrato	Población	*Proporción*	Muestra
1	42,4M	*41,0%*	*410*
2	37,6M	*36,3%*	*363*
3	23,5M	*22,7%*	*227*

2. Muestreo estratificado uniforme

Hablaremos de una **afijación uniforme** cuando asignamos el mismo tamaño de muestra a todos los estratos definidos, sin importar el peso que tienen esos estratos en la población. Siguiendo con el ejemplo anterior, un muestreo estratificado uniforme definiría la siguiente muestra por estrato:

Estrato	Población	*Proporción*	Muestra
1	42,4M	*41,0%*	*334*
2	37,6M	*36,3%*	*333*
3	23,5M	*22,7%*	*333*

Esta técnica favorece los estratos que tienen menos peso en la población, equiparándolos en importancia a los estratos más relevantes. Globalmente, reduce la eficiencia de nuestra muestra (menor precisión en los resultados), pero como contrapartida permite estudiar características particulares de cada estrato con mayor precisión. En nuestro ejemplo, si queremos emitir alguna afirmación específica sobre la población del estrato 3 (mayores de 44 años), podremos hacerlo con menor nivel de error muestral si empleamos una muestra de 333 unidades que si lo hacemos con una muestra de 227 (como ocurría en el muestreo estratificado proporcional).

3. Muestreo estratificado óptimo (respecto a la desviación estándar)

En este caso, el tamaño de los estratos en la muestra no guardará proporcionalidad con la población. Por el contrario, se define el tamaño de los estratos proporcionalmente a la desviación estándar de las variables objeto de estudio. Es decir, se toman estratos de mayor tamaño en los estratos con mayor variabilidad interna para representar mejor en el total de la muestra los grupos poblacionales más difíciles de estudiar.

Eficiencia de los diferentes muestreos estratificados

Las preguntas inevitables son: ¿cuándo conviene emplear la estratificación?, ¿qué tipo de estratificación es más conveniente?

- El **muestreo estratificado proporcional produce siempre menor o igual error muestral que el muestreo aleatorio simple**, es decir, es más preciso. La igualdad se produce cuando las medias o las proporciones que estamos analizando son iguales en todos los estratos. Por lo tanto, la estratificación produce más beneficio cuanto más diferentes sean los estratos entre sí
- El **muestreo estratificado óptimo es siempre igual o más preciso que el muestreo estratificado proporcional**. Ambos métodos son igual de precisos cuando las desviaciones típicas dentro de cada estrato son iguales, en cuyo caso ambos métodos

son totalmente equivalentes. Por lo tanto, la estratificación óptima produce más beneficio cuanto más diferencias existan entre las desviaciones dentro de cada grupo, situación en la que podremos reducir el tamaño muestral de los grupos más homogéneos en beneficio de los más heterogéneos. Como contrapartida, es un método más complejo y que requiere tener mucha información a priori de la muestra que estudiamos, algo que normalmente no tenemos.

Tamaños de muestra requeridos por cada técnica

Vemos que la estratificación puede proporcionar beneficios. Si estas técnicas pueden emplearse para estimar de forma más precisa ya sean medias (p.e. media de cigarrillos consumidos por los fumadores de México) o proporciones (p.e. proporción de la población de México que fuma), también pueden permitirnos reducir el tamaño de muestra requerido para lograr una estimación con un nivel de error determinado.

La siguiente tabla resume el tamaño de muestra requerido al emplear cada técnica, en función del error máximo que estamos dispuestos a aceptar y de las características del propio universo, que consideraremos de tamaño infinito (si fuese finito, debe aplicarse un factor de corrección).

	Tamaño de muestra para estimar una proporción	Tamaño de muestra para estimar una media
Muestreo aleatorio simple	$\frac{Z^2 p(1-p)}{e^2}$	$\frac{Z^2 \sigma^2}{e^2}$
Muestreo estratificado proporcional	$\frac{Z^2 \sum_{h=1}^{L} W_h p_h (1-p_h)}{e^2}$	$\frac{Z^2 \sum_{h=1}^{L} W_h \sigma_h^2}{e^2}$
Muestreo estratificado óptimo	$\frac{Z^2 (\sum_{h=1}^{L} W_h \sqrt{p_h (1-p_h)})^2}{e^2}$	$\frac{Z^2 (\sum_{h=1}^{L} W_h \sigma_h)^2}{e^2}$

Para interpretar el cuadro anterior es necesario tener en cuenta lo siguiente:

- **Z** = Es la desviación del valor medio que aceptamos para lograr el nivel de confianza deseado. En función del nivel de confianza que busquemos, usaremos un valor determinado que viene dado por la forma que tiene la distribución de Gauss. Los valores más frecuentes son:

Nivel de confianza 90% -> Z=1,645

Nivel de confianza 95% -> Z=1,96

Nivel de confianza 99% -> Z=2,575

- **L** es el número de estratos en que particionamos la muestra y **h** es un índice que se refiere a un estrato concreto. Por lo tanto, h puede variar entre 1 y L estratos.
- **p** es la proporción que buscamos en el total de la población (p.e. % de fumadores). Por lo tanto, **(1-p)** es la proporción de la muestra complementaria, la que no cumple el criterio buscado (no fumadores). Del mismo modo, $\mathbf{p_h}$ es dicha proporción dentro de cada uno de los estratos.
- $\boldsymbol{\sigma}^2$ es la varianza que el dato buscado (en el caso de estimar medias) tiene en el total de la población. Asimismo, $\boldsymbol{\sigma_h}^2$ es la varianza dentro de cada estrato.
- **e** es el margen de error aceptado.
- **Wh** es el peso que el estrato tiene en la muestra (tamaño del estrato respecto al total de la muestra). Si hablamos de estratificación proporcional, cada Wh es igual a la proporción que ese estrato representa en la población. Si hablamos de estratificación óptima, cada Wh se calcula en función de la dispersión dentro de cada estrato.

Es posible demostrar a partir de las fórmulas anteriores que los diferentes métodos de estratificación sólo reducen el tamaño de la muestra si los valores de **p** y **σ** varían entre estratos. De lo contrario, todas las expresiones son equivalentes. Veamos un ejemplo: si tomamos la expresión de tamaño de muestra requerido para estimar una media mediante un muestreo estratificado óptimo (ignorando el parámetro Z en este caso)

$$\frac{\left(\sum_{h=1}^{L} W_h \sigma_h\right)^2}{e^2}$$

y consideramos que todas las varianzas de los estratos son iguales ($\sigma_h=\sigma$) y que el tamaño de los estratos es idéntico ($W_h=1/L$), el resultado que obtenemos es

$$\frac{\left(\sum_{h=1}^{L} W_h \sigma_h\right)^2}{e^2} = \frac{\left(\sum_{h=1}^{L} \frac{1}{L}\sigma\right)^2}{e^2} = \frac{\left(\frac{L}{L}\sigma\right)^2}{e^2} = \frac{\sigma^2}{e^2}$$

1. Muestreo aleatorio simple

Las encuestas por muestreo consisten en extraer de una población finita de N unidades, subpoblaciones de un tamaño fijado de antemano. Si todas las unidades son indistinguibles, el número de muestras de tamaño n viene dado por:

$$\binom{N}{n} = \frac{N!}{n!(N-n)!} = {}_{N}C_{n}$$

Por ejemplo, si la población contiene 5 unidades A, B, C, D, E; existen 10 muestras diferentes de tamaño 3, que son:

ABC, ABD, ABE, ACD, ACE

ADE, BCD, BCE, BDE. CDE

Debe notarse que la misma letra no ocurre dos veces en la misma muestra; y, también, que el orden de los elementos no tiene importancia, las seis muestras ABC, ACB, BAC, BCA, CAB, CBA son consideradas como iguales.

El muestreo aleatorio simple es un método de selección de n unidades sacadas de N, de tal manera que cada una de las muestras tiene la misma probabilidad de ser elegida.

En la práctica una muestra aleatoria simple es extraída de la siguiente forma:

Se numeran las unidades de la población del 1 al N, y por medio de una tabla de números aleatorios o colocando los números 1 a N en una urna, se extraen sucesivamente n números. Las unidades que llevan estos números constituyen la muestra.

El método elegido debe de verificar que en cualquier fase de la obtención de la muestra cada individuo que no ha sido sacado previamente, tiene la misma probabilidad de ser elegido[1].

Es fácil ver que cada una de las ${}_NC_n$ muestras tiene igual posibilidad de obtenerse.

Cuando un número ha sido sacado de la urna, éste no es reemplazado, ya que esto daría lugar a que la misma unidad entrara en la muestra más de una vez. Por esta razón el muestreo es descrito como sin reemplazo. El muestreo con reemplazo, es totalmente factible, aunque rara vez es usado, ya que no se ve la conveniencia de tener el mismo individuo dos veces en la misma muestra.

2. Muestreo estratificado aleatorio.

En este tipo de muestreo, la población de N unidades es dividida en subpoblaciones de N_1, N_2, …, N_L unidades, respectivamente. Estas subpoblaciones no se sobreponen y juntas forman la totalidad de la población, por lo que

$N_1 + N_2 + \ldots + N_L = N$

Las subpoblaciones son llamados estratos. Para obtener un beneficio completo de la estratificación se debe de conocer N_h. Una vez que han sido determinados los estratos, se saca una muestra de cada uno, la extracción se realiza de forma independiente en cada estrato. Los tamaños de la muestra dentro de los estratos son representados por n_1, n_2, …, n_L.

Si se toma una muestra aleatoria simple en cada estrato, el procedimiento completo es conocido como muestreo estratificado aleatorio.

La estratificación es una técnica común. Hay muchas razones para realizarla; las principales son:

Si se desea cierta precisión en alguna subdivisión, es necesario tratarla como si fuera una “población” por sí sola.

Las conveniencias de tipo administrativo.

La diversidad de determinados grupos (por ejemplo, hoteles, hospitales, prisiones, etc.) hace necesario un enfoque diferente al de las personas normales. O, por ejemplo, las grandes compañías conviene separarlas en un estrato diferente, para las pequeñas empresas se puede utilizar un tipo de muestreo por áreas.

La estratificación puede dar lugar a una ganancia en precisión de los estimadores de la población. Esto ocurre cuando una población heterogénea es dividida en subpoblaciones cada una de las cuales es internamente homogénea.

Muestreo sistemático

Este método de muestreo consiste en lo siguiente: Supóngase que las N unidades de la población se numeran en algún orden de 1 a N. Para seleccionar una muestra de n unidades tomamos una al azar de las k primeras unidades, a continuación elegimos la que viene k unidades siguientes y así sucesivamente. Por ejemplo, si k =30 y la primera unidad elegida es la 19, las subsiguientes unidades serán los números 49, 79, 109, etc. La selección de la primera unidad determina la muestra completa. Este tipo de muestreo se llama muestra sistemática de cada k-ésima unidad.

Las ventajas de este método sobre el aleatorio simple son:

Es más fácil obtener la muestra y ejecutarlo con menos errores.

Intuitivamente aparece como más preciso que el muestreo simple aleatorio. En efecto, estratifica la población en n sustratos, los cuales consisten en las primeras k unidades, las segundas k unidades, etc. Eligiendo una unidad por estrato. La diferencia está en que en el muestreo estratificado la unidad dentro de cada sustrato se elige al azar, en este siempre está en la misma posición relativa.

Una variante del muestreo sistemático consiste en escoger cada unidad en el centro del estrato; esto es, en lugar de empezar la secuencia con un número al azar escogido del *1* al *k*, tomamos el número inicial como $(k+1)/2$ si k es impar y $(k+2)/2$ si k es par.

4. Muestreo por conglomerado.

La población está dividida en áreas lo más heterogenias posibles internamente y lo más homogéneas posibles entre sí. Selecciona al azar un conglomerado que será el que formará la muestra.

Hay dos razones principales para la extensa aplicación del muestreo por conglomerado. En muchos países no hay listas completas ni al día de las personas, fincas, casas, etc en una región geográfica grande. Sin embargo, a partir de mapas de la región, la misma puede ser subdividida en segmentos de tierra con límites fácilmente identificables en las zonas rurales, o en unidades de superficie como manzanas en zonas urbanas. En EE.UU y Europa se toman a menudo estos conglomerados, porque resuelven el problema de construir una lista de unidades de muestreo.

Aun cuando se dispongan de listas consideraciones económicas pueden apuntar hacia la elección de una unidad conglomerada mayor. Para un tamaño de muestra dado una unidad pequeña usualmente da resultados más precisos que una unidad grande. Por ejemplo, una simple muestra al azar de 600 casas cubre una ciudad más uniformemente que 20 manzanas de 30 casas cada una. Pero obviamente se incurren en más gasto seleccionando 600 casas al azar y viajando por ellas que localizando 20 manzanas y la visita de todas las casas de las mismas. Cuando el costo es contrapesado con la precisión, la unidad mayor puede ser superior. En muchas decisiones prácticas el tipo de unidad puede tener alguna conveniencia o desventaja especial. Por ejemplo, elegir unidades pequeñas al muestrear una cosecha puede introducir un sesgo debido a la incertidumbre de los límites exactos de la unidad.

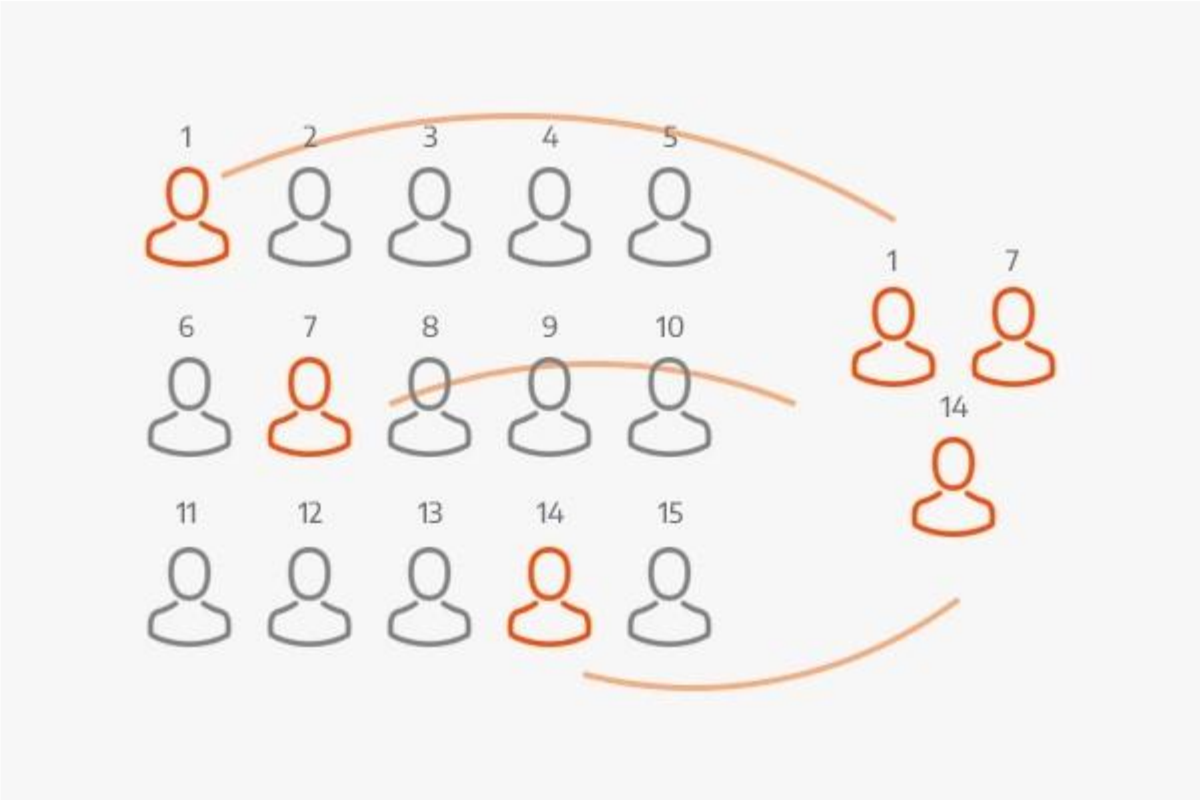
¿Qué es el muestreo aleatorio simple?

El muestreo aleatorio simple es un tipo de muestreo probabilístico que se aplica al tomar una muestra en la que todos los elementos del universo tienen la misma probabilidad de ser seleccionadas.

Esta funciona a través de un método de «sorteo» en el cual, entre un universo de individuos, se le asigna a cada integrante un número con el que puede ser escogido.

Por ejemplo, si se necesita seleccionar una muestra de 50 personas en un universo de 1000, se le asignará a esas 1000 un número y a modo de sorteo se seleccionarán 50 números al azar conformando de esta manera la muestra requerida.

Dicho procedimiento garantiza que todos los elementos muestrales tengan la misma importancia y probabilidad de ser seleccionados y llegar a formar parte de la muestra.



Selección al azar del muestreo aleatorio simple

Ventajas y desventajas del muestreo aleatorio simple

A la hora de utilizar el muestreo aleatorio simple, este presenta diversas ventajas y desventajas.

Ventajas

Al utilizar este procedimiento obtendrás las siguientes ventajas:

- La principal ventaja de seleccionar el muestreo aleatorio simple es lo sencillo que es para armar las muestras.
- Es un método donde se toma de forma equitativa la selección de las muestras a partir de una población.
- Todos los individuos de la población en general tienen igualdad de oportunidades de ser seleccionado.
- La población es representativa, siendo el único margen de error la suerte y pasa a llamarse error de muestreo.
- Es el mejor método a la hora de explicar los resultados ya que su selección es aleatoria e imparcial.
- Por su representatividad obtenida, se pueden realizar generalizaciones a partir de los resultados de las muestras con respecto a la población.

Desventajas

El muestreo aleatorio simple es uno de los más eficaces y prácticos, es por eso que sus desventajas son muy pocas.

- Se requiere de una lista completa de todos los miembros de la población.
- Esta lista debe estar debidamente elaborada, completa y actualizada.
- En las poblaciones grandes es mucho más difícil disponer de todos estos datos, por lo que se recomienda usar otra técnica.

Ejemplos del muestreo aleatorio simple

Una consultora que realiza auditorías sobre diversas organizaciones decide realizar una investigación en una empresa que posee 500 empleados.

Para realizar dicho estudio, necesita realizar una muestra estadística de 50 personas (10%), el método que selecciona el investigador para realizarla es el de muestreo aleatorio simple, por lo tanto, los pasos a seguir por el investigador son los siguientes:

1. Enlistar con número a los 500 empleados de la orga nización.

2. Elije 50 números al azar.

3. Los 50 números que correspondan a un empleado serán los individuos que conformen la muestra.

De esta forma queda conformada la muestra para poder realizar la investigación, de tal manera que todos los empleados de la org anización tuvieron las mismas posibilidades de ser seleccionados para conformar dicha muestra.

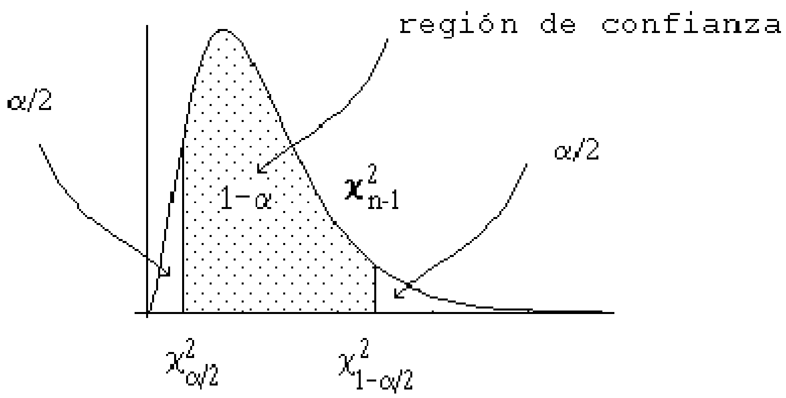
Intervalo de confianza para la varianza

Para estimar un intervalo de confianza para la varianza, nos ayudaremos de la siguiente propiedad de la distribución χ^2:

$$\chi^2_{n-1} = \sum_{i=1}^{n} \frac{(X_i - \overline{X})^2}{\sigma^2} = \frac{(n-1)\,\hat{S}^2}{\sigma^2} \rightsquigarrow \chi^2_{n-1}$$

Consideremos dos cuantiles de esta distribución que nos dejen una probabilidad $1-\alpha$ en la ``zona central" de la distribución (cf. figura 8.7):

Figura: Cuantiles de la distribución χ^2_{n-1}.



$$\begin{cases} \mathcal{P}\left[\chi^2_{n-1} < \chi^2_{n-1,\alpha/2}\right] = \frac{\alpha}{2} \\ \\ \mathcal{P}\left[\chi^2_{n-1} > \chi^2_{n-1,1-\alpha/2}\right] = \frac{\alpha}{2} \end{cases} \Longrightarrow \mathcal{P}\left[\chi^2_{n-1,\alpha/2} \leq \chi^2_{n-1} \leq \chi^2_{n-1,1-\alpha/2}\right] = 1-\alpha$$

Entonces un intervalo de confianza al nivel $1-\alpha$ para la varianza de una distribución gaussiana (cuyos parámetros desconocemos) lo obtenemos teniendo en cuenta que existe una probabilidad $1-\alpha$ de que:

$$\chi^2_{n-1,\alpha/2} \leq \chi^2_{n-1} \leq \chi^2_{n-1,1-\alpha/2} \quad \Rightarrow$$

$$\Rightarrow \quad \chi^2_{n-1,\alpha/2} \leq \frac{(n-1)\,\hat{\mathcal{S}}^2}{\sigma^2} \leq \chi^2_{n-1,1-\alpha/2}$$

$$\Rightarrow \quad \frac{(n-1)\hat{\mathcal{S}}^2}{\chi^2_{n-1,1-\alpha/2}} \leq \sigma^2 \leq \frac{(n-1)\hat{\mathcal{S}}^2}{\chi^2_{n-1,\alpha/2}}$$

Por tanto el intervalo que buscamos es

$$\sigma^2 \in \left[\frac{(n-1)\hat{\mathcal{S}}^2}{\chi^2_{n-1,1-\alpha/2}} \quad , \quad \frac{(n-1)\hat{\mathcal{S}}^2}{\chi^2_{n-1,\alpha/2}}\right]$$

Ejemplo

En un ejemplo anterior se estudiaba la altura de los individuos de una ciudad, obteniéndose en una muestra de tamaño 25 los siguientes valores:

$$\begin{aligned} \overline{x} &= 170 \text{ cm} \\ \mathcal{S} &= 10 \text{ cm} \end{aligned}$$

Calcular un intervalo de confianza con $\alpha = 0{,}05$ para la varianza σ^2 de la altura de los individuos de la ciudad.

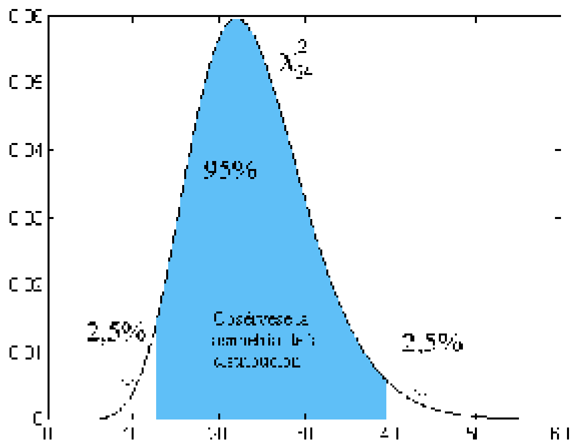
Solución:

Para estimar un intervalo de confianza para σ^2 (varianza poblacional) el estadístico que nos resulta �til es:

$$\chi^2 = \frac{(n-1)\,\hat{\mathcal{S}}^2}{\sigma^2} \rightsquigarrow \chi^2_{n-1}$$

Entonces el intervalo de confianza que buscamos lo obtenemos mediante (cf. figura 8.8)

Figura: Percentiles del 2,5% y del 97,5% para la distribución χ^2_{24}.



$$\chi^2_{n-1,\alpha/2} \leq \chi^2 \leq \chi_{n-1,1-\alpha^2} \iff \chi^2_{24;0,025} = 12,4 \leq \frac{24 \cdot 10,206^2}{\sigma^2} \leq \chi_{24;0,075} = 39'4$$

$$\iff \sigma^2 \in [63,45;201,60]$$

Por tanto, para el valor poblacional de la desviación típica tenemos que

$$7,96 \leq \sigma \leq 14,199$$

Con una confianza del 95%, que por supuesto contiene a las estimaciones puntuales $S = 10$ y $\hat{S} = 10,206$ calculados sobre la muestra.

UNIDAD III

PRUEBAS DE HIPÓTESIS CON UNA MUESTRA

La hipótesis de investigación que se genera en todo proyecto se define como la proposición o explicación tentativa del fenómeno investigado o la postulación de lo que se busca o se trata de probar. La hipótesis de trabajo está integrada por enunciados formales que declaran lo que el investigador quiere probar. De tal manera que, para refutar o confirmar un problema de investigación, se debe plantear una hipótesis que intenta proponer o explicar la relación entre dos variables, y ésta se debe apoyar siempre en conocimientos organizados y sistematizados.

El diseño de una investigación clínica debe tomar en cuenta si el estudio pretende generar una hipótesis para ser probada en estudios futuros o probar hipótesis específicas sobre las que el investigador tiene cierta evidencia de que sus observaciones puedan ser ciertas. Los estudios que generan hipótesis se conocen como exploratorios; los que prueban diferentes hipótesis se conocen como confirmatorios. Por supuesto que un solo estudio puede tener aspectos exploratorios y confirmatorios.

Justificación de la prueba de hipótesis

La prueba de hipótesis es un método esencial para la toma de decisiones. La decisión relaciona la elección entre dos enunciados competitivos y mutuamente excluyentes, respecto de uno o más parámetros de la población. Los enunciados competitivos se conocen como hipótesis nula y alternativa, respectivamente. Con base a lo anterior, es necesario señalar los atributos principales que debe poseer una hipótesis: Debe hacer referencia a una situación real. Las variables que se presentan en su planteamiento deben ser precisas, comprensibles y concretas.

La relación entre las variables debe ser clara, verosímil y lógica.

Los términos y las relaciones planteadas deben ser observables y medibles.

Las variables deben estar relacionadas con técnicas disponibles para probarlas.

Tipos y clases de hipótesis

En esencia, existen dos tipos diferentes de hipótesis:

Hipótesis nula (H0). Es definida como una manifestación que reclama la ausencia de la diferencia entre valores o variables supuestas o hipotéticas y la media de la población. Esta hipótesis refuta, niega o plantea lo contrario de la hipótesis de investigación y suele plantear que no existen diferencias.

Hipótesis de investigación (H1). Se le conoce también como hipótesis de trabajo, alternativa (HA) o estadística, y es una manifestación en desacuerdo de la hipótesis nula.

Estos dos tipos de hipótesis pueden ser de cuatro clases diferentes:

Descriptiva. Este tipo de hipótesis sólo intenta describir el valor de las variables que se van a observar en el contexto o en la manifestación de otra variable.

Correlaciónales. Son hipótesis que especifican la correlación entre dos variables. Las hipótesis de este tipo pueden establecer asociación, predicción o ser explicativas, pero nunca causales. Diferenciales. Son hipótesis que intentan definir diferencias entre grupos.

3.2 Hipótesis nula y alternativa.

Las hipótesis nula y alternativa son dos enunciados mutuamente excluyentes acerca de una población. Una prueba de hipótesis utiliza los datos de la muestra para determinar si se puede rechazar la hipótesis nula.

Hipótesis nula (H0)

La hipótesis nula indica que un parámetro de población (tal como la media, la desviación estándar, etc.) es igual a un valor hipotético. La hipótesis nula suele ser una afirmación inicial que se basa en análisis previos o en conocimiento especializado.

Hipótesis alternativa (H1)

La hipótesis alternativa indica que un parámetro de población es más pequeño, más grande o diferente del valor hipotético de la hipótesis nula. La hipótesis alternativa es lo que usted podría pensar que es cierto o espera probar que es cierto.

Hipótesis unilaterales y bilaterales

La hipótesis alternativa puede ser unilateral o bilateral.

Bilateral

Utilice una hipótesis alternativa bilateral (también conocida como hipótesis no direccional) para determinar si el parámetro de población es mayor que o menor que el valor hipotético. Una prueba bilateral puede detectar cuándo el parámetro de población difiere en cualquier dirección, pero tiene menos potencia que una prueba unilateral.

Unilateral

Utilice una hipótesis alternativa unilateral (también conocida como hipótesis direccional) para determinar si el parámetro de población difiere del valor hipotético en una dirección específica. Usted puede especificar la dirección para que sea mayor que o menor que el valor hipotético. Una prueba unilateral tiene mayor potencia que una prueba bilateral, pero no puede detectar si el parámetro de población difiere en la dirección opuesta.

Ejemplos de hipótesis bilaterales y unilaterales

Bilateral

Un investigador tiene los resultados de una muestra de estudiantes que presentaron un examen nacional en una escuela secundaria. El investigador desea saber si las calificaciones de esa escuela difieren del promedio nacional de 850. Una hipótesis alternativa bilateral (también conocida como hipótesis no direccional) es adecuada porque el investigador está interesado en determinar si las calificaciones son menores que o mayores que el promedio nacional. ($H0$: $\mu = 850$ vs. $H1$: $\mu \neq 850$)

Unilateral

Un investigador tiene los resultados de una muestra de estudiantes que tomaron un curso de preparación para un examen nacional. El investigador desea saber si los estudiantes preparados tuvieron puntuaciones por encima del promedio nacional de 850. Una hipótesis alternativa unilateral (también conocida como hipótesis direccional) se puede utilizar porque el investigador plantea la hipótesis de que las puntuaciones de los estudiantes preparados son mayores que el promedio nacional. ($H0$: $\mu = 850$ vs. $H1$: $\mu > 850$)

3.3 Error tipo I y error 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_0cuando 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 se pone en venta el medicamento menos efectivo en lugar del medicamento más efectivo.

Cuando realice las pruebas de hipótesis, considere los riesgos de cometer errores de tipo I y tipo II. Si las consecuencias de cometer un tipo de error son más graves o costosas que cometer el otro tipo de error, entonces elija un nivel de significancia y una potencia para la prueba que reflejen la gravedad relativa de esas consecuencias.

PRUEBAS DE HIPÓTESIS

Si queremos decidir entre dos hipótesis que afectan a un cierto parámetro de la población, a partir de la información de la muestra usaremos el contraste de hipótesis, cuando optemos por una de estas dos hipótesis, hemos de conocer una medida del error cometido, es decir, cuantas veces de cada cien nos equivocamos.

En primer lugar, veremos cómo se escribirían las hipótesis que queremos contrastar:

H0 se llama hipótesis nula y es lo contrario de lo que sospechamos que va a ocurrir (suele llevar los signos igual, mayor o igual y menor o igual)

H1 se llama hipótesis alternativa y es lo que sospechamos que va a ser cierto (suele llevar los signos distinto, mayor y menor)

Los contrastes de hipótesis pueden ser de dos tipos:

Bilateral: En la hipótesis alternativa aparece el signo distinto.

Unilateral: En la hipótesis alternativa aparece o el signo > o el signo <.

Podemos aceptar una hipótesis cuando en realidad no es cierta, entonces cometeremos unos errores, que podrán ser de dos tipos:

Error de tipo I: Consiste en aceptar la hipótesis alternativa cuando la cierta es la nula.

Error de tipo II: Consiste en aceptar la hipótesis nula cuando la cierta es la alternativa.

Estos errores los aceptaremos si no son muy grandes o si no nos importa que sean muy grandes.

Alfa: Es la probabilidad de cometer un error de tipo I.

Beta: Es la probabilidad de cometer un error de tipo II.

De los dos, el más importante es alfa que llamaremos nivel de significación y nos informa de la probabilidad que tenemos de estar equivocados si aceptamos la hipótesis alternativa. Debido a que los dos errores anteriores a la vez son imposibles de controlar, vamos a fijarnos solamente en el nivel de significación, este es el que nos interesa ya que la hipótesis alternativa que estamos interesados en probar y no queremos aceptarla si en realidad no es cierta, es decir, si aceptamos la hipótesis alternativa queremos equivocarnos con un margen de error muy pequeño.

El nivel de significación lo marcamos nosotros. Si es grande es más fácil aceptar la hipótesis alternativa cuando en realidad es falsa. El valor del nivel de significación suele ser un 5%, lo que significa que 5 de cada 100 veces aceptamos la hipótesis alternativa cuando la cierta es la nula.

Solamente vamos a estudiar el contraste bilateral para la media.

CONTRASTE DE HIPÓTESIS BILATERAL PARA LA MEDIA

Si se cumple una de las siguientes hipótesis:

El tamaño de la muestra es mayor de 30 y la variable sigue un modelo normal.

El tamaño de la muestra es mayor de 100.

Estudiaremos el siguiente contrate de hipótesis bilateral:

Calculamos los siguientes valores:

Valor experimental que se calcula a partir de la muestra.

Valor teórico y es el valor que en la distribución N (0,1) deja a su derecha un área de alfa/2 para un nivel de significación alfa. Es el valor z que definíamos ala principio del tema.

La regla de decisión fijado el nivel de significación, alfa, es la siguiente:

Si se acepta la hipótesis alternativa, llegamos a la conclusión de que la hipótesis es cierta. Si se acepta la hipótesis nula, en realidad no podemos afirmar que sea cierta, sino que la hipótesis alternativa no es cierta, ya que el margen de error con el que se acepta la hipótesis nula es muy grande.

Actividad 21. Un equipo de psicólogos han comprobado que en cierta población infantil, el tiempo (en minutos) empleado en realizar determinada actividad manual, sigue un modelo

Normal de probabilidad. Un grupo de 36 niños, seleccionados aleatoriamente en dicha población, realizaron esa actividad manual en un tiempo medio de 6,5 minutos con una desviación típica muestral de 1,5 minutos. A partir de esta información, para un nivel de significación del 1% (alfa=0,01) ¿podíamos rechazar la hipótesis de que el tiempo medio en la población es de 7 minutos? Utiliza la escena siguiente.

No podemos aceptar la hipótesis alternativa, y por lo tanto no podemos rechazar la hipótesis de que el tiempo medio en la población es de 7 minutos.

Actividad 22. El gerente de una empresa selecciona aleatoriamente entre sus trabajadores una muestra de 169 y anota el número de horas de trabajo que cada uno de ellos ha perdido por causa de accidentes laborales en el año 2001. A partir de la información obtenida determina, en esos 169 trabajadores, un número medio de horas perdidas por accidentes laborales en el 2001 de 36,5 horas. Sabiendo que:

Donde representa el número de horas perdidas por el i-ésimo trabajador.

a) ¿Podríamos rechazar, con un nivel de significación del 1% la hipótesis de que el número medio de horas perdidas a causa de accidentes laborales en esa empresa durante el año 2001 fue de 35 horas?

b) ¿Y para un nivel de significación del 5%?

3.4 Pruebas de hipótesis Z para la media (desviación estándar poblacional conocida).

Dentro del estudio de la inferencia estadística, se describe como se puede tomar una muestra aleatoria y a partir de esta muestra estimar el valor de un parámetro poblacional en la cual se puede emplear el método de muestreo y el teorema del valor central lo que permite explicar cómo a partir de una muestra se puede inferir algo acerca de una población, lo cual nos lleva a definir y elaborar una distribución de muestreo de medias muestrales que nos permite explicar el teorema del límite central y utilizar este teorema para encontrar las probabilidades de obtener las distintas medias maestrales de una población.

Pero es necesario tener conocimiento de ciertos datos de la población como la media, la desviación estándar o la forma de la población, pero a veces no se dispone de esta información.

En este caso es necesario hacer una estimación puntual que es un valor que se usa para estimar un valor poblacional. Pero una estimación puntual es un solo valor y se requiere un intervalo de valores a esto se denomina intervalo de confianza y se espera que dentro de este intervalo se encuentre el parámetro poblacional buscado. También se utiliza una estimación mediante un intervalo, el cual es un rango de valores en el que se espera se encuentre el parámetro poblacional

En nuestro caso se desarrolla un procedimiento para probar la validez de una aseveración acerca de un parámetro poblacional este método es denominado Prueba de hipótesis para una muestra.

2.- HIPOTESIS Y PRUEBA DE HIPOTESIS

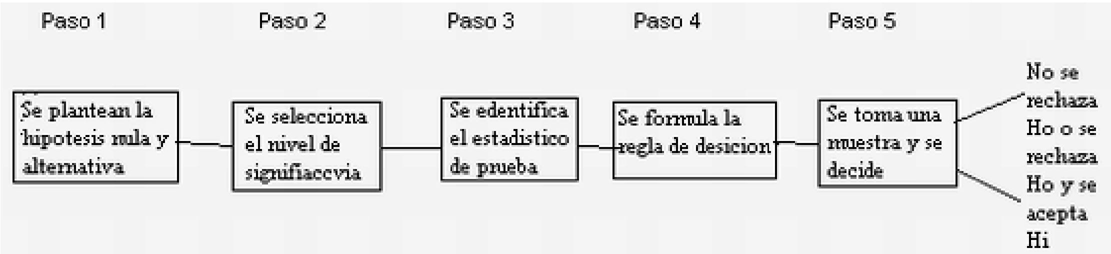
Tenemos que empezar por definir que es una hipótesis y que es prueba de hipótesis.

Hipótesis es una aseveración de una población elaborado con el propósito de poner a prueba, para verificar si la afirmación es razonable se usan datos.

En el análisis estadístico se hace una aseveración, es decir, se plantea una hipótesis, después se hacen las pruebas para verificar la aseveración o para determinar que no es verdadera.

Por tanto, la prueba de hipótesis es un procedimiento basado en la evidencia maestral y la teoría de probabilidad; se emplea para determinar si la hipótesis es una afirmación razonable.

Prueba de una hipótesis: se realiza mediante un procedimiento sistemático de cinco paso:



Siguiendo este procedimiento sistemático, al llegar al paso cinco se puede o no rechazar la hipótesis, pero debemos de tener cuidado con esta determinación ya que en la consideración de estadística no proporciona evidencia de que algo sea verdadero. Esta prueba aporta una clase de prueba más allá de una duda razonable. Analizaremos cada paso en detalle

Objetivo de la prueba de hipótesis.

El propósito de la prueba de hipótesis no es cuestionar el valor calculado del estadístico (muestral), sino hacer un juicio con respecto a la **diferencia** entre estadístico de muestra y un valor planteado del parámetro.

3.- Procedimiento sistemático para una prueba de hipótesis de una muestra

.**Paso 1: Plantear la hipótesis nula Ho y la hipótesis alternativa H1.**

Cualquier investigación estadística implica la existencia de hipótesis o afirmaciones acerca de las poblaciones que se estudian.

La hipótesis nula (Ho) se refiere siempre a un valor especificado del parámetro de población, no a una estadística de muestra. La letra H significa hipótesis y el subíndice cero no hay diferencia. Por lo general hay un "no" en la hipótesis nula que indica que "no hay cambio" Podemos rechazar o aceptar Ho.

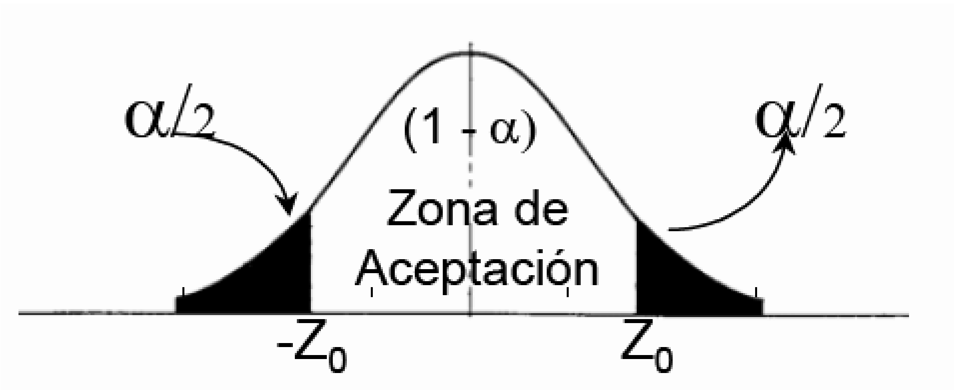
La hipótesis nula es una afirmación que no se rechaza a menos que los datos maestrales proporcionen evidencia convincente de que es falsa. El planteamiento de la hipótesis nula siempre contiene un signo de igualdad con respecto al valor especificado del parámetro.

La hipótesis alternativa (H1) es cualquier hipótesis que difiera de la hipótesis nula. Es una afirmación que se acepta si los datos maestrales proporcionan evidencia suficiente de que la hipótesis nula es falsa. Se le conoce también como la hipótesis de investigación. El planteamiento de la hipótesis alternativa nunca contiene un signo de igualdad con respecto al valor especificado del parámetro.

Paso 2: Seleccionar el nivel de significancia.

Nivel de significancia: Probabilidad de rechazar la hipótesis nula cuando es verdadera. Se le denota mediante la letra griega α, también es denominada como nivel de riesgo, este término es más adecuado ya que se corre el riesgo de rechazar la hipótesis nula, cuando en realidad es verdadera. Este nivel esta bajo el control de la persona que realiza la prueba.

Si suponemos que la hipótesis planteada es verdadera, entonces, el nivel de significación indicará la probabilidad de no aceptarla, es decir, estén **fuera** de área de aceptación. El **nivel de confianza** (1-α), indica la probabilidad de aceptar la hipótesis planteada, cuando es verdadera en la población.



La distribución de muestreo de la estadística de prueba se divide en dos regiones, una región de rechazo (conocida como región crítica) y una región de no rechazo (aceptación). Si la estadística de prueba cae dentro de la región de aceptación, no se puede rechazar la hipótesis nula.

La región de rechazo puede considerarse como el conjunto de valores de la estadística de prueba que no tienen posibilidad de presentarse si la hipótesis nula es verdadera. Por otro lado, estos valores no son tan improbables de presentarse si la hipótesis nula es falsa. El valor crítico separa la región de no rechazo de la de rechazo.

Tipos de errores

Cualquiera sea la decisión tomada a partir de una prueba de hipótesis, ya sea de aceptación del Ho o de la Ha, puede incurrirse en error:

Un error tipo I se presenta si la hipótesis nula Ho es rechazada cuando es verdadera y debía ser aceptada. La probabilidad de cometer un error tipo I se denomina con la letra alfa α

Un error tipo II, se denota con la letra griega β se presenta si la hipótesis nula es aceptada cuando de hecho es falsa y debía ser rechazada.

En cualquiera de los dos casos se comete un error al tomar una decisión equivocada.

En la siguiente tabla se muestran las decisiones que pueden tomar el investigador y las consecuencias posibles.

	Investigador	
Hipotesis nula	Se acepta Ho	Se rechaza Ho
Ho es verdadrea	Decision correcta	Error tipo I
Ho es falsa	Error tipo II	Decision orrecta

Para que cualquier ensayo de hipótesis sea bueno, debe diseñarse de forma que minimice los errores de decisión. En la práctica un tipo de error puede tener más importancia que el otro, y así se tiene a conseguir poner una limitación al error de mayor importancia. La única forma de reducir ambos tipos de errores es incrementar el tamaño de la muestra, lo cual puede ser o no ser posible.

La probabilidad de cometer un error de tipo II denotada con la letra griega beta β, depende de la diferencia entre los valores supuesto y real del parámetro de la población. Como es más fácil encontrar diferencias grandes, si la diferencia entre la estadística de muestra y el correspondiente parámetro de población es grande, la probabilidad de cometer un error de tipo II, probablemente sea pequeña.

El estudio y las conclusiones que obtengamos para una población cualquiera, se habrán apoyado exclusivamente en el análisis de una parte de ésta. De la probabilidad con la que estemos dispuestos a asumir estos errores, dependerá, por ejemplo, el tamaño de la muestra requerida. Las contrastaciones se apoyan en que los datos de partida siguen una distribución normal

Existe una relación inversa entre la magnitud de los errores α y β: conforme a aumenta, β disminuye. Esto obliga a establecer con cuidado el valor de a para las pruebas estadísticas.

Lo ideal sería establecer α y β.En la práctica se establece el nivel α y para disminuir el Error β se incrementa el número de observaciones en la muestra, pues así se acortan los límites de confianza respecto a la hipótesis planteada. La meta de las pruebas estadísticas es rechazar la hipótesis planteada. En otras palabras, es deseable aumentar cuando ésta es verdadera, o sea, incrementar lo que se llama **poder de la prueba (1- β)** La aceptación de la hipótesis planteada debe interpretarse como que la información aleatoria de la muestra disponible no permite detectar la falsedad de esta hipótesis.

Paso 3: Cálculo del valor estadístico de prueba

Valor determinado a partir de la información muestral, que se utiliza para determinar si se rechaza la hipótesis nula., existen muchos estadísticos de prueba para nuestro caso utilizaremos los estadísticos z y t. La elección de uno de estos depende de la cantidad de muestras que se toman, si las muestras son de la prueba son iguales a 30 o mas se utiliza el estadístico z, en caso contrario se utiliza el estadístico t.

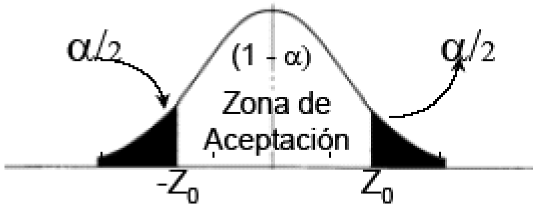
Tipos de prueba

a) Prueba bilateral o de dos extremos: la hipótesis planteada se formula con la igualdad

Ejemplo

$H0 : \mu = 200$

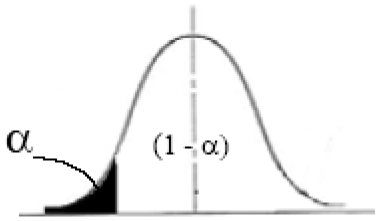
$H1 : \mu \neq 200$

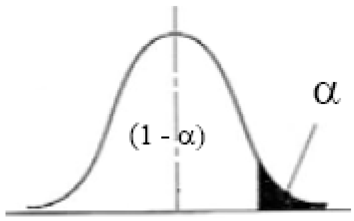


b) Pruebas unilateral o de un extremo: la hipótesis planteada se formula con ≥ o ≤

$H0 : \mu \geq 200$ $H0 : \mu \leq 200$

$H1 : \mu < 200$ $H1 : \mu > 200$





En las pruebas de hipótesis para la media (μ), cuando se conoce la desviación estándar (σ) poblacional, o cuando el valor de la muestra es grande (30 o más), el valor estadístico de prueba es z y se determina a partir de:

$$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

El valor estadístico z, para muestra grande y desviación estándar poblacional desconocida se determina por la ecuación:

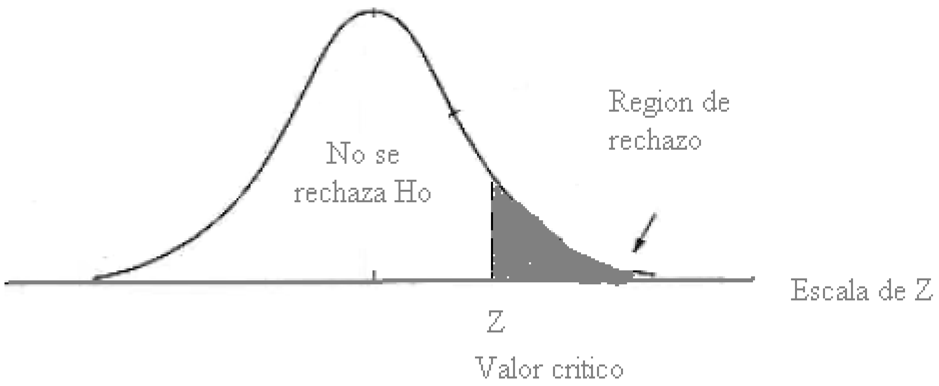
$$Z = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

En la prueba para una media poblacional con muestra pequeña y desviación estándar poblacional desconocida se utiliza el valor estadístico t.

$$t = \frac{\bar{x} - \mu}{S / \sqrt{n}}$$

Paso 4: Formular la regla de decisión

SE establece las condiciones específicas en la que se rechaza la hipótesis nula y las condiciones en que no se rechaza la hipótesis nula. La región de rechazo define la ubicación de todos los valores que son tan grandes o tan pequeños, que la probabilidad de que se presenten bajo la suposición de que la hipótesis nula es verdadera, es muy remota



Distribución muestral del valor estadístico z, con prueba de una cola a la derecha

Valor crítico: Es el punto de división entre la región en la que se rechaza la hipótesis nula y la región en la que no se rechaza la hipótesis nula.

Paso 5: Tomar una decisión.

En este último paso de la prueba de hipótesis, se calcula el estadístico de prueba, se compara con el valor crítico y se toma la decisión de rechazar o no la hipótesis nula. Tenga presente que en una prueba de hipótesis solo se puede tomar una de dos decisiones: aceptar o rechazar la hipótesis nula. Debe subrayarse que siempre existe la posibilidad de rechazar la hipótesis nula cuando no debería haberse rechazado (error tipo I). También existe la posibilidad de que la hipótesis nula se acepte cuando debería haberse rechazado (error de tipo II).

4.- Ejemplo en la cual se indica el procedimiento para la prueba de hipótesis

Ejemplo

El jefe de la Biblioteca Especializada de la Facultad de Ingeniería Eléctrica y Electrónica de la UNAC manifiesta que el número promedio de lectores por día es de 350. Para confirmar o no este supuesto se controla la cantidad de lectores que utilizaron la biblioteca durante 30 días. Se considera el nivel de significancia de 0.05

Datos:

Día	Usuarios	Día	Usuarios	Día	Usuario
1	356	11	305	21	429

2	427	12	413	22	376
3	387	13	391	23	328
4	510	14	380	24	411
5	288	15	382	25	397
6	290	16	389	26	365
7	320	17	405	27	405
8	350	18	293	28	369
9	403	19	276	29	429
10	329	20	417	30	364

Solución: Se trata de un problema con una media poblacional: muestra grande y desviación estándar poblacional desconocida.

Paso 01: Seleccionamos la hipótesis nula y la hipótesis alternativa

Ho: μ=350

Ha: $\mu\neq$ 350

Paso 02: Nivel de confianza o significancia 95%

α=0.05

Paso 03: Calculamos o determinamos el valor estadístico de prueba

De los datos determinamos: que el estadístico de prueba es t, debido a que el número de muestras es igual a 30, conocemos la media de la población, pero la desviación estándar de la población es desconocida, en este caso determinamos la desviación estándar de la muestra y la utilizamos en la formula reemplazando a la desviación estándar de la población.

$$Z = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

Calculamos la desviación estándar muestral y la media de la muestra empleando Excel, lo cual se muestra en el cuadro que sigue.

Columna1	
Media	372.8
Error típico	9.56951578
Mediana	381
Moda	405
Desviación estándar	52.4143965
Varianza de la muestra	2747.26897
Curtosis	0.36687081
Coeficiente de asimetría	0.04706877
Rango	234
Mínimo	276
Máximo	510
Suma	11184
Cuenta	30

Nivel de confianza (95.0%)	19.571868

Paso 04: Formulación de la regla de decisión.

La regla de decisión la formulamos teniendo en cuenta que esta es una prueba de dos colas, la mitad de 0.05, es decir 0.025, está en cada cola. El área en la que no se rechaza Ho esta entre las dos colas, es por consiguiente 0.95. El valor crítico para 0.05 da un valor de Zc = 1.96.

Por consiguiente, la regla de decisión: es rechazar la hipótesis nula y aceptar la hipótesis alternativa, si el valor Z calculado no queda en la región comprendida entre -1.96 y +1.96. En caso contrario no se rechaza la hipótesis nula si Z queda entre -1.96 y +1.96.

Paso 05: Toma de decisión.

En este último paso comparamos el estadístico de prueba calculado mediante el Software Minitab que es igual a Z = 2.38 y lo comparamos con el valor critico de Zc = 1.96. Como el estadístico de prueba calculado cae a la derecha del valor critico de Z, se rechaza Ho. Por tanto, no se confirma el supuesto del Jefe de la Biblioteca.

3.5 Pruebas para proporciones.

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

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

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

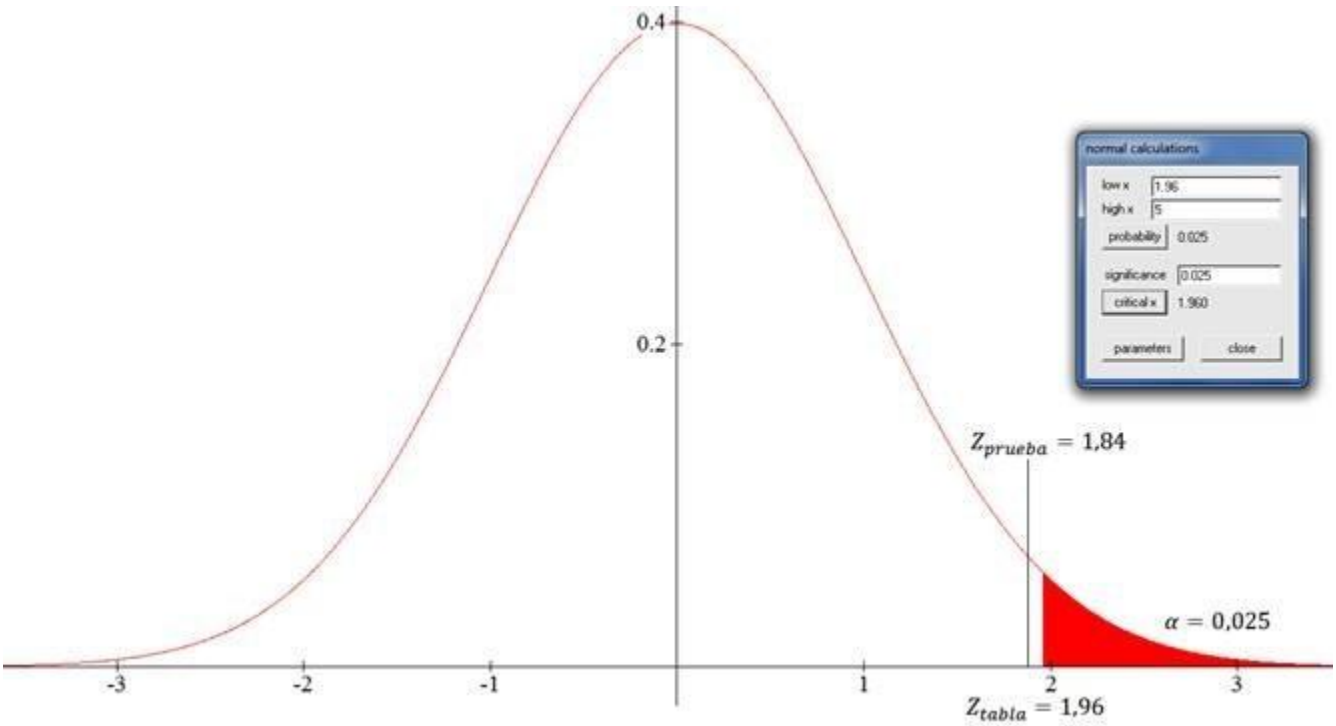
Por lo tanto se debe utilizar la fórmula con el factor finito de corrección.

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

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:

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

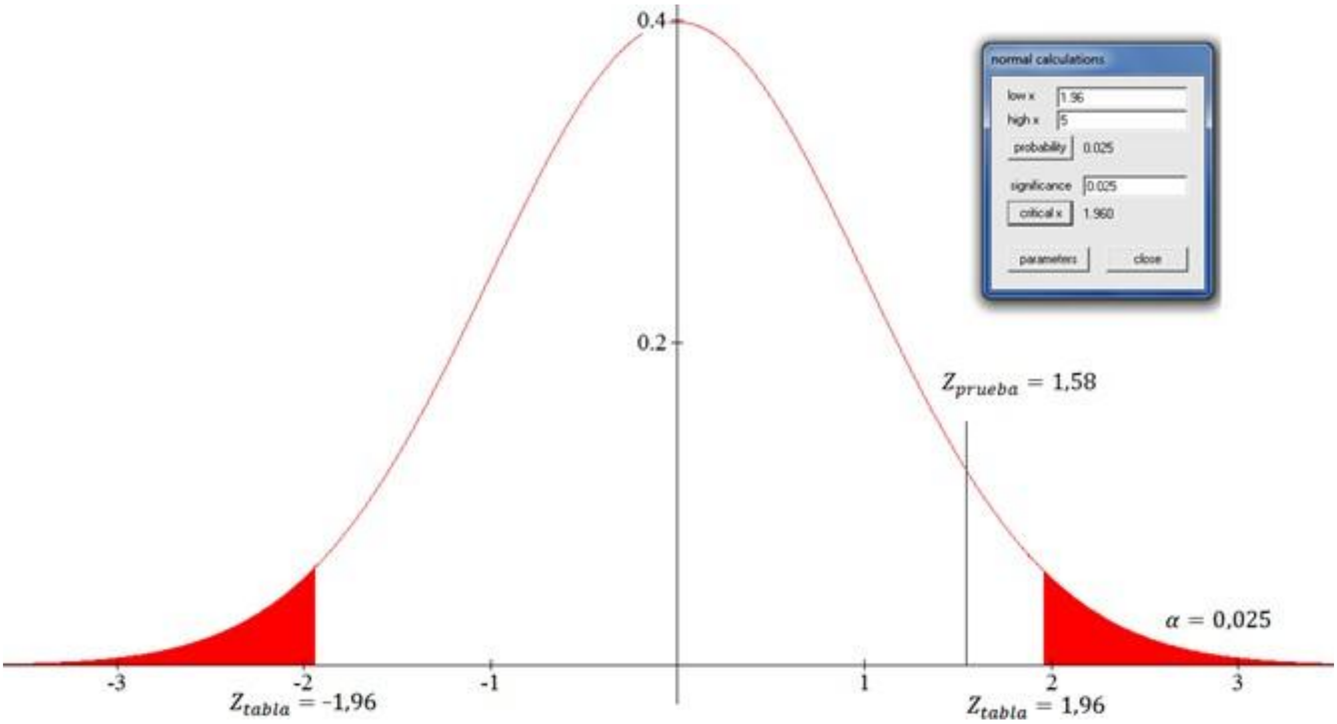
$$Z_{prueba} = \frac{\frac{18}{20} - \frac{14}{20}}{\sqrt{0{,}8(1-0{,}8)\left(\frac{1}{20} + \frac{1}{20}\right)}} = 1{,}58$$

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

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

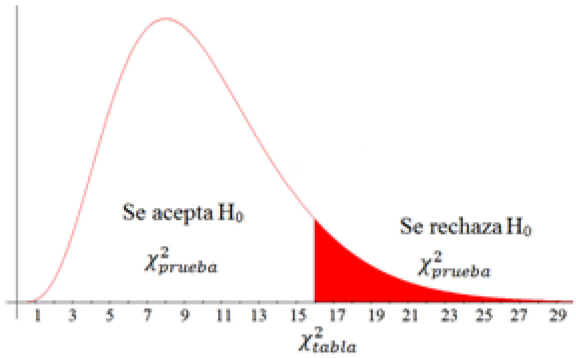
Por lo tanto el valor estadístico de prueba para este caso es la prueba *ji cuadrado* o conocida también como *chi cuadrado*

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.

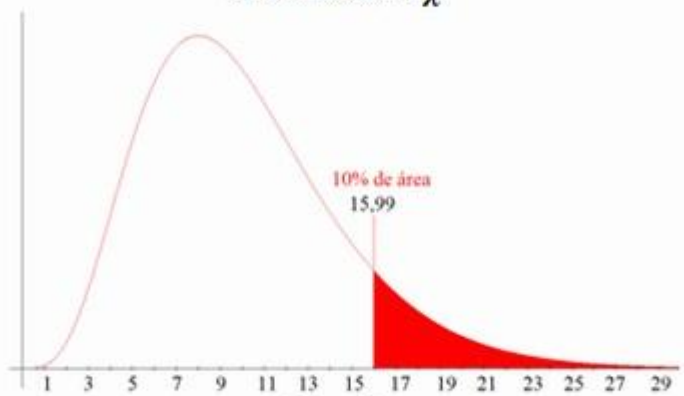
Solución:

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 $\chi^2_{tabla} = 21,026$

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D
1	r	4		
2	k	5		
3				
4	(r-1)(k-1)	12	=(B1-1)*(B2-1)	
5				
6	α	0,05		
7	χ^2_{tabla}	21,026	=PRUEBA.CHI.INV(B6;B4)	
8				
9	α	0,01		
10	χ^2_{tabla}	26,2170	=PRUEBA.CHI.INV(B9;B4)	

2) La siguiente tabla muestra las frecuencias observadas y las frecuencias esperadas al lanzar un dado 60 veces. Contrastar la hipótesis de que el dado es bueno, con un nivel de significación de 0,01.

Cara del dado	1	2	3	4	5	6
Frecuencia observada	6	8	9	15	14	8
Frecuencia esperada	10	10	10	10	10	10

Solución:

$r = 2$

$k = 6$

$$\alpha = 0{,}01$$

Las hipótesis son:

H_0:Todas las proporciones de la población son iguales.

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

Los grados de libertad se calculan aplicando la fórmula:

$Grados\ de\ libertad = (2-1)(6-1)$

$Grados\ de\ libertad = 5$

Con lectura en la tabla con 5 grados de libertad y 0,01 de área se obtiene $\chi^2_{tabla} = 15{,}086$

Calculando χ^2_{prueba} se obtiene:

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

$$\chi^2_{prueba} = \frac{(6-10)^2}{10} + \frac{(8-10)^2}{10} + \frac{(9-10)^2}{10} + \frac{(15-10)^2}{10} + \frac{(14-10)^2}{10} + \frac{(8-10)^2}{10}$$

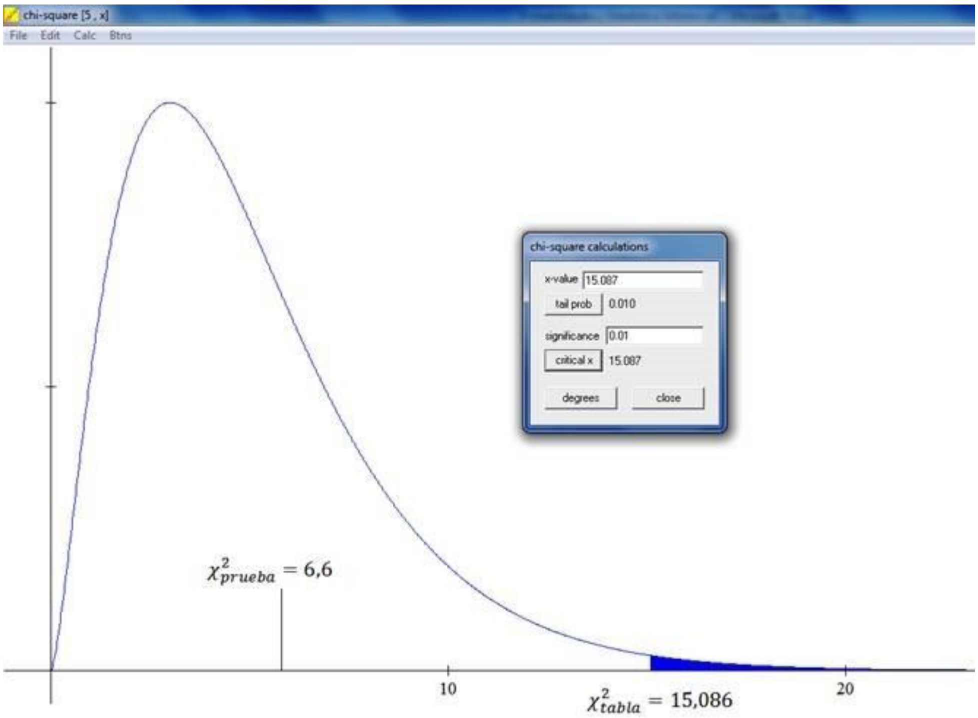
$$\chi^2_{prueba} = 1{,}6 + 0{,}4 + 0{,}1 + 2{,}5 + 1{,}6 + 0{,}4$$

$$\chi^2_{prueba} = 6{,}6$$

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E	F	G
1	Cara del dado	1	2	3	4	5	6
2	Frecuencia observada	6	8	9	15	14	8
3	Frecuencia esperada	10	10	10	10	10	10
4							
5	α	0,01					
6	r	2	=CONTAR(B2:B3)				
7	k	6	=CONTAR(B2:G2)				
8	(r-1)(k-1)	5	=(B6-1)*(B7-1)				
9	χ^2_{tabla}	15,086	=PRUEBA.CHI.INV(B5;B8)				
10							
11	Probabilidad de χ^2_{prueba}	0,2521	=PRUEBA.CHI(B2:G2;B3:G3)				
12	$\chi^2_{prueba} = \sum \frac{(o_i - e_i)^2}{e_i}$	6,6	=PRUEBA.CHI.INV(B11;B8)				
13							

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



Decisión:

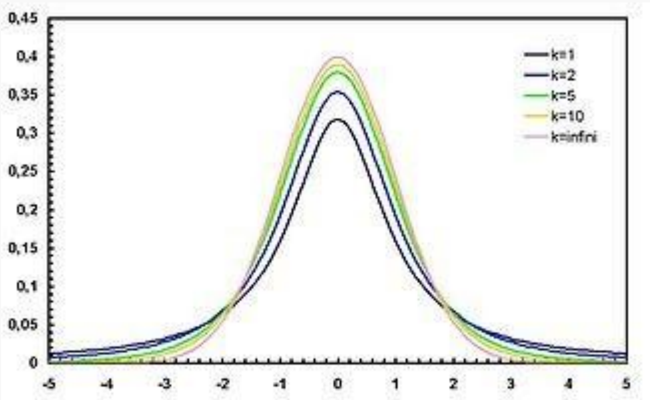
H_0 es aceptada, ya que χ^2_{prueba} (6,6) es menor que χ^2_{tabla}(15,086), por lo tanto, se concluye que todas las proporciones de la población son iguales, es decir, el dado es bueno.

UNIDAD IV

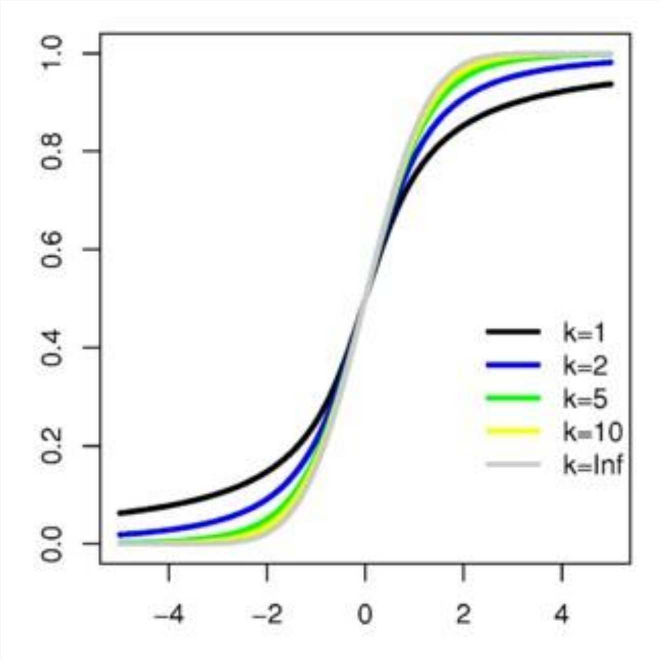
PRUEBAS DE HIPÓTESIS CON DOS MUESTRAS Y VARIAS MUESTRAS DE DATOS NUMÉRICOS.

4.1.- Distribuciones normal y t de Student.

Distribución t de student



Función de densidad de probabilidad



Función de distribución de probabilidad

Parámetros	grados de libertad (real)
Dominio	
Función de densidad (pdf)	
Función de distribución (cdf)	donde es la función hipergeométrica
Media	para , indefinida para otros valores
Mediana	

Moda	
Varianza	para , indefinida para otros valores
Coeficiente de simetría	para
Curtosis	para
Entropía	• : función digamma, • : función beta
Función generadora de momentos (mgf)	(No definida)
[editar datos en Wikidata]	

En probabilidad y estadística, la distribución t (de Student) es una distribución de probabilidad que surge del problema de estimar la media de una población normalmente distribuida cuando el tamaño de la muestra es pequeño.

Aparece de manera natural al realizar la prueba t de Student para la determinación de las diferencias entre dos varianzas muestrales y para la construcción del intervalo de confianza para la diferencia entre las partes de dos poblaciones cuando se desconoce la desviación típica de una población y esta debe ser estimada a partir de los datos de una muestra.

Fue desarrollada por William Sealy Gosset, bajo el seudónimo *Student*.

Caracterización

La distribución t de Student es la distribución de probabilidad del cocient

Donde

- *Z* es una variable aleatoria distribuida según una normal típica (de media nula y varianza 1).

- *V* es una variable continua que sigue una distribución χ^2 con grados de libertad.
- *Z* y *V* son independientes

Si μ es una constante no nula, el cociente es una variable aleatoria que sigue

la distribución t de Student no central con parámetro de no-centralidad .

Aparición y especificaciones de la distribución *t de Student*

Supongamos que $X_1,..., X_n$ son variables aleatorias independientes distribuidas normalmente, con media μ y varianza σ^2. Sea la media muestral. Entonces sigue una distribución normal de media 0 y varianza 1.

Sin embargo, dado que la desviación estándar no siempre es conocida de antemano, Gosset estudió un cociente relacionado, es la cuasi varianza maestral y

demostró que la función de densidad de *T* es donde es igual a $n - 1$.

La distribución de *T* se llama ahora la distribución-*t* de Student.

El parámetro representa el número de *grados de libertad*. La distribución depende

de , pero no de o , lo cual es muy importante en la práctica. Intervalos de confianza derivados de la distribución *t* de Student

El procedimiento para el cálculo del intervalo de confianza basado en la t de Student consiste en estimar la desviación típica de los datos S y calcular el error estándar de la

media: ,siendo entonces el intervalo de confianza para la media: .Es este resultado el que se utiliza en el test de Student: puesto que la diferencia de las medias de muestras de dos distribuciones normales se distribuye también normalmente, la distribución *t* puede usarse para examinar si esa diferencia puede razonablemente suponerse igual a cero. Para

efectos prácticos el valor esperado y la varianza son: y para Historia La distribución de Student fue descrita en el año 1908 por William Sealy Gosset. Gosset trabajaba en una fábrica de cerveza, Guinness, que prohibía a sus empleados la publicación de artículos científicos debido a una difusión previa de secretos industriales. De ahí que Gosset publicase sus resultados bajo el seudónimo de *Student*.[1]

Distribución t de Student no estandarizada

La distribución t puede generalizarse a 3 parámetros, introduciendo un parámetro

vocacional y otro de escala . El resultado es una distribución t de Student

Las pruebas de significación estadística sirven para comparar variables entre distintas muestras. Si la distribución de la muestra es normal se aplican los llamados tests paramétricos. Si la distribución no puede asumirse normal se aplican las pruebas no paramétricas. Hay que tener siempre en cuenta que los tests paramétricos son más potentes y dan más información que los no paramétricos, por lo que, si pueden usarse, se prefieren. El uso indiscriminado de muestras de distribución fuera de la normalidad conlleva el peligro de obtener conclusiones erróneas.

En general, con pocos datos, es preferible, si no es difícil ni conlleva un alto coste de tiempo y dinero, realizar más determinaciones para poder aplicar pruebas paramétricas al lograr una distribución normal. El teorema del límite nos dice que si el tamaño de la muestra es suficiente, la distribución siempre tiende a ser normal, lo que juega a nuestro favor. También hay que tener presente que la mayoría de las veces no hay suficiente tiempo y dinero para realizar un número elevado de pruebas para calcular la variancia de la población (Ec. 1.1), por lo que se recurre a la variancia de la muestra (Ec. 1.2)

$\sigma^2 = (\Sigma(x_i-\mu)^2)/N$ (Ec. 1.1)

$s^2 = [\Sigma(x_i-x)^2]/n-1$ (Ec. 1.2)

x_i: valor puntual o media de un conjunto pequeño de datos (muestra).
μ: media de un conjunto infinito de datos (población) o valor aceptado como verdadero.
x: media de un conjunto finito de datos de la muestra.

Es muy importante tener en cuenta que en las pruebas de significación estadística siempre se plantea la hipótesis nula "H_0" (no hay diferencias significativas entre los estadísticos de las muestras comparadas), y la hipótesis alternativa "H_1" (hay diferencias significativas). Se obtiene mucha mayor información cuando se puede rechazar la hipótesis nula, lo que quiere decir que los estadísticos de las muestras que se comparan son diferentes entre sí con una probabilidad mayor del 95%. Si no se puede rechazar la hipótesis nula ($p > 0,05$) se pierde mucha información porque no se puede decir que sean iguales, ni que sean diferentes porque la probabilidad es menor del 95%.

Para analizar la dispersión se usa el concepto de "**cuadrados medios (CM)**". El cuadrado medio es la suma de los cuadrados de las diferencias de los valores individuales con respecto a un valor central (generalmente la media), partido por los grados de libertad que tiene esa muestra. Elevar al cuadrado cada diferencia tiene la ventaja de que hacemos positivas todas las diferencias, porque en realidad lo que queremos valorar es la distancia de los valores al valor central, sin importarnos si están por arriba o por debajo. El dividir por los grados de libertad, sencillamente nos permite comparar sumatorios de cuadrados "**SC**" entre grupos con distinto tamaño de muestra. Sería absurdo decir que una muestra con 50 valores tiene más dispersión que otra de 5 valores porque tiene un SC mayor, evidentemente la suma de 50 diferencias será mayor que la suma de 5, a no ser que las diferencias en la muestra con 5 valores sean muy grandes. Si recapitulamos un poco podemos imaginar la variancia como un cuadrado medio (Ec. 1.3)

$s^2 = [\Sigma(x_i\text{-}x)^2]/n\text{-}1$ (Ec. 1.3)

Brevemente se exponen en la tabla 1 métodos estadísticos que se pueden aplicar para comparar distintos tipos de variables.

Tabla 1. Relación entre variables "V"

V. independiente	V. dependiente	Representación	Análisis estadístico
Cualitativa	Cualitativa	Tabla de contingencia	Comparación de proporciones[1]
Cualitativa (binaria o de más de 2 categorías)	Cuantitativa	Tabla	Comparación de medias[2]
Cuantitativa	Cuantitativa	Diagrama de dispersión	Modelo de regresión[3]

1 Ejemplo: prueba de $\chi 2$.
2 Ejemplo: prueba de t de Student o análisis de variancia (F de Snedecor).
3 Ejemplo : prueba de regresión lineal por Ajuste de Mínimos Cuadrados.

Para comprobar si la relación entre variables en una muestra es significativa, es decir, es poco probable que dicha relación se explique por fluctuaciones aleatorias del muestreo, se realizan las pruebas de significación estadística.

Distribución "T" de Student

Descrita por William S. Gosset en 1908. Publicaba bajo el pseudónimo de "Student" mientras trabajaba para la cervecería Guinnes en Irlanda. Está diseñada para probar hipótesis en estudios con muestras pequeñas (menores de 30)

La fórmula general para la T de Student es la siguiente:

$$t = \frac{X - \mu}{s/\sqrt{n}}$$

En donde el numerador representa la diferencia a probar y el denominador la desviación estándar de la diferencia llamado también Error Estándar. En esta fórmula t representa al valor estadístico que estamos buscando X barra es el promedio de la variable analizada de la muestra, y miu es el promedio poblacional de la variable a estudiar. En el denominador tenemos a s como representativo de la desviación estándar de la muestra y n el tamaño de ésta.

Grados de libertad: El número de grados de libertad es igual al tamaño de la muestra (número de observaciones independientes) menos 1.

gl = df = (n – 1)

Si pudiera expresar en un cierto número de pasos para resolver un problema de t de student tendría que declarar los siguientes:

Paso 1. Plantear la hipótesis nula (H_0) y la hipótesis alternativa (H_1). La hipótesis alternativa plantea matemáticamente lo que queremos demostrar, en tanto que la hipótesis nula plantea exactamente lo contrario.

Paso 2. Determinar el nivel de significancia (rango de aceptación de la hipótesis alternativa), a.

Se considera un nivel alfa de: 0.05 para proyectos de investigación; 0.01 para aseguramiento de la calidad; y 0.10 para estudios o encuestas de mercadotecnia.

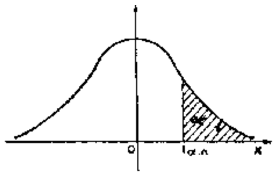
Paso 3. Evidencia muestral, se calcula la media y la desviación estándar a partir de la muestra.

Paso 4. Se aplica la distribución T de Student para calcular la probabilidad de error por medio de la fórmula general presentada al principio y se contrasta con el valor T obtenido de la tabla correspondiente.

Paso 5. En base a la evidencia disponible se acepta o se rechaza la hipótesis alternativa. Si la probabilidad de error (p) es mayor que el nivel de significancia se rechaza la hipótesis alternativa. Si la probabilidad de error (p) es menor que el nivel de significancia se acepta la hipótesis alternativa.

Por supuesto que al final lo que tenemos que contrastar es el valor de T que hayamos obtenido en el problema contra el valor T crítico que obtenemos de la tabla de T de Student.

Aquí puedes bajar la tabla de T de Student: Tabla T de Student un extremo y Tabla T de Student de dos extremos: Tabla T de Student dos extremos.



$\alpha/2$ / gl	0,40	0,30	0,20	0,10	0,050	0,025	0,010	0,005	0,001	0,0005
1	0,325	0,727	1,376	3,078	6,314	12,71	31,82	63,66	318,3	636,6
2	0,289	0,617	1,061	1,886	2,920	4,303	6,965	9,925	22,33	31,60
3	0,277	0,584	0,978	1,638	2,353	3,182	4,541	5,841	10,22	12,94
4	0,271	0,569	0,941	1,533	2,132	2,776	3,747	4,604	7,173	8,610
5	0,267	0,559	0,920	1,476	2,015	2,571	3,365	4,032	5,893	6,859
6	0,265	0,553	0,906	1,440	1,943	2,447	3,143	3,707	5,208	5,959
7	0,263	0,549	0,896	1,415	1,895	2,365	2,998	3,499	4,785	5,405
8	0,262	0,546	0,889	1,397	1,860	2,306	2,896	3,355	4,501	5,041
9	0,261	0,543	0,883	1,383	1,833	2,262	2,821	3,250	4,297	4,781
10	0,260	0,542	0,879	1,372	1,812	2,228	2,764	3,169	4,144	4,587
11	0,260	0,540	0,876	1,363	1,796	2,201	2,718	3,106	4,025	4,437
12	0,259	0,539	0,873	1,356	1,782	2,179	2,681	3,055	3,930	4,318
13	0,259	0,538	0,870	1,350	1,771	2,160	2,650	3,012	3,852	4,221
14	0,258	0,537	0,868	1,345	1,761	2,145	2,624	2,977	3,787	4,140
15	0,258	0,536	0,866	1,341	1,753	2,131	2,602	2,947	3,733	4,073
16	0,258	0,535	0,865	1,337	1,746	2,120	2,583	2,921	3,686	4,015
17	0,257	0,534	0,863	1,333	1,740	2,110	2,567	2,898	3,646	3,965
18	0,257	0,534	0,862	1,330	1,734	2.101	2,552	2,878	3,611	3,922
19	0,257	0,533	0,861	1,328	1,729	2,093	2,539	2,861	3,579	3,883
20	0,257	0,533	0,860	1,325	1,725	2,086	2,528	2,845	3,552	3,850
21	0,257	0,532	0,859	1,323	1,721	2,080	2,518	2,831	3,527	3,819
22	0,256	0,532	0,858	1,321	1,717	2,074	2,508	2,819	3,505	3,792
23	0,256	0,532	0,858	1,319	1,714	2,069	2,500	2,807	3,485	3,767
24	0,256	0.531	0,857	1,318	1,711	2,064	2,192	2,797	3,467	3,745
25	0,256	0,531	0,856	1,316	1,708	2,060	2,485	2,787	3,450	3,725
26	0,256	0,531	0,856	1,315	1,706	2,056	2,479	2,779	3,435	3,707
27	0,256	0,531	0,855	1,314	1,703	2,052	2,473	2,771	3,421	3,690
28	0,256	0,530	0,855	1,313	1,701	2,048	2,467	2,763	3.408	3.674
29	0,256	0,530	0,854	1,311	1,699	2.045	2,462	2.756	3,396	3.659
30	0,256	0,530	0,854	1,310	1,697	2,042	2,457	2,750	3,385	3,646
40	0,255	0,259	0,851	1,303	1,648	2,021	2,423	2,704	3,307	3,551
50	0,255	0,528	0,849	1,298	1,676	2,009	2,403	2,678	3,262	3,495
60	0,254	0,527	0,848	1,296	1,671	2,000	2,390	2,660	3,232	3,460
80	0,254	0,527	0,846	1,292	1,664	1,990	2,374	2,639	3,195	3,415
100	0,254	0,256	0,845	1,290	1,660	1,984	2,365	2,626	3,174	3,389
200	0,254	0,525	0,843	1,286	1,653	1,972	2,345	2,601	3,131	3,339
500	0,253	0,525	0,842	1,283	1,648	1,965	2,334	2,586	3,106	3,310
∞	0,253	0,524	0,842	1,282	1,645	1,960	2,326	2,576	3,090	3,291

Si el resultado del problema cae en la región de H0 se acepta ésta, de lo contrario se rechaza. Por supuesto, si rechazas H0 aceptarás H1.

En la gráfica precedente se aprecian las regiones de aceptación y de rechazo con respecto a H0.

Ejercicio 1: Se aplica una prueba de autoestima a 25 personas quienes obtienen una calificación promedio de 62.1 con una desviación estándar de 5.83. Se sabe que el valor correcto de la prueba debe ser mayor a 60. ¿Existe suficiente evidencia para comprobar que no hay problemas de autoestima en el grupo seleccionado?

Paso 1. Hipótesis alternativa: la que se va a comprobar. El grupo no tiene problemas de autoestima. Valor de prueba para determinar autoestima mayor a 60. Hipótesis nula, lo contrario a la hipótesis alternativa.

H1 > 60;

H0 =< 60.

Paso 2. Determinar el nivel de significancia alfa: alfa = 0.05.

Paso 3. Resultados de la evidencia muestral: X = 62.1; s = 5.83

Paso 4. Aplicar la distribución de probabilidad calculando T:

$$T = \frac{62.1 - 60}{5.83/\sqrt{25}}$$

El resultado de la ecuación es 1.8. Dado que 1.8 es mayor que 1.7109 cae en la región de H1 y se acepta la hipótesis alternativa. Si buscamos el valor de 1.8 bajo la curva normal encontraremos que es de 0.0359 el cual es menor que 0.05. La conclusión es que no hay problemas de autoestima en el grupo estudiado. Esto con el diseño de la investigación presentado.

Veamos en video como resolver problemas de T de student con Excel y con SPSS:

Ejemplo 2: Suponga que Ud. tiene una técnica que puede modificar la edad a la cual los niños comienzan a hablar. En su localidad, el promedio de edad en la cual un niño emite su primera palabra es de 13.0 meses. No se conoce la desviación estándar poblacional. Usted aplica dicha técnica a una muestra aleatoria de 15 niños. Los resultados arrojan que la edad media muestral en la que se pronuncia la primera palabra es de 11.0 meses, con una desviación estándar de 3.34. Pruebe la hipótesis de que la técnica afecta la edad en que los niños empiezan a hablar con un nivel de significancia alfa del 0.05.

Aquí las preguntas de la investigación serían ¿Cuáles son las hipótesis nula y la alternativa? y si con el procesamiento estadístico se puede afirmar que la técnica es efectiva para modificar la edad en que los niños empiezan a hablar.

Hipótesis nula: La técnica no afecta la edad en que los niños comienzan a hablar, matemáticamente sería, $H0 = 13.0$

Hipótesis alternativa: La técnica afecta la edad en que los niños comienzan a hablar, matemáticamente sería, $H1 \neq 13.0$

$$Tp = \frac{11 - 13}{3.34/\sqrt{15}}$$

El resultado de Tp es -2.32. Si lo comparamos con el resultado de T crítico o Tc obtenido de tablas con un nivel de significancia alfa de 0.05 y 14 grados de libertad para dos extremos, el resultado de Tc es 2.145

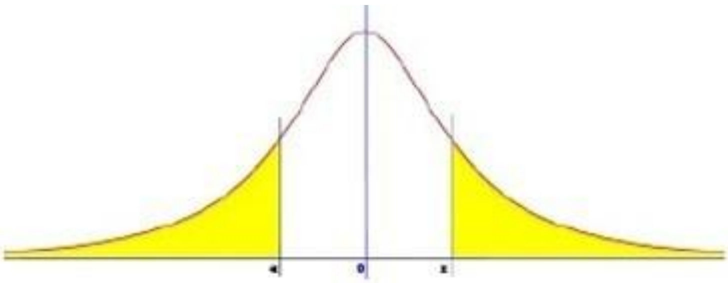


Tabla distribución t. Dos colas, probabilidad dentro(%)/fuera(0.00) del intervalo $\mu \pm t_{n-1}\sigma/\sqrt{n}$

Valor de t para un intervalo de confianza de	*90%*	*95%*	*98%*	*99%*
Valor crítico de |t| para valores de P de número de grados de libertad	*0.10*	*0.05*	*0.02*	*0.01*
1	6.31	12.71	31.82	63.66
2	2.92	4.30	6.96	9.92
3	2.35	3.18	4.54	5.84
4	2.13	2.78	3.75	4.60
5	2.02	2.57	3.36	4.03
6	1.94	2.45	3.14	3.71
7	1.89	2.36	3.00	3.50
8	1.86	2.31	2.90	3.36
9	1.83	2.26	2.82	3.25
10	1.81	2.23	2.76	3.17
12	1.78	2.18	2.68	3.05
14	1.76	2.14	2.62	2.98
16	1.75	2.12	2.58	2.92
18	1.73	2.10	2.55	2.88
20	1.72	2.09	2.53	2.85
30	1.70	2.04	2.46	2.75
50	1.68	2.01	2.40	2.68
∞	1.64	1.96	2.33	2.58

Con los resultados anteriores se rechaza la hipótesis nula y se decide que, la técnica afecta la edad en que los niños comienzan a hablar con un nivel de significancia de 0.05. El valor P correspondiente si lo buscamos en la curva normal de probabilidades sería de 0.010, por debajo del nivel de significancia.

Tabla de valores bajo la curva normal.

Ejemplo 3. Una profesora del programa de estudios para la mujer cree que la cantidad de cigarrillos fumados por las mujeres se ha incrementado en años recientes. Un censo realizado hace dos años con mujeres de una ciudad vecina mostró que el número promedio de cigarrillos fumados diariamente por una mujer era de 5.4 con una desviación estándar de 2.5. Para evaluar esta hipótesis, la profesora determinó el número de cigarrillos fumados diariamente por una muestra aleatoria de 120 mujeres que viven actualmente en la ciudad donde habita. Los datos muestran que el número de cigarrillos fumados diariamente por las 120 mujeres tiene una media de 6.1 y una desviación estándar de 2.7. Con esa información y un nivel de significancia de 0.05, ¿tiene razón la profesora al afirmar que la cantidad de cigarrillos fumados por las mujeres se ha incrementado?

$$Tp = \frac{6.1 - 5.4}{2.7/\sqrt{120}}$$

Los resultados de la ecuación muestran una Tp de 2.9 que, contrastada con la Tc obtenida de tablas para un extremo que resulta en 1.6449 cae en la región de rechazo de H0. Si calculamos P en tablas para 2.90 es 0.002, muy por debajo del 0.05 del nivel de significancia.

Ejercicio 4. La siguiente es una tabla de resultados del coeficiente intelectual entre niños que tienen buenas calificaciones en lectura y de aquellos que tienen bajas calificaciones en lectura. A un nivel de significancia del 0.05, ¿hay diferencia significativa entre el coeficiente intelectual entre los grupos? Utilice la prueba de T de Student contrastando las hipótesis contra el valor crítico.

No Progresivo	Coeficiente Intelectual de Buenos Lectores	Coeficiente Intelectual de Malos Lectores
1	105	94

2	110	95
3	100	93
4	102	93
5	103	92
6	104	95
7	108	100
Media		
DS		

Resuelva el problema anterior utilizando Excel y SPSS.

Igual y se puede utilizar una calculadora en línea específica para T de Student.

Ejercicios resueltos de prueba de hipótesis

1) Una empresa está interesada en lanzar un nuevo producto al mercado. Tras realizar una campaña publicitaria, se toma la muestra de 1 000 habitantes, de los cuales, 25 no conocían el producto. A un nivel de significación del 1% ¿apoya el estudio las siguientes hipótesis?

- a. Más del 3% de la población no conoce el nuevo producto.
- b. Menos del 2% de la población no conoce el nuevo producto

Datos:

n = 1000

x = 25

$$p = \frac{25}{1000} = 0,025$$

$$\alpha = 1\% = 0,01$$

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

Dónde:

x = ocurrencias

n = observaciones

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

p_0 = proporción propuesta

Solución:

a)

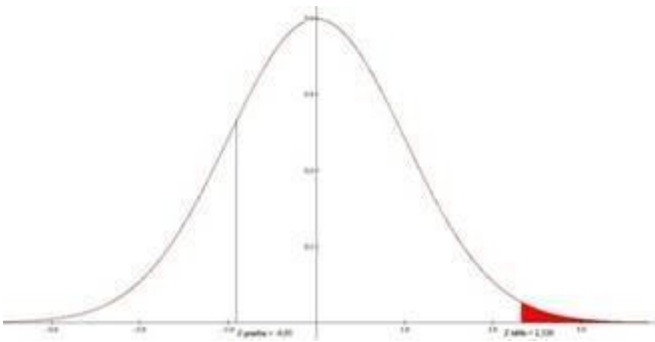
$p_o = 3\% = 0{,}03$

$Hipótesis\ Nula: H_0: p = p_0$

$Hipótesis\ Alternativa: H_1: p > p_0$

a = 0,01 $\Longrightarrow z_{tabla} = 2{,}326$

$$z_{prueba} = \frac{\frac{x}{n} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{\frac{25}{1000} - 0{,}03}{\sqrt{\frac{0{,}03(1-0{,}03)}{1000}}} = \frac{-0{,}005}{0{,}00539} = -0{,}93$$



Ho es aceptada, ya que zprueba (-0,93) es menor que ztabla (2,326), por lo que no es cierto que más del 3% de la población no conoce el nuevo producto.

En Excel

	A	B	C	D	E
1	Nivel de significación	0,01			
2	Z tabla	-2,32634787	=INV.NORM.ESTAND(B1)		
3	Z tabla	2,32634787	=B2*-1		
4	n	1000			
5	x	25			
6	po	0,03			
7					
8	$z_{prueba} = \frac{\frac{x}{n} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$	-0,93	=(B5/B4-B6)/RCUAD(B6*(1-B6)/B4)		
9					

b)

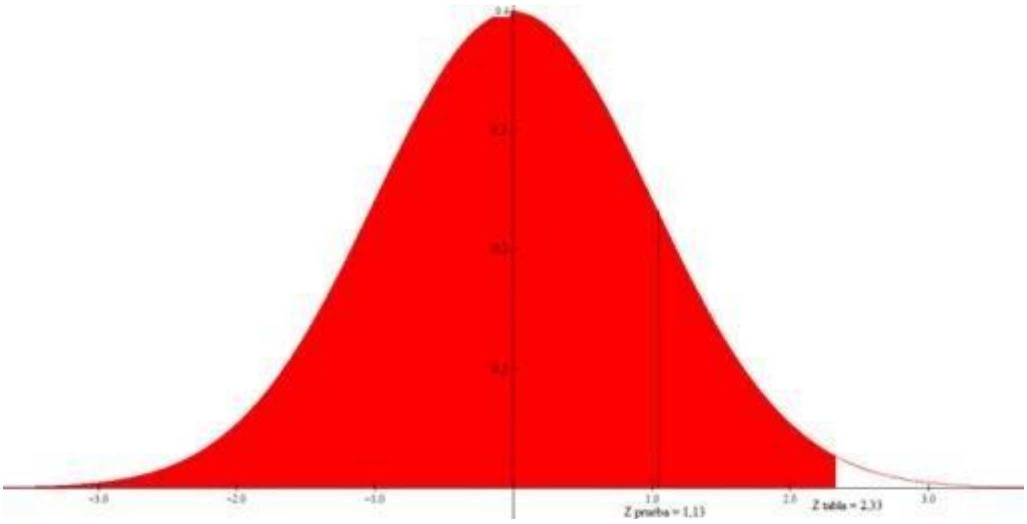
$p_o = 2\% = 0{,}02$

$Hipótesis\ Nula: H_0 : p = p_0$

$Hipótesis\ Alternativa: H_1 : p < p_0$

a = 0,01 $\Longrightarrow z_{tabla} = 2{,}326$

$$z_{prueba} = \frac{\frac{x}{n} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{\frac{25}{1000} - 0{,}02}{\sqrt{\frac{0{,}02(1-0{,}02)}{1000}}} = \frac{0{,}005}{0{,}00443} = 1{,}13$$



H0 es rechazada, ya que zprueba (1,13) es menor que ztabla (2,326), por lo que es cierto que menos del 2% de la población no conoce el nuevo producto.

2) Cuando las ventas medias, por establecimiento autorizado, de una marca de relojes caen por debajo de las 170,000 unidades mensuales, se considera razón suficiente para lanzar una campaña publicitaria que active las ventas de esta marca. Para conocer la evolución de las ventas, el departamento de marketing realiza una encuesta a 51 establecimientos autorizados, seleccionados aleatoriamente, que facilitan la cifra de ventas

del último mes en relojes de esta marca. A partir de estas cifras se obtienen los siguientes resultados: media = 169.411,8 unidades., desviación estándar = 32.827,5 unidades. Suponiendo que las ventas mensuales por establecimiento se distribuyen normalmente; con un nivel de significación del 5 % y en vista a la situación reflejada en los datos. ¿Se considerará oportuno lanzar una nueva campaña publicitaria?

Datos:

$\mu = 170000$

n = 51

$\bar{x} = 169441{,}8$

$\sigma = 32827{,}5$

$\alpha = 5\% = 0{,}05$

$$Z_{prueba} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

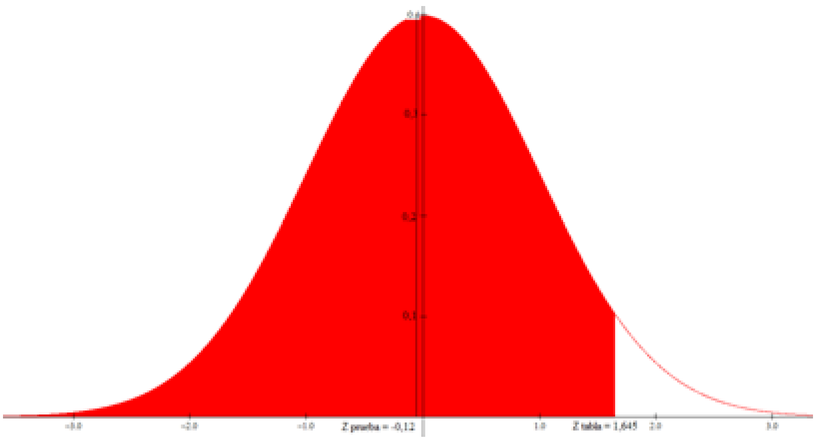
Solución:

Ho: (= 170000

H1: (< 170000

a = 0,05 $\Longrightarrow z_{tabla} = 1{,}645$

$$Z_{prueba} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{169441{,}8 - 170000}{\frac{32827{,}5}{\sqrt{51}}} = \frac{-555{,}2}{4596{,}769} = -0{,}12$$



Se rechaza Ho, porque zprueba (-0,12) es menor que ztabla (1,645), por lo tanto se acepta H1: (< 170000, y se debe considerar oportuno lanzar una nueva campaña publicitaria.

En Excel

	A	B	C	D
1	Nivel de significación	0,05		
2	Z tabla	-1,645	=INV.NORM.ESTAND(B1)	
3	Z tabla	1,645	=B2*-1	
4	μ	170000		
5	$\bar{x}$	169441,8		
6	n	51		
7	σ	32827,5		
8	$Z_{prueba} = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$	-0,12	=(B5-B4)/(B7/RCUAD(B6))	
9				
10				

3) Un gerente de ventas de libros universitarios afirma que en promedio sus representantes de ventas realiza 40 visitas a profesores por semana. Varios de estos representantes piensan que realizan un número de visitas promedio superior a 40. Una muestra tomada al azar durante 8 semanas reveló un promedio de 42 visitas semanales y una desviación estándar de 2 visitas. Utilice un nivel de confianza del 99% para aclarar esta cuestión.

Datos:

(= 40

$\bar{x} = 42$

n = 8

$S = 2$

Nivel de confianza del 99%

Nivel de significación = (100%-99%)/2 = 0,5% = 0,005

$$t_{prueba} = \frac{\bar{x}-\mu}{\frac{S}{\sqrt{n}}}$$

Solución:

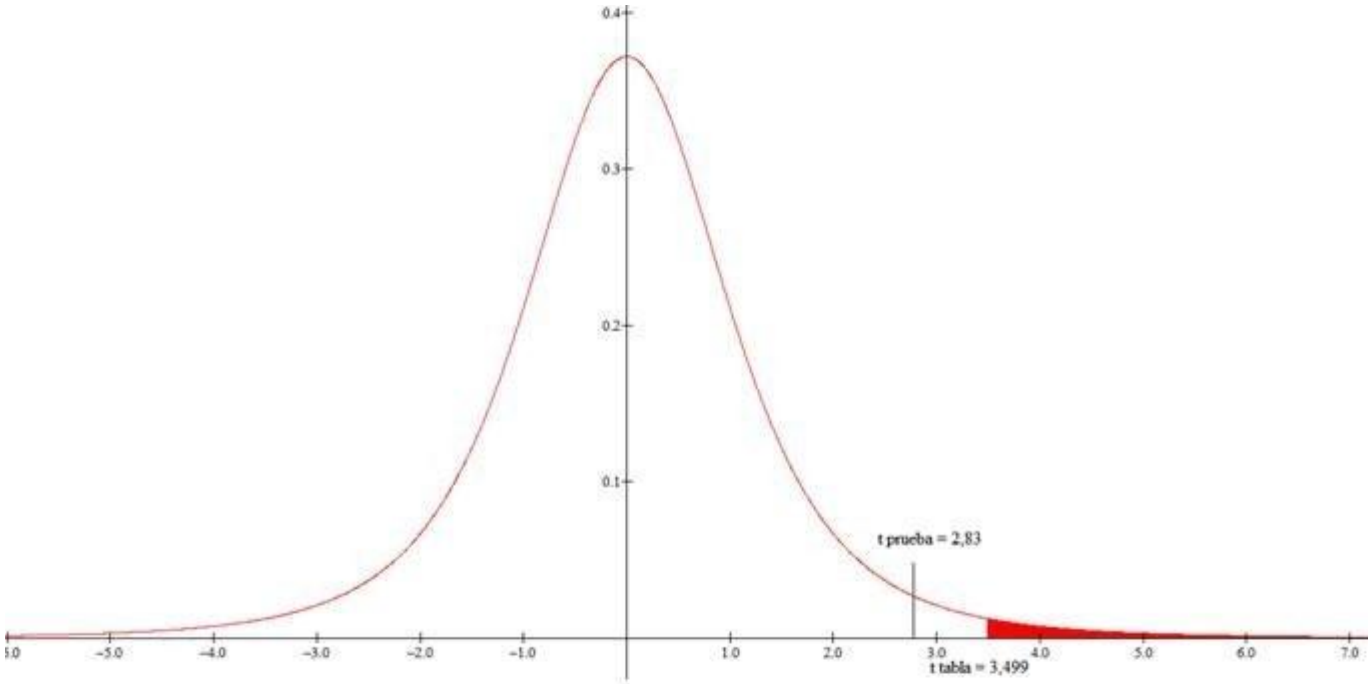
H0: (= 40

H1: (> 40

Grados de libertad: n-1 = 8-1 =7

a = 0,005 $\Longrightarrow t_{tabla} = 3{,}499$

$$t_{prueba} = \frac{\bar{x}-\mu}{\frac{S}{\sqrt{n}}} = \frac{42-40}{\frac{2}{\sqrt{8}}} = \frac{2}{0{,}7071} = 2{,}83$$



H0 es aceptada, ya que tprueba (2,83) es menor que ttabla (3,499), por lo que no es acertado pensar que están realizando un número de visitas promedio superior a 40.

En Excel

	A	B	C	D
1	Nivel de significación	0,005		
2	n	8		
3	grados de libertad: n-1	7	=B2-1	
4	t tabla	-3,499	=INV.T(B1;B3)	
5	t tabla	3,499	=B4*-1	
6	$\bar{x}$	42		
7	μ	40		
8	S	2		
9	$t_{prueba} = \dfrac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}$	2,83	=(B6-B7)/(B8/RCUAD(B2))	
10				
11				

4) Un investigador de mercados y hábitos de comportamiento afirma que el tiempo que los niños de tres a cinco años dedican a ver la televisión cada semana se distribuye normalmente con una media de 22 horas y desviación estándar 6 horas. Frente a este estudio, una empresa de investigación de mercados cree que la media es mayor y para probar su hipótesis toma una muestra de 64 observaciones procedentes de la misma población, obteniendo como resultado una media de 25. Si se utiliza un nivel de significación del 5%. Verifique si la afirmación del investigador es realmente cierta.

Datos:

$\mu = 22$

$\sigma = 6$

n = 64

$\bar{x} = 25$

a = 5% = 0,05

$$Z_{prueba} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

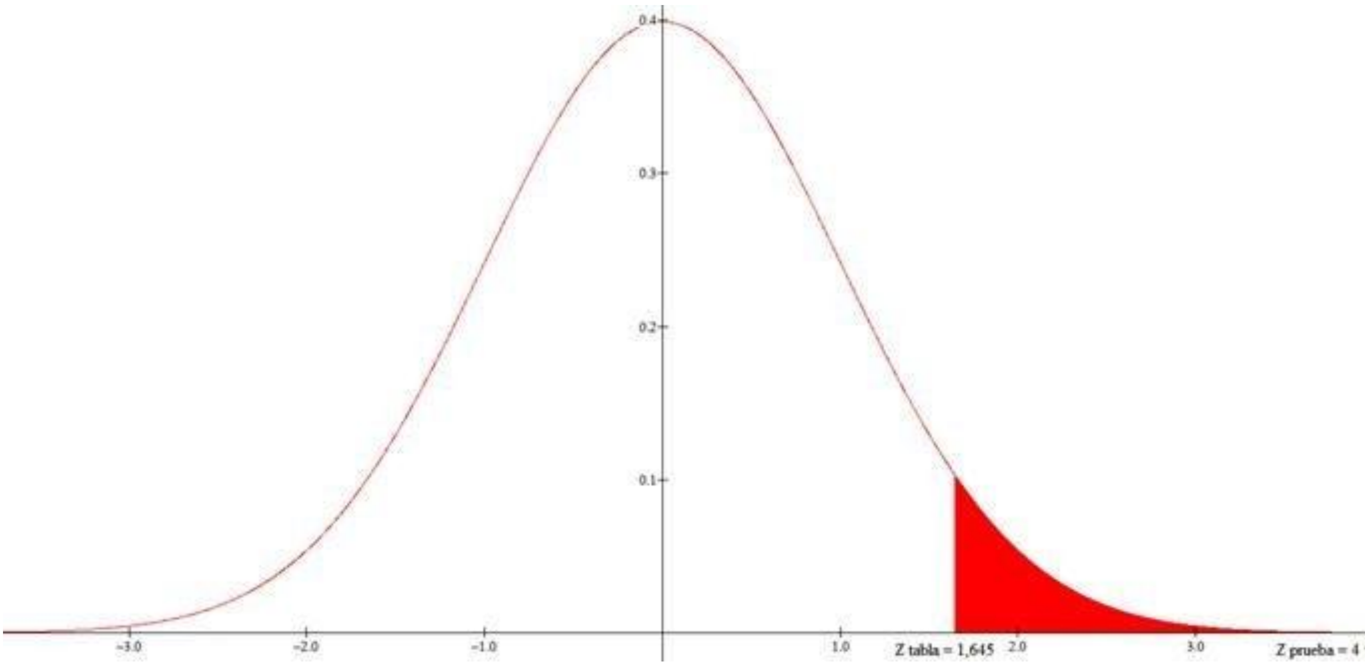
Solución:

H0: (= 22

H1: (> 22

a = 0,05 $\Longrightarrow z_{tabla} = 1{,}645$

$$Z_{prueba} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{25 - 22}{\frac{6}{\sqrt{64}}} = \frac{3}{0{,}75} = 4$$



Se rechaza Ho, porque zprueba (4) es mayor que ztabla (1,645), por lo tanto el tiempo que los niños de tres a cinco años dedican a ver la televisión es mayor de 22 horas, lo que implica que la empresa de investigación de mercados tiene la razón.

En Excel

	A	B	C	D
1	Nivel de significación	0,05		
2	Z tabla	-1,645	=INV.NORM.ESTAND(B1)	
3	Z tabla	1,645	=B2*-1	
4	μ	22		
5	$\bar{x}$	25		
6	n	64		
7	σ	6		
8	$Z_{prueba} = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$	4	=(B5-B4)/(B7/RCUAD(B6))	
9				
10				

5) Una organización que provee de alimentos a centros escolares le encarga a usted la siguiente investigación de su mercado interno. Está interesada en comparar los gastos resultantes de elaborar un plato muy usual, según el tipo de batería de cocina utilizado. Los gastos son de dos tipos: de energía, X1 (ya que los materiales y el diseño de la batería pueden hacer variar el tiempo necesario de cocción) y de condimentos, X2 (algunas baterías aconsejan la utilización de cantidades más pequeñas de aceite, sal, líquidos...). Se hicieron 5 pruebas con cada tipo de batería obteniéndose los siguientes resultados (gastos en dólares):

a)

Batería A		Batería B		Batería C	
X_1	X_2	X_1	X_2	X_1	X_2
32	48	28	46	31	43
30	45	26	39	30	42
31	45	30	45	32	44
33	47	28	47	30	44
29	44	29	48	27	41

b)

Batería A		Batería B		Batería C	
X_1	X_2	X_1	X_2	X_1	X_2
32	48	32	45	31	45
31	47	29	44	28	46
26	39	30	45	28	47
29	48	31	43	30	42
32	44	30	44	27	41

Utilizando un análisis ANOVA, ¿qué puede inferirse a partir de los datos recabados?

Solución

1) Tomando los datos de la tabla a) con respecto a X1

Observación: n	Muestra: k		
	A	B	C
1	32	28	31
2	30	26	30
3	31	30	32
4	33	28	30
5	29	29	27
Total	155	141	150

Cálculo de las medias aritméticas

$$\bar{x} = \frac{\sum x_i}{n}$$

$$\bar{x}_1 = \frac{155}{5} = 31$$

$$\bar{x}_2 = \frac{141}{5} = 28{,}2$$

$$\bar{x}_3 = \frac{150}{5} = 30$$

Llenando la siguiente tabla para calcular la varianza muestral

Observación: n	Muestra: k			$(x_1 - \bar{x}_1)^2$	$(x_2 - \bar{x}_2)^2$	$(x_3 - \bar{x}_3)^2$
	A	B	C			
1	32	28	31	1	0,04	1
2	30	26	30	1	4,84	0
3	31	30	32	0	3,24	4
4	33	28	30	4	0,04	0
5	29	29	27	4	0,64	9
Total	155	141	150	10	8,8	14

$$s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$$

$$s_1^2 = \frac{10}{5-1} = 2{,}5$$

$$s_2^2 = \frac{8{,}8}{5-1} = 2{,}2$$

$$s_3^2 = \frac{14}{5-1} = 3{,}5$$

Tomando en cuenta los cálculos de las varianzas se evidencia que la batería B es la que tiene menos varianza, por lo que para el gasto de energía, ésta batería es la mejor.

Estimación interna de varianza (within estimate) s_w^2

$$s_w^2 = \frac{s_1^2 + s_2^2 + s_3^2 + \cdots + s_k^2}{k}$$

$$s_w^2 = \frac{2{,}5 + 2{,}2 + 3{,}5}{3} = \frac{8{,}2}{3} = 2{,}733$$

Estimación intermediante de varianza (between estimate) s_x^2

$$s_x^2 = n \cdot s_{\bar{x}}^2$$

Donde

$$s_{\bar{x}}^2 = \frac{\sum(\bar{x} - \bar{\bar{x}})^2}{k-1}$$

$$\bar{\bar{x}} = \frac{\sum \bar{x}_i}{k} = \frac{31 + 28{,}2 + 30}{3} = \frac{89{,}2}{3} = 29{,}733$$

$\bar{x}$	$(\bar{x} - \bar{\bar{x}})^2$

31	1,604
28,2	2,351
30	0,071
Total	4,026

Varianza de las medias aritméticas

$$s_{\bar{x}}^2 = \frac{\sum(\bar{x} - \bar{\bar{x}})^2}{k-1} = \frac{4{,}026}{3-1} = 2{,}013$$

Estimación intermediante de varianza

$$s_x^2 = n \cdot s_{\bar{x}}^2 = 5 \cdot 2{,}013 = 10{,}065$$

Planteamiento de hipótesis

H0: Todas las proporciones de la población son iguales.

H1: No todas las proporciones de la población son iguales.

$$F_{prueba} = \frac{s_x^2}{s_w^2} = \frac{10{,}065}{2{,}733} = 3{,}68$$

F tabla

Grados de libertad:

Numerador: k-1 = 3-1 =2

Denominador: k(n-1) = 3(5-1) =12

Nivel de significación del 1%

$$F_{tabla} = 6{,}93$$

Como Fprueba es menor que Ftabla , H0 se acepta, por lo tanto no existen diferencias reales entre la baterías.

En Excel

	A	B	C	D	E	F	G
1	n	A	B	C	$(x_1-\bar{x}_1)^2$	$(x_2-\bar{x}_2)^2$	$(x_3-\bar{x}_3)^2$
2	1	32	28	31	1	0,04	1
3	2	30	26	30	1	4,84	0
4	4	31	30	32	0	3,24	4
5	4	33	28	30	4	0,04	0
6	5	29	29	27	4	0,64	9
7	Total	155	141	150	10	8,8	14
8	media	31	28,2	30			
9							
10	$s^2=\frac{\sum(x_i-\bar{x})^2}{n-1}$		2,5				
11			2,2				
12	n	5	3,5				
13							
14	$s_w^2=\frac{s_1^2+s_2^2+s_3^2+\cdots+s_k^2}{k}$			2,733333			
15							
16			$(\bar{x}-\bar{\bar{x}})^2$				
17		31	1,604444				
18		28,2	2,351111				
19		30	0,071111				
20	media d me	29,73333333	4,026667				
21							
22		$s_{\bar{x}}^2=\frac{\sum(\bar{x}-\bar{\bar{x}})^2}{k-1}$	2,013333				
23							
24							
25							
26		$s_x^2=n\cdot s_{\bar{x}}^2$	10,06667	=B12*C22			
27							
28							
29		$F_{prueba}=\frac{s_x^2}{s_w^2}$	3,682927				
30							
31	α	0,01					
32	Grados de libertad						
33	Numerador	k-1	2				
34	Denominad	k(n-1)	12				
35	k	3					
36	n	5					
37							
38	F tabla	6,93	=DISTR.F.INV(B23;C25;C26)				

2) Tomando los datos de la tabla a) con respecto a X2

Observación: n	Muestra: k		
	A	B	C
1	48	46	43
2	45	39	42
3	45	45	44
4	47	47	44
5	44	48	41
Total	229	225	214

Cálculo de las medias aritméticas

$$\bar{x}=\frac{\sum x_i}{n}$$

$$\bar{x}_1=\frac{229}{5}=45{,}8$$

$$\bar{x}_2=\frac{225}{5}=45$$

$$\bar{x}_3 = \frac{214}{5} = 42{,}8$$

Llenando la siguiente tabla para calcular la varianza muestral

Observación: n	Muestra: k			$(x_1 - \bar{x}_1)^2$	$(x_2 - \bar{x}_2)^2$	$(x_3 - \bar{x}_3)^2$
	A	B	C			
1	48	46	43	4,84	1	0,04
2	45	39	42	0,64	36	0,64
3	45	45	44	0,64	0	1,44
4	47	47	44	1,44	4	1,44
5	44	48	41	3,24	9	3,24
Total	229	225	214	10,8	50	6,8

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

$$s_1^2 = \frac{10{,}8}{5-1} = 2{,}7$$

$$s_2^2 = \frac{50}{5-1} = 12{,}5$$

$$s_3^2 = \frac{6{,}8}{5-1} = 1{,}7$$

Tomando en cuenta los cálculos de las varianzas se evidencia que la batería C es la que tiene menos varianza, por lo que para el gasto de condimentos, ésta batería es la mejor.

Estimación interna de varianza (within estimate) s_w^2

$$s_w^2 = \frac{s_1^2 + s_2^2 + s_3^2 + \cdots + s_k^2}{k}$$

$$s_w^2 = \frac{2{,}7 + 12{,}5 + 1{,}7}{3} = 5{,}633$$

Estimación intermediante de varianza (between estimate) s_x^2

$$s_x^2 = n \cdot s_{\bar{x}}^2$$

Donde

$$s_{\bar{x}}^2 = \frac{\sum (\bar{x} - \bar{\bar{x}})^2}{k-1}$$

$$\bar{\bar{x}} = \frac{\sum \bar{x}_i}{k} = \frac{45{,}8 + 45 + 42{,}8}{3} = 44{,}533$$

$\bar{x}$	$(\bar{x} - \bar{\bar{x}})^2$
45,8	1,604

45	0,218
42,8	3,004
Total	4,827

Varianza de las medias aritméticas

$$s_{\bar{x}}^2 = \frac{\sum(\bar{x}-\bar{\bar{x}})^2}{k-1} = \frac{4{,}827}{3-1} = 2{,}413$$

Estimación intermediante de varianza

$$s_x^2 = n \cdot s_{\bar{x}}^2 = 5 \cdot 2{,}413 = 12{,}065$$

Planteamiento de hipótesis

H0: Todas las proporciones de la población son iguales.

H1: No todas las proporciones de la población son iguales.

$$F_{prueba} = \frac{s_x^2}{s_w^2} = \frac{12{,}065}{5{,}633} = 2{,}142$$

F tabla

Grados de libertad:

Numerador: k-1 = 3-1 =2

Denominador: k(n-1) = 3(5-1) =12

Nivel de significación del 1%

$$F_{tabla} = 6{,}93$$

Como Fprueba es menor que Ftabla , H0 se acepta, por lo tanto no existen diferencias reales entre la baterías.

3) Tomando los datos de la tabla b) con respecto a X1

Observación: n	Muestra: k		
	A	B	C
1	32	32	31
2	31	29	28
3	26	30	28
4	29	31	30
5	32	30	27
Total	150	152	144

Cálculo de las medias aritméticas

$$\bar{x} = \frac{\sum x_i}{n}$$

$$\bar{x}_1 = \frac{150}{5} = 30$$

$$\bar{x}_2 = \frac{152}{5} = 30{,}4$$

$$\bar{x}_3 = \frac{144}{5} = 28{,}8$$

Llenando la siguiente tabla para calcular la varianza muestral

Observación: n	Muestra: k			$(x_1 - \bar{x}_1)^2$	$(x_2 - \bar{x}_2)^2$	$(x_3 - \bar{x}_3)^2$
	A	B	C			
1	32	32	31	4	2,56	4,84
2	31	29	28	1	1,96	0,64
3	26	30	28	16	0,16	0,64
4	29	31	30	1	0,36	1,44
5	32	30	27	4	0,16	3,24
Total	150	152	144	26	5,2	10,8

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

$$s_1^2 = \frac{26}{5-1} = 6{,}5$$

$$s_2^2 = \frac{5{,}2}{5-1} = 1{,}3$$

$$s_3^2 = \frac{10{,}8}{5-1} = 2{,}7$$

Tomando en cuenta los cálculos de las varianzas se evidencia que la batería B es la que tiene menos varianza, por lo que para el gasto de energía, ésta batería es la mejor.

Estimación interna de varianza (within estimate) s_w^2

$$s_w^2 = \frac{s_1^2 + s_2^2 + s_3^2 + \cdots + s_k^2}{k}$$

$$s_w^2 = \frac{6{,}5 + 1{,}3 + 2{,}7}{3} = 3{,}5$$

Estimación intermediante de varianza (between estimate) s_x^2

$$s_x^2 = n \cdot s_{\bar{x}}^2$$

Donde

$$s_{\bar{x}}^2 = \frac{\sum (\bar{x} - \bar{\bar{x}})^2}{k-1}$$

$$\bar{\bar{x}} = \frac{\sum \bar{x}_i}{k} = \frac{30 + 30{,}4 + 28{,}8}{3} = 29{,}733$$

$\bar{x}$	$(\bar{x} - \bar{\bar{x}})^2$

30	0,071
30,4	0,444
28,8	0,871
Total	1,387

Varianza de las medias aritméticas

$$s_{\bar{x}}^2 = \frac{\sum(\bar{x} - \bar{\bar{x}})^2}{k-1} = \frac{1,387}{3-1} = 0,693$$

Estimación intermediante de varianza

$$s_x^2 = n \cdot s_{\bar{x}}^2 = 5 \cdot 0,693 = 3,465$$

Planteamiento de hipótesis

H0: Todas las proporciones de la población son iguales.

H1: No todas las proporciones de la población son iguales.

$$F_{prueba} = \frac{s_x^2}{s_w^2} = \frac{3,465}{3,5} = 0,99$$

F tabla

Grados de libertad:

Numerador: k-1 = 3-1 =2

Denominador: k(n-1) = 3(5-1) =12

Nivel de significación del 1%

$$F_{tabla} = 6,93$$

Como Fprueba es menor que Ftabla , H0 se acepta, por lo tanto no existen diferencias reales entre la baterías.

4) Tomando los datos de la tabla b) con respecto a X2

Observación: n	Muestra: k		
	A	B	C
1	48	45	45
2	47	44	46
3	39	45	47
4	48	43	42
5	44	44	41
Total	226	221	221

Cálculo de las medias aritméticas

$$\bar{x} = \frac{\sum x_i}{n}$$

$$\bar{x}_1 = \frac{226}{5} = 45{,}2$$

$$\bar{x}_2 = \frac{221}{5} = 44{,}2$$

$$\bar{x}_3 = \frac{221}{5} = 42{,}2$$

Llenando la siguiente tabla para calcular la varianza muestral

Observación: n	Muestra: k			$(x_1 - \bar{x}_1)^2$	$(x_2 - \bar{x}_2)^2$	$(x_3 - \bar{x}_3)^2$
	A	B	C			
1	48	45	45	7,84	0,64	0,64
2	47	44	46	3,24	0,04	3,24
3	39	45	47	38,44	0,64	7,84
4	48	43	42	7,84	1,44	4,84
5	44	44	41	1,44	0,04	10,24
Total	226	221	221	58,8	2,8	26,8

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

$$s_1^2 = \frac{58{,}8}{5-1} = 14{,}7$$

$$s_2^2 = \frac{2{,}8}{5-1} = 0{,}7$$

$$s_3^2 = \frac{26{,}8}{5-1} = 6{,}7$$

Tomando en cuenta los cálculos de las varianzas se evidencia que la batería B es la que tiene menos varianza, por lo que para el gasto de condimentos, ésta batería es la mejor.

Estimación interna de varianza (within estimate) s_w^2

$$s_w^2 = \frac{s_1^2 + s_2^2 + s_3^2 + \cdots + s_k^2}{k}$$

$$s_w^2 = \frac{14{,}7 + 0{,}7 + 6{,}7}{3} = 7{,}367$$

Estimación intermediante de varianza (between estimate) s_x^2

$$s_x^2 = n \cdot s_{\bar{x}}^2$$

Donde

$$s_{\bar{x}}^2 = \frac{\sum (\bar{x} - \bar{\bar{x}})^2}{k-1}$$

$$\bar{\bar{x}} = \frac{\sum \bar{x}_i}{k} = \frac{45{,}2 + 44{,}2 + 42{,}2}{3} = 44{,}533$$

$\bar{x}$	$(\bar{x}-\bar{\bar{x}})^2$
45,2	0,444
44,2	0,111
44,2	0,111
Total	0,667

Varianza de las medias aritméticas

$$s_{\bar{x}}^2 = \frac{\sum(\bar{x}-\bar{\bar{x}})^2}{k-1} = \frac{0{,}667}{3-1} = 0{,}333$$

Estimación intermediante de varianza

$$s_x^2 = n \cdot s_{\bar{x}}^2 = 5 \cdot 0{,}333 = 1{,}67$$

Planteamiento de hipótesis

H0: Todas las proporciones de la población son iguales.

H1: No todas las proporciones de la población son iguales.

$$F_{prueba} = \frac{s_x^2}{s_w^2} = \frac{1{,}67}{7{,}367} = 0{,}23$$

F tabla

Grados de libertad:

Numerador: k-1 = 3-1 =2

Denominador: k(n-1) = 3(5-1) =12

Nivel de significación del 1%

$$F_{tabla} = 6{,}93$$

Como Fprueba es menor que Ftabla , H0 se acepta, por lo tanto no existen diferencias reales entre la baterías.

Resumen de las varianzas

Varianza	Tabla a)			Tabla b)		
	A	B	C	A	B	C
X_1	2,5	2,2	3,5	6,5	1,3	2,7
X_2	2,7	12,5	1,7	14,7	0,7	6,7
Total	5,2	14,7	5,2	21,2	2	9,4

Media aritmética de las varianzas

A:

$$\bar{X} = \frac{5{,}2 + 21{,}2}{4} = 6{,}6$$

B:

$$\bar{X} = \frac{14{,}7 + 2}{4} = 4{,}175$$

C:

$$\bar{X} = \frac{5{,}2 + 9{,}4}{4} = 7{,}3$$

II) Pruebas de significación

II.A) Pruebas paramétricas

1. Prueba t de Student

Con esta prueba se pretende averiguar si dos muestras que tienen medias iguales, provienen de la misma población.

Hipótesis nula "H_0" → $\mu_1 = \mu_2$;
Hipótesis alternativa "H_1" → $\mu_1 \neq \mu_2$

La prueba permite comparar la media con su valor verdadero o bien las medias de dos poblaciones. Se basa en los límites de confianza "LC" para el promedio x de n mediciones repetidas (Ec. 2.1). A partir de dicha ecuación tenemos:

$\mu = x \pm t(s/\sqrt{n})$ (Ec. 2.1) → $x - \mu = \pm t\, s/\sqrt{n}$ (Ec. 2.2)

$s/\sqrt{n}$: error estándar "EE" o desviación estándar "DE" de la distribución muestral de medias. Como las medias son $\sqrt{n}$ veces más probables que los resultados aislados, la DE de las medias es $\sqrt{n}$ veces menor que la DE de resultados aislados, siendo **n** el número de determinaciones con las que se calcula la media.

t: "t de student" (tabla 2). Es un parámetro tabulado que depende de los grados de libertad de la muestra (n-1) "gl" y del intervalo de confianza que se quiera (generalmente 95%).

Tabla 2. Valores críticos de t (p= 0.05, entre paréntesis p=0.1).

gl	1	2	3	4	5	6	7	8	9
t	12.71	4.3	3.18	2.78	2.57	2.45	2.36	2.31	2.26
	(6.31)	(2.92)	(2.35)	(2.13)	(2.02)	(1.94)	(1.89)	(1.86)	(1.83)
gl	**10**	**12**	**14**	**16**	**18**	**20**	**30**	**50**	**α**
T	2.23	2.18	2.14	2.12	2.1	2.09	2.04	2.01	1.96
	(1.81)	(1.78)	(1.76)	(1.75)	(1.73)	(1.72)	(1.70)	(1.68)	(1.64)

Nota: *los valores críticos de t son estimados para una prueba de 2 colas. Para una prueba de una cola se toma el valor que corresponde a p=0.1, es decir, el doble del valor de p deseado (0.05).*

Si x - μ obtenida en la muestra a comparar es menor que la calculada para un cierto nivel de probabilidad, no se rechaza la hipótesis nula de que x y μ sean iguales; es decir, sus diferencias son debidas a errores aleatorios y no existe un error sistemático significativo.

Para comparar 2 medias experimentales el proceso es semejante. Se ha de tener en cuenta si los datos de las 2 muestras están apareados o no (figura 1):

Figura 1. Esquema del cálculo de t para datos apareados y no apareados

A) t student con datos apareados

X_{11}	X_{21}	D_1
X_{12}	X_{22}	D_2
...	...	...
X_{1i}	X_{2i}	D_i
		$\bar{D}_i = \sum D_i/i$

$S = \sqrt{\sum (D_i - \bar{D}_i)^2/(n-1)}$

$t = \bar{D}_i/EE \Rightarrow EE = s/\sqrt{n}$

Si $p \leq 0.05$ rechazar H_0

B) t student con datos no apareados

X_{11}	X_{21}
X_{12}	X_{22}
...	...
X_{1i}	X_{2i}
$\bar{X}1$	$\bar{X}2$
S_1	S_2

$EE_{(x_1-x_2)} = \sqrt{(\bar{s})^2/(1/n_1+1/n_2)}$

siendo $(\bar{s})^2 = (n_1-1)\,s_1^2+n_2-1)s_2^2)/(n_1+n_2-2)$

$t=(x_1-x_2)/EE$[t tiene (n_1+n_2-2)° de libertad]

Si $p \leq 0.05$ rechazar H_0

* Datos apareados: tienen la ventaja de permitir trabajar simplificando a una sola muestra (cuyos valores corresponden a la diferencia "D_i" entre cada par de datos apareados). Sustituimos x - μ (Ec. 2.2) por Di - 0 porque el valor real de las diferencias, suponiendo

que las dos muestras tienen la misma media, es 0. La DE se calcula con la muestra de diferencias.

* Datos no apareados: como no se puede simplificar a una sola muestra, se ha de introducir el concepto de desviación estándar ponderada "sp" (Ec. 2.3). En la **ecuación 2.2** se sustituye s por s_p y x - μ por $x_1 - x_2$ y el tamaño de muestra "n" se sustituye por N ponderado "$(N_1 + N_2)/ N_1N_2$".

$S_p=\sqrt{[S(x_1 - x_1)2+S(x_2 - x_2)2+ \ldots]/(n_1+ n_2+ \ldots - N_s)}$ (Ec. 2.3)

n_1, n_2, ...: el tamaño de las muestras.
N_s: número de muestras.
$(n_1+n_2+\ldots-N_s)$: número de grados de libertad.

Ejemplo 1: se analizaron dos sueros control (A y B) para la determinación de la glucemia. Se realizó sobre cada uno de ellos 5 determinaciones (tabla 3a) y se quiere determinar si estos dos sueros control son diferentes en relación al nivel de glucosa.

Tabla 3a

	glucosa mg/dl				
Suero A	75	80	82	79	80
Suero B	81	90	85	83	87

Aunque el número de determinaciones es reducido podemos suponer que si realizáramos más determinaciones la distribución sería normal (teorema central del límite). Realizamos la prueba t de student de datos no apareados, ya que aunque las dos muestras tienen el mismo tamaño provienen del análisis de dos sueros supuestamente diferentes (tabla 3b):

Tabla 3b

	$\bar{X}$	$\bar{X}_B - \bar{X}_A$	t (8 g l*)	s ponderada	N ponderado	t $s_P N_P$
Muestra A	79.2					
Muestra B	85.2	6	2.31**	9.45	2.5	13.8

* grados de libertad (ver **Ec. 2.3**). ** valor de t correspondiente (**tabla 2**).

Como la diferencia de las medias es menor que 13.8, puede decirse que las dos muestras son significativamente iguales ($p < 0.05$).

2. Pruebas de una y dos colas

En las "**pruebas bilaterales o de dos colas**" se comparan dos muestras para saber si difieren entre sí, sin preguntarse cuál de ellas tiene mayor estadístico (Ej. media). Si se pretende evaluar qué muestra tiene el estadístico mayor (sesgo positivo) se realiza una "**prueba unilateral o de una cola**". Para un tamaño "n" determinado y un nivel de probabilidad concreto, los valores críticos de ambas pruebas difieren. Suponiendo una población simétrica, la probabilidad de la prueba unilateral es la mitad de la probabilidad de la prueba bilateral. Por ello, para encontrar el valor adecuado para una significación del 95% ($p=0.05$) en una prueba de una cola, se busca en la columna de $p=0.1$ de la tabla de pruebas bilaterales.

Tabla 4. Valores críticos de F para una prueba de dos colas (p= 0.05)

	1	2	3	4	5	6	7	8	9	10	12	15	20
1	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3	968.6	976.7	984.9	993.1
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43	39.45
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25	14.17
4	12.22	10.65	9.979	9.605	9.364	9.197	9.074	8.980	8.905	8.844	8.751	8.657	8.560
5	10.01	8.434	7.764	7.388	7.146	6.978	6.853	6.757	6.681	6.619	6.525	6.428	6.329
6	8.813	7.260	6.599	6.227	5.988	5.820	5.695	5.600	5.523	5.461	5.366	5.269	5.168
7	8.073	6.542	5.890	5.523	5.285	5.119	4.995	4.899	4.823	4.761	4.666	4.568	4.467
8	7.571	6.059	5.416	5.053	4.817	4.652	4.529	4.433	4.357	4.295	4.200	4.101	3.999
9	7.209	5.715	5.078	4.718	4.484	4.320	4197	4.102	4.026	3.964	3.868	3.769	3.667
10	6.937	5.456	4.826	4.468	4.236	4.072	3.950	3.855	3.779	3.717	3.621	3.522	3.419
11	6.724	5.256	4.630	4.275	4.044	3.861	3.759	3.664	3.588	3.526	3.430	3.330	3.226
12	6.554	5.096	4.474	4.121	3.891	3.728	3.607	3.512	3.436	3.374	3.277	3.177	3.073
13	6.414	4.965	4.347	3.996	3.767	3.604	3.473	3.388	3.312	3.350	3.153	3.053	2.948
14	6.298	4.857	4.242	3.892	3.663	3.501	3.380	3285	2.209	3.147	3.050	2.949	2.844
15	6.200	4.775	4.153	3.804	3.576	3.415	3.293	3.199	3.123	3.060	2.963	2.862	2.756
16	6.115	4.687	4.077	3.729	3.502	3.341	3.219	3.125	3.049	2.986	2.889	2.788	2.681
17	6.042	4.619	4.011	3.665	3.438	3.277	3.156	3.061	2.985	2.922	2.825	2.723	2.616
18	5.978	4.560	3.954	3.608	3.382	3.221	3.100	3.005	2.929	2.866	2.769	2.667	2.559
19	5.922	4.508	3.903	3.559	3.333	3.172	3.051	2.956	2.880	2.817	2.720	2.617	2.509
20	5.871	4.461	3.859	3.515	3.289	3.128	3.007	2.913	2.837	2.774	2.676	2.573	2.469

Tabla tomada de HR Neave. Elementary Statistics Tables. George Allen & Unwin Ltd.

La decisión de utilizar una prueba de una o dos colas, depende del grado de conocimiento del sesgo positivo o negativo que se tenga a priori. Nunca debe decidirse después de realizar el experimento, pues la decisión está influenciada por los resultados.

Ejemplo 2: antes de analizar la vitamina A por cromatografía se realiza una extracción líquido-líquido. Si se quiere evaluar la recuperación de la vitamina A en el proceso de extracción, el sesgo será forzosamente negativo, pues nunca puede extraerse más de lo que hay. En este caso se aplicará una prueba de una cola.

3. Comparación de variancias por contraste de Fisher

Para comparar las variancias de dos muestras (S_1^2 y S_2^2) se plantea la hipótesis nula y la alternativa.

Hipótesis nula "H_0" → $S_1^2 = S_2^2$

Hipótesis alternativa "H_1" → $S_1^2 \neq S_2^2$

Dos muestras tienen variancias diferentes cuando la razón de sus variancias "F", colocando en el numerador la variancia mayor para que siempre sea mayor de uno, excede el valor crítico F tabulado. El valor crítico de F se escoge de la tabla 5 según los tamaños de muestra (n1, n2) y el nivel de significación deseado (generalmente 95%). Hay que tener en cuenta que aunque sólo se exponga la tabla de dos colas, se debe aplicar la tabla de valores F para pruebas de una o dos colas según el caso.

$F=S1^2/S2^2$ (Ec. 2.4) → $F \leq F_{tabulado}$ (g l=n_1-1; n_2-1 y 95% significación) no se puede rechazar H_0 con una probabilidad de error menor del 5% ($p < 0.05$).

Tabla 5

Muestra	Pureza en %				Ti	Ti²
A	0.3	0.2	0.2	0.6	1.3	1.69
B	-0.6	-0.3	0.5	0.5	0.1	0.01
C	-0.3	0.0	0.0	-0.2	-0.5	0.25
D	0.6	0.7	0.0	-0.2	1.1	1.21
E	-0.3	0.0	0.0	-0.2	-0.5	0.25
	n=4 y k=5 por lo que N=20				ΣT=1.5	ΣT²=3.41
	$\Sigma_i \Sigma_j X_{ij}^2$= 2.84				$\overline{X}_T$=1.5/20=0.075	

Ti: sumatorio de cada muestra; Ti²: Ti elevado al cuadrado; ΣT: sumatorio de los Ti; ΣT²: ΣT elevado al cuadrado; k: número de muestras; n: número de determinaciones en cada muestra; $\Sigma \Sigma_j X_{ij}^2$: sumatorio de todos los datos elevados al cuadrado; $\overline{X}_T$: media global de todos los datos, que coincide con ΣT/N.

4. Análisis de variancia (Anova) de un factor

Está basada en la prueba de Cochran para estudiar la homogeneidad de las variancias de poblaciones origen de varias muestras. Si las muestras proceden de la misma población tendrán variancias semejantes, por lo que un análisis de variancias permitirá comparar sus medias. Es una prueba bilateral de comparación por *contraste de Fisher* de k medias, en la que se compara las variancias de las k medias (s^2y), con la variancia que debiera obtenerse si las k muestras procedieran de poblaciones con igual media [Sp^2/n → siendo Sp la desviación estándar ponderada calculada entre las distintas muestras (Ec. 2.3)].

$F = s^2y / s^2/n$ = CMMedias/CMResiduales (Ec. 2.5)

El ANOVA de un factor se aplica para análisis donde, además del error aleatorio inevitable en las mediciones, hay un factor controlado (Ej. temperatura) o aleatorio (Ej. toma de muestra al azar). Hay que indicar que el ANOVA es una generalización de la "prueba t de student" (t y F son equivalentes al comparar sólo 2 grupos).

Planteamos la hipótesis nula y alternativa:

$Ho: M_1=M_2=M_3= \ldots =M_k$; $H_1: M_1 \neq M_2 \neq M_3 \neq \ldots \neq M_k$ (siendo M_k, la media k-ésima).

Se define el tamaño de la población total como:

$N= n_1+n_2+n_3+ \ldots +n_k$.

Para una mejor comprensión pondremos un ejemplo para realizar el cálculo sencillo "con papel y lápiz":

Ejemplo 3: se dan los resultados de cuatro análisis de purezas en % realizados sobre 5 muestras (tabla 5). Los datos se manejan respecto a una pureza basal de 98.5% para manejar números pequeños y simplificar los cálculos. Se pretende comparar si las medias son significativamente iguales:

Tabla 6: Fuentes de variación

	Sumatoria de cuadrados "SC"	grados libertad (gl)	CM*
Intermuestra	$\Sigma T^2/n - \Sigma T/N = 3.41/4 - 1.5/20 = 0.778$	k-1=4	$CM_{Medias}=0.195$
Intramuestra**	2.77-0.778= 1.99	19-4=15	$CM_{Residual}=0.133$
Total	$\Sigma i\, \Sigma j X_{ij}^2 - \Sigma T/N = 2.84 - 1.5/20 = 2.77$	N-1=19	CMTotal=0.146

**CM (cuadrado medio)= SC/g l. ** los valores intramuestra se calculan por diferencia entre el total y el intermuestra, ya que la variancia total es la suma de la variancia inter e intramuestra.*

F= CMMedias/CMResiduales= 0.195/0.133= 1.47. Se compara con el F crítico tabulado para una cola porque es de suponer que la variación intermuestra, si las muestras proceden de distinta población, sea mayor que la variación intramuestra. Si Fcalculado > Ftabulado →p ≤ 0.05 se puede rechazar H0 y admitir H1, osea que las muestras provienen de distinta población.

5. Regresión y correlación

Ambas técnicas permiten analizar la relación entre dos variables cuantitativas. Es muy común la confusión entre regresión y correlación a pesar de que son completamente distintas.

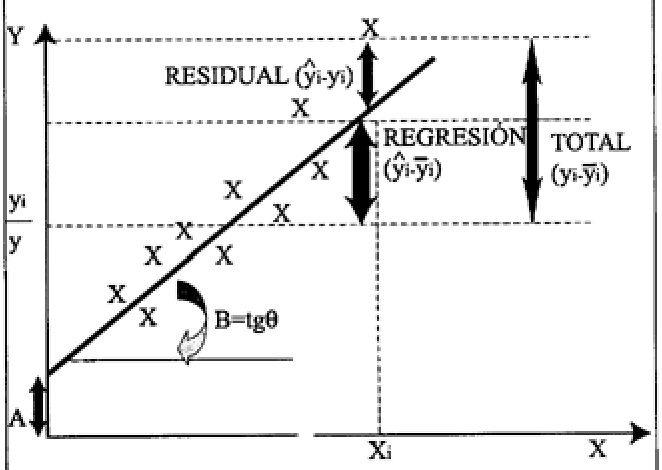
Tabla 7: Comparación entre regresión y correlacción

REGRESIÓN	CORRELACIÓN
Se evalúa la dependencia de una variable "efecto" con respecto a varias variables independientes "causas"	Se evalúa la concordancia entre dos variables independientes (ambas juegan un papel simétrico)
La variable independiente puede ser incluso controlada	Muy exigente. Las variables y su distribución conjunta han de seguir una ley normal
La intensidad de la relación entre la variable dependiente y las independientes se evalúa con r^2 (Ec. 2.11)	La concordancia entre las dos variables se evalúa con r^2

5A) Correlación por ajuste de una recta con el criterio de mínimos cuadrados

A partir de la matriz de datos con n pares de valores (x_i, y_i) se pueden representar los pares de valores por puntos en un diagrama de ejes cartesianos (figura 2). El eje de abscisas representa la variable X y el eje de ordenadas la variable Y. Los pares de valores (x_i, y_i) se representan por los puntos de intersección de las rectas que perpendicularmente a los ejes X e Y, pasan por los puntos x_ie y_i de dichos ejes. Para buscar la recta que mejor se ajusta al conjunto de puntos representados y evaluar el grado de ajuste a dicha recta, se determinan los coeficientes A y B de la recta de ajuste (Ec. 2.6) que hagan mínima la suma de los residuales "ei" al cuadrado (Ec. 2.7).

Figura 2. Correlación por ajuste de mínimos cuadrados



X: pares de valores (Xi, Yi)
Para el conjunto de los sujetos se demuestra que la variación total se descompone como SC→$\Sigma(y_i-\bar{y})^2=\Sigma(\hat{y}_i-\bar{y})^2+\Sigma(y_i-\hat{y}_i)^2$

y = A + Bx (Ec. 2.6)

Los residuales "e_i" representan la distancia, perpendicular a la recta, de los puntos a la recta.

$\Sigma e_i^2 = \Sigma 1^n [y_i-(A+Bx_i)]^2$ (Ec. 2.7)

Como las derivadas de una función al alcanzar los valores mínimos deben ser cero, los valores A y B se calculan (Ec. 2.8) haciendo cero las derivadas parciales del sumatorio de residuales (Ec. 2.7) con respecto a A y B, ordenada en el origen y pendiente[1] de la recta de correlación (figura 2):

$A= y-B x$, $B = [\Sigma x_i y_i - \Sigma x_i \Sigma y_i/n]/\Sigma x_i^2-(\Sigma x_i)^2/n$ (Ec. 2.8)

[1]Recordar que la pendiente es el valor de la tangente del ángulo que forma la recta con el eje de abscisas.

Nota: es importante saber que la recta de correlación pasa por el punto (x, y) siendo x, y las medias de los datos x_i e y_i. Aplicando la ecuación punto-pendiente de una recta podemos obtener la ecuación de la recta de mínimos cuadrados:

$y= y + B(x- x)$

5.a.1. Descomposición de la suma de cuadrados

Como se observa en la figura 2 para un sujeto "i" cualquiera, se verifica que:

$(y_i - y)\text{TOTAL}= (\hat{}y_i - y)_{\text{REGRESIÓN}}+(y_i-\hat{}y_i)_{\text{RESIDUAL}}$ (Ec. 2.9)

y_i: valor aislado de y del sujeto i.
y: media del conjunto de valores y_i. Se supone el valor verdadero.
$\hat{}y_i$: valor de y del sujeto i, calculado por la ecuación de la recta de mínimos cuadrados.
Total: variación total, distancia desde y_i al valor medio de y.
Regresión: variación explicada por la recta de regresión.
Residual: variación no explicada.

Y para el conjunto de los sujetos se demuestra que la variación total se descompone como:

$\Sigma(y_i - y)^2= \Sigma(\hat{}y_i - y)^2+\Sigma(y_i-\hat{}y_i)^2$ (Ec. 2.10)

5.a.2 Valoración del ajuste de la recta

En el laboratorio es muy común el ajuste de datos experimentales a una recta, por ejemplo para construir la recta de calibración de una determinada técnica. El ajuste perfecto se da cuando la variancia residual es nula y todos los puntos están sobre la recta de regresión. El ajuste nulo se da cuando la variación explicada por la regresión es cero.

Se define el coeficiente de correlación "r^2" como la relación entre la variación explicada y la variación total.

r^2 = SCRegresión/SCTotal = B^2 (SCx/SCy) (Ec. 2.11)

El coeficiente toma valor de 1 cuando el ajuste es perfecto y 0 cuando es nulo. Es importante indicar que un ajuste nulo no quiere decir ausencia de relación, ya que sólo indica ausencia de relación lineal (puede existir una relación de tipo parabólico, exponencial...). Hay que indicar que es preferible hablar de r^2 que de r porque r varía desde 1 hasta +1, correspondiendo con una pendiente de la recta negativa o positiva, sin embargo r^2 siempre toma valores positivos.

En algunas circunstancias se obtienen valores bajos de "r". Para evaluar si el coeficiente es significativo se debe considerar el número de pares de valores usados en su cálculo. El método más simple es calcular un valor de t usando la ecuación 2.12.

$t= [|r|\sqrt{(n-2)}]/\sqrt{(1-r^2)}$ (Ec. 2.12)

El valor de t calculado se compara con el tabulado para el nivel de significación deseado, usando una prueba de t de dos colas con (n-2) gl. Si t calculado es mayor que el tabulado, H0 se rechaza (hay correlación significativa). H0: no existe correlación entre x e y.

Hay que indicar que los cálculos anteriores permiten obtener la "recta de regresión de y sobre x" es decir, la recta que evalúa cómo varía y cuando x se ajusta a los valores elegidos. La "recta de regresión de x sobre y", la que supone que todos los errores ocurren en x, no coincide con la anterior salvo cuando r =1.

5.a.3. Errores de la pendiente y ordenada en el origen de la recta de regresión

Son errores importantes por lo que es muy útil evaluarlos. Para ello se define el estadístico sy/x a partir de los residuos de la tabla 8.

Tabla 8: Fuentes de variación en un análisis de correlación (y= A + Bx)*

fuente	sumas de cuadrados	grados de libertad	variancias ´O CM
Regresión	$\Sigma(\hat{y}_i - \bar{y})^2 = B^2\Sigma(x_i - \bar{x})^2 = B^2SC_x$*	1**	SC/gl
Residual	$\Sigma(y_i - \hat{y}_i)^2 = SC_y - B^2SC_x$	n-2	SC/gl
total	$\Sigma(y_i - \bar{y})^2 = SC_y$	n-1	SC/gl

* A es una constante y no participa en la suma de cuadrados.
** La variación explicada por la regresión tiene 1 grado de libertad porque los valores $\hat{y}_i$ están definidos por la ecuación de una recta. Las rectas son monodimensionales y sólo tienen un grado de libertad.

$s_{y/x}= \sqrt{[\Sigma(y_i\text{-}\hat{}y_i)^2 / (n-2)]}$ (Ec.2.13)

Las DE de los parámetros A y B de la recta de mínimos cuadrados son:

$SB=s_{y/x}/\sqrt{S(x_i-x)^2}$ (Ec.2.14)
$SA=s_{y/x}\sqrt{[\Sigma x_i^2 / n\Sigma(x_i-x)^2]}$ (Ec.2.15)

SB y SA se pueden utilizar para estimar los límites de confianza de la pendiente y ordenada en el origen ($B \pm ts_B/\sqrt{n}$; $A \pm ts_A/ \sqrt{n}$), donde t se obtiene para (n-2) gl.

Una vez ajustada la recta de calibración, la señal de una muestra cualquiera y0 se extrapola para calcular la concentración x0. El cálculo del error en la estimación de x0 es complejo, pero podemos emplear la siguiente fórmula aproximada:

$S_{x0}=s_{y/x}/B\sqrt{[1/m+1/n+[(y_0-y)^2 /B^2\Sigma(x_i-x)2]]}$ (Ec.2.16)

m: nº de lecturas realizadas para calcular y0.
n: nº de puntos de la recta de calibración.

A partir de s se puede calcular los límites de confianza de x0. Hay que indicar que S_{x0} es menor a medida que y_0 se aproxima a y, de ahí la razón de elegir unas concentraciones para la recta de calibración que hagan que las muestras problema estén próximas a y. El aumento excesivo de m y n genera mucho trabajo adicional y no disminuye apreciablemente S_{x0} por lo que suele ser suficiente m=4 y n=6.

5B) Regresión lineal

La regresión lineal simple consta de dos etapas bien diferenciadas. En la primera etapa, meramente descriptiva, se utiliza el ajuste por mínimos cuadrados para hallar la ecuación de la recta que se ajuste mejor a los datos "recta de regresión". La segunda etapa, inferencial, estima la dependencia de la variable dependiente respecto a la independiente.

5.b.1. Vertiente descriptiva o correlación

Etapas:

1. Cálculo de la recta de regresión por el ajuste de mínimos cuadrados (ver sección superior).
2. Descomposición de la suma de cuadrados.
3. Valoración de la bondad del ajuste por el coeficiente "r^2".

Hay dos índices estadísticos muy importantes: la covariancia y el coeficiente de correlación de Pearson.

* Covariancia "S_{xy}": promedio de los productos de las desviaciones de las dos variables $P_{xy}=(x_i-x)(y_i- y)$. Puesto que las desviaciones tienen signo, la covariancia será positiva cuando la nube de puntos en la representación y *vs x* tenga una forma lineal ascendente y negativa en el caso contrario. La unidad de medida de la covariancia es el producto de las unidades de medida de x e y.

$S_{xy}= \Sigma(x_i-x)(y_i-y)/n-1=SP_{xy}/n-1$ (Ec. 2.17)

S_{xy}: covarianca.

SP_{xy}: sumatorio de los productos cruzados de las desviaciones de las dos variables → $\Sigma(x_i-x)(y_i-y)$

* <u>Coeficiente de correlación de Pearson</u> "r_{xy}": se obtiene estandarizando la covariancia para eliminar la influencia de las unidades de medida, por lo que es una magnitud adimesional. Se define como:

$r_{xy}= S_{xy/SxSy}$ (Ec. 2.18)

r_{xy} toma valores entre 1 y +1. 1: relación lineal perfecta negativa; +1: relación lineal perfecta positiva y 0: relación nula. **Nota importante**: el cuadrado del coeficiente de Pearson coincide con el cociente entre la variación explicada por la recta de regresión y la variación total (coeficiente de correlación, ecuación 2.11).

5.b.2. Vertiente inferencial o regresión

Como se ha dicho, la regresión lineal pretende deducir la relación lineal entre una variable dependiente, y otras independientes que la condicionan. El modelo exige que los datos de la población cumplan los supuestos de linealidad, homocedasticidad, independencia entre las observaciones y normalidad de las distribuciones condicionadas de la variable "y". Estos 4 supuestos se pueden resumir en que los términos de error o residuales para cada una de las distribuciones condicionales deben ser variables aleatorias independientes, distribuidas según la Ley Normal.

* <u>Linealidad</u>: Las medias de "y" están sobre la recta: $\mu_i= \alpha + ßx_i$
* <u>Homocedasticidad</u>: las variancias de la variable dependiente no se modifican con la variable independiente:

$\sigma^2 y(x1)= \sigma^2 y(x2)... = \sigma^2$

* Ausencia de autocorrelación: un valor y_i no influye sobre otro valor y_j. Al igual que en el caso de la correlación, se estima el modelo de regresión mediante una recta:

$\mu_i = \alpha + \beta x_i + \varepsilon_i$ (Ec. 2.19)

ε_i: error residual.

Estimación de los parámetros del modelo de regresión: se demuestra que los coeficientes A y B del ajuste de mínimos cuadrados son los mejores estimadores lineales de los parámetros a y b de la población. Así, α=A (ordenada en el origen de la recta de correlación, ecuación 2.7) y β=B (pendiente de la recta, ecuación 2.8).

En el trabajo de laboratorio muchas veces se necesitan técnicas que detecten errores sistemáticos en los resultados obtenidos con un método respecto a otro de referencia. Una posible relación entre dos métodos es una función lineal:

$Y = \alpha + \beta x + \varepsilon$ (Ec. 2.20)

Donde α representa el error sistemático constante, β el error sistemático proporcional y ε el error no sistemático (aleatorio) de los métodos.

5C) Modelo de regresión ortogonal

En el apartado anterior vimos que α y β se estimaban por el ajuste de mínimos cuadrados suponiendo que X estaba desprovista de error. En este caso se debe considerar también el error de X, por lo que α y β se estiman de forma que la suma de cuadrados de las distancias perpendiculares desde cada punto hasta la recta de regresión sea mínima: "modelo de regresión ortogonal".

A_0= y B_0 x (Ec. 2.21)

$B_0 = (SC_y SC_x)/2SP_{xy} + \sqrt{\{[(SC_y SC_x)/2SP_{xy}]^2 + 1\}}$ (Ec. 2.22)

II. B) Pruebas no paramétricas

Las pruebas paramétricas requieren que los datos sigan la distribución normal. Muchas veces nos interesa aplicar métodos que no requieran dicha hipótesis:

* En muchos experimentos se utilizan muestras pequeñas de datos no pudiéndose aplicar el teorema central del límite.

* Hay variables que, aunque se recojan numerosos datos, claramente no siguen una distribución normal. Así por ejemplo las concentraciones de anticuerpos en suero de una muestra de pacientes sigue una distribución log-normal.

* Se emplean cuando interesa aplicar métodos más sencillos para una valoración inicial. Esta razón va perdiendo peso pues día a día son más numerosos los laboratorios que disponen de programas estadísticos de ordenador que simplifican enormemente la aplicación de tests paramétricos.

1. PRUEBAS BASADAS EN TABLAS DE CONTINGENCIA. ESTUDIOS RETROSPECTIVOS CASO-CONTROL

El mejor diseño de un estudio es el diseño experimental, diseño aleatoriamente controlado de tipo prospectivo, en el que los individuos se asignan de forma aleatoria, bien al grupo de tratamiento o bien a un grupo control. Pero si el tratamiento conlleva un cierto peligro no se puede aplicar a humanos por razones éticas. En estos casos se realiza un *estudio prospectivo de grupos* "*cohortes*", se confrontan individuos expuestos con individuos análogos en todo lo que se pueda imaginar salvo en la exposición. Cuando la exposición se valora difícilmente o cuando la frecuencia de la enfermedad es muy pequeña resulta más práctico realizar un estudio retrospectivo por medio de ***tablas de contingencia***, en el que se considera un grupo indiscriminado de individuos con enfermedad (casos) y sin ella (controles), y se observa la exposición a la que han estado sometidos.

1.a. Prueba de chi-cuadrado de Karl-Pearson

Consiste en un contraste de frecuencias o contajes observados (O) con los teóricos o esperados (E) representados en la tabla de contingencia de dimensiones 2x2 de la figura 3. Así, se define el parámetro chi cuadrado de la siguente forma:

$\chi^2 = \Sigma(O-E)^2/E$ (Ec. 2.23.a)

Figura 3. Tabla 2x2. Prueba $\aleph^2$
Se tabula, como ejemplo, los niños nacidos en un hospital durante un año en función de su peso y de las semanas de gestación

		Prematuros 0	Prematuros 1	
Bajo peso	0	**352** 342,7(E)	**35** 44.3 (E)	$\sum_i$=387 O
	0	**12** 21.3 (E)	**12** 2.7 (E)	$\sum_i$=24 O
		$\sum_c$=364 **O**	$\sum_c$=47 **O**	$\sum_t$=411

Prematuros: 1→semanas de gestación < 36;0→≥36.
Bajo peso: 1→< 2500 g; 0→ ≥2500 g.
O: observados; E: esperados

Las frecuencias se ajustan a la distribución de Poisson, que tiene una propiedad poco habitual, la variancia coincide con la media. χ^2, por lo tanto, es una razón de tipo señal-ruido entre el cuadrado de la diferencia entre las frecuencias observada y esperada respecto a la media esperada.

Nota importante: la prueba sólo se puede aplicar cuando el N° total de observaciones es mayor de 50 y las frecuencias individuales esperadas no son menores de 5.

Se compara el χ^2 calculado con el valor crítico tabulado (figura 3) en función del nivel de significación y del número de grados de libertad [gl de cada cada variable= N° de clases 1, siendo la clase cada una de las posibilidades en la variable; gl para la prueba es la multiplicación de los gl de cada variable. En el ejemplo de la figura 3 el número de gl=1]. Si el chi cuadrado calculado supera al tabulado, se puede rechazar H_0 y admitir H_1 (es decir, es poco verosímil ($p< 0.05$) que los resultados obtenidos sean debidos al azar).

Tabla 9

Valores críticos de $\aleph^2$ (p = 0.05)

número g l	Valor crítico	número g l	Valor crítico
1	3.84	6	12.59
2	5.99	7	14.07
3	7.81	8	15.51
4	9.49	9	16.92
5	11.07	10	18.31

Tabla tomada de HR Neave. Elementary Statistics Tables. George Allen & Unwin Ltd.

Nota: cuando hay sólo dos clases (un grado de libertad) o la frecuencia observada en una o varias clases es menor del 5%, se puede aplicar la "corrección de Yates" que implica sustituir cada O-E por |O-E |-0.5.

Ejemplo 4. Se recogen los datos de peso y semanas de gestación de los niños nacidos en un hospital durante un año. Se tabulan en una tabla dos por dos, clasificándolos en 4 clases en función del bajo peso (0: peso >= 2500 g; 1: peso < 2500 g) y prematuridad (0 >= 36 semanas de gestación; 1: < 36). A partir de los datos recogidos "observados" calcular los datos esperados e indicar si la diferencia con respecto a los observados es debida al azar.

Primero se comprueba que hay más de 50 observaciones y que las frecuencias observadas son mayores del 5%. Los valores esperados se calculan por medio de una simple regla de tres. Para el grupo de niños bajo peso "0", prematuros "0", el valor esperado se calcularía de la siguiente forma:

Si en 411 recién nacidos (totales) hay 364 prematuros, en 387 recién nacidos con bajo peso habrá X recién nacidos prematuros. Aplicando una regla de tres se obtiene el valor esperado:

X= 364*387/411=342.7.

De igual forma se calcula el resto de valores esperados, que se indican en la figura 3.

Cálculo de $\aleph^2$ (tabla 10.a):

Tabla 10.a

Fecuencia O	Frecuencia E	O-E	$(O-E)^2/E$
352	342.7	9.3	0.252
35	44.3	-9.3	1.952
12	21.3	-9.3	4.060
12	2.7	9.3	32.033
			$\aleph^2$= 38,299

Para 3 grados de libertad (4 clases 1) $\aleph^2$ crítico tabulado es 7.81. Como el valor de $\aleph^2$ calculado es mayor que el valor crítico, se puede rechazar H0 con una probabilidad de error menor del 5%, por lo que significativamente hay diferencias que no son debidas al azar.

1.b Prueba de chi-cuadrado de McNemar "$\aleph^{2}$M"

En el contraste de observaciones emparejadas se calcula $\aleph^2$ de McNemar (ecuación 2.23.b) construyendo una tabla 2x2 diferente, en la que las celdas muestran las cuatro posibilidades de los pares caso-control: ambos estan expuestos, sólo los casos están expuestos, sólo los controles están expuestos, ninguno está expuesto. Si no hubiera ninguna relación entre la enfermedad y la exposición sería de esperar que hubiera igual frecuencia cuando los casos están expuestos y los controles no, como en la situación contraria, por lo que el valor esperado en este caso es la media de los valores que no están en la diagonal principal. En el ejemplo 5 (tabla 10.c) el valor esperado sería E= b+c/2= 14.

$\aleph^2_M$ se calcula a partir de la ecuación 2.23.a considerando dicho valor esperado y aplicando la corrección de Yates

$\aleph^2_M$= [(b-c)-1]^2/(b+c) (Ec. 2.23.b)

Si $\aleph^2_M$ es mayor que el tabulado

Ejemplo 5: se realizó un estudio caso-control para estudiar el efecto potencial del consumo regular de tabaco sobre el carcinoma vesical. Para ello se contrastaron 50 pacientes con dicho cáncer frente a 50 individuos control. En los casos había 35 que consumían más de 10 cigarrillos al día, así como 12 de los asociados a ellos en el grupo control. En los casos que no fumaban se asociaban 5 controles que fumaban. Calcular la $\aleph^2$ de McNemar.

La tabla 10.b sería la tabla clásica de contingencia, pero para el análisis de McNemar construimos la tabla 10.c.

Tabla 10.b

Consumo de Tabaco

		Sí	No	Total
Carcinoma vesical	Sí	35	15	50
	No	17	33	50
	Total	**52**	**48**	**100**

Tabla 10.c

Control fumador

		Sí	No	Total
Caso fumador	Sí	a=12	b=23	35
	No	c=5	d=10	15
	Total	**17**	**33**	**50**

χ^2_M= [(b-c)-1]$_2$/(b+c) = 17^2/28= 10.32, valor mucho mayor que el tabulado para 1 gl (tabla 9) se rechaza H0 con una probabilidad de error menor del 5%, por lo que significativamente hay diferencias, no debidas al azar.

1.c Prueba de chi-cuadrado de Mantel-Haenszel "χ^2_{M-H}"

χ^2 de Karl-Pearson equivale desde el punto de vista paramétrico a un t-test para dos categorías o un ANOVA para más de 2 categorías. χ^2 de McNemar equivale a un t-test de datos apareados. En el caso de un ANOVA de dos factores, un estadístico no paramétrico equivalente es χ^2 de Mantel-Haenszel.

1.d Test exacto de Fisher

Cuando las frecuencias observadas son menores del 5%, deben de aplicarse otros test alternativos a la prueba de chi cuadrado. Una posibilidad es valorar teóricamente la probabilidad de observar un valor tan extremo, o incluso más extremo que el observado realmente. A partir de la Ley binomial (Ec. 2.26 y 2.27), se puede demostrar que la probabilidad de que aparezca una configuración determinada en una tabla 2x2 (tabla 11) viene dada por la ecuación 2.24.

Tabla 11: Tabla 2x2

a	b
c	d

a, b, c, d: número de casos en cada clase.

$p(k)= (a+b)!\,(c+d)!\,(a+c)!(b+d)!/N!a!b!c!d!$ (Ec. 2.24)

K: es el menor número entre a, b, c, d.

$x!$: factorial de x: $x!=x(x-1)(x-2)(x-3)\ \dots\ 1$

Se indica p(k) porque la probabilidad usada en el estadístico es toda la probabilidad de la cola, es decir desde k hasta el final de la distribución $p(K\rightarrow 0)$ (ecuación 2.25).

$p(K\rightarrow 0)= p(k)+p(k-1)+p(k-2)+p(k-3)+ \dots + p(0)$ (Ec. 2.25)

Nota: para K- i (i=1, 2, ... k) se sustituye el valor de k en la tabla 2x2 por k-i y se suma +i en las celdas adyacentes para que las sumas parciales y totales de las celdas sean las mismas (ver ejemplo 6).

Si $p(k\rightarrow 0)$ supera 0.05 no se puede rechazar la hipótesis nula, que las diferencias son debidas al azar.

Ejemplo 6: Se han tabulado, como indica la figura 4 los niños nacidos en un hospital durante un año en función de las semanas de gestación y aparición del síndrome de dificultad respiratoria "SDR". ¿Hay diferencias significativas entre los niños prematuros y los no prematuros en cuanto al SDR?.

Figura 4. Tabla 2x2. Test exacto de Fisher

Se tabula, como ejemplo, los niños nacidos en un hospital durante un año en función de las semanas de gestación y aparición del síndrome de dificultad respiratoria "SDR"

		Prematuros 0	Prematuros 1	
SDR	0	**363** a)	**41** b)	$\sum_f$=404 O
SDR	1	1 c)	6 d)	$\sum_f$=7 O
		$\sum_c$=364 O	$\sum_c$=47 O	$\sum_T$=411

Prematuros: 1→semanas de gestación < 36;0→≥36.
SDR: 1→ positivo; 0→ negativo.
O: observados; E: esperados

De las cuatro clases el menor número de casos se da en recién nacidos no prematuros con SDR (prematuros=0, SDR=1), en la que sólo hay un caso (como en una clase hay menos de 5 % de casos no puede aplicarse la prueba de χ^2). Se calcula (Ec. 2.24) la probabilidad de tener de forma aleatoria 1 sólo caso o ninguno en esa clase (la probabilidad de la cola):

p(1→0)= p(1) + p(0)= (404! 7! 364! 47!) / (411! 363! 41! 1! 6!) + (404! 7! 364! 47!) / (411! 364! 41! 0! 7!)= $1.05*10^{-5}$

Como la probabilidad de que las diferencias sean debidas al azar es mucho menor de 0.05, se puede afirmar que sí hay diferencias significativas.

2. PRUEBA DE LOS SIGNOS

A cada valor experimental se le resta la mediana postulada y se considera el signo de cada resultado, ignorando los resultados iguales a la mediana. Para contrastar si una preponderancia de signos, en uno de los dos sentidos, es significativa, utilizamos la Ley binomial. Esta ley establece que la probabilidad de que aparezcan r signos menos, entre n signos, viene dada por:

$P(r)={}^nC_r p^r q^{(n-r)}$ (Ec. 2.26)

${}^nC_r$: nº de combinaciones de r elementos de un total de n elementos.

${}^nC_r$= n!/[r!(n-r)!] (Ec. 2.27)

p: probabilidad de que aparezca un signo menos en un resultado.

q: probabilidad de que en un resultado aparezca un signo más:

$q^{(n-r)}=1-p^{r}$.

Si P(r) calculado es mayor que 0.05, no se puede rechazar H0, es decir, que los datos proceden de la mediana postulada.

Ejemplo 7: un medicamento requiere que tenga una mediana del 5% en porcentaje del componente A. Analizando varios lotes del mismo se encontraron las siguientes concentraciones: 4.9, 4.8, 5.6, 6.1, 5.8, 6.2. ¿Cumple el medicamento la normativa con respecto al componente A?.

1. Restando la mediana a cada resultado: -0.1, -0.2, 0.6, 1.1, 0.8, 1.2. Se observa que predominan las diferencias positivas. Para contrastar si el déficit de signos negativos es significativo utilizamos la ecuación 2.26. El número de combinaciones de 2 elementos negativos en un total de 6 elementos será, según la ecuación 2.27, ${}^{6}C2$ = 6!/2!4!=15. Por la definición de mediana es evidente que la mitad de valores están debajo de ella y la otra mitad encima, lo que indica que la probabilidad de tener un signo más, p, o menos, q, en uno de los resultados es 1/2. Aplicando la ecuación 2.26 la probabilidad de que aparezcan 2 o menos signos negativos entre 6 signos es:

. 2. Para conocer si los datos difieren significativamente de la mediana realizamos una prueba de dos colas, en la que calculamos la probabilidad de obtener 2 o menos signos idénticos (ambos positivos o negativos) cuando se toman al azar 6 resultados que será el doble de P(≤2); 0.6875.

3. Puesto que el valor experimental es mayor de 0.05, no se puede rechazar H_0, que los datos proceden de una población con mediana 5, con una probabilidad de error menor del 5%.

La prueba de los signos puede utilizarse como una alternativa no paramétrica a la prueba t por parejas, y también para indicar si existe una tendencia en los valores con el tiempo (restando a valores iniciales, valores posteriores).

3. PRUEBA DE RACHAS DE WALD-WOLFOWITZ

En otros casos nos interesa, además de evaluar el número de signos (+) ó (-), si éstos aparecen en una secuencia aleatoria o a rachas. Racha se puede definir como cada serie de valores con el mismo signo y tiene un significado semejante al significado popular de rachas, de buena o mala suerte. Una distribución aleatoria de signos debería tener un número de rachas determinado, porque si se acumulan mucho los signos positivos o negativos, quiere decir que ello no es debido al azar sino a que hay una tendencia no explicable por errores aleatorios. De igual modo, si el número de rachas es excesivo quiere decir que hay un continuo y regular cambio de signos positivos y negativos, tampoco explicable por el azar, sino por algún factor periódico que afecta a nuestras mediciones.

Esta prueba, por lo tanto, evalúa si el número de rachas es demasiado pequeño o grande para aceptar H_0 (distribución aleatoria de los signos). Se tabulan para un nivel de significación determinado y un determinado número total de signos positivos y negativos, el número de rachas por debajo y por encima para las que se puede rechazar la hipótesis nula (tabla 12)

Tabla 12: Prueba de Rachas Wald-Wolfowitz (significación del 95%)

Nº signos (+)	Nº signos (-)	Rechazar H0 si Nº rachas es menor que:	Rechazar H0 si Nº rachas es mayor que:
2	12-20	3	*
3	6-14	3	*
3	15-20	4	*
4	5-6	3	8
4	7	3	*
4	8-15	4	*
5	5	3	9
5	6	4	9
5	7-8	4	10
5	9-12	4	*
5	13-18	5	*
6	6	4	10
6	7-8	4	11
6	9-12	5	12
6	13-18	6	*
7	7	4	12
7	8	5	12
7	9	5	12
7	10-12	6	13
8	8	5	13
8	9	6	13
8	10-11	6	14
8	12-15	7	15

Tabla adaptada de Swed FS y Eisenhar C. Ann. Math. Statist. 1943, 14, 66.
** La prueba no puede ser aplicada.*

4. PRUEBAS BASADAS EN EL RECORRIDO DE LOS RESULTADOS

En estas pruebas se usa el recorrido intercuartílico o recorrido total "w" de la muestra como medida de dispersión.

Recorrido "w" se puede definir como la distancia entre el valor inferior y el superior de un conjunto de datos. Recorrido intercuartílico es la distancia entre el P25 y el P75*.

* Un parámetro muy útil para situar puntos en la curva normal y en general en cualquier tipo de distribución, son los percentiles. El percentil de orden k representa el valor de la variable que deja por debajo el k% de los sujetos de la población. Así la **mediana** es la percentil 50 "P_{50}", que es medida de la tendencia central. Un tipo de percentiles muy usados son los cuartiles "$Q_1=P_{25}$, $Q_2=P_{50}$, $Q_3=P_{75}$", que indican, respectivamente, las posiciones en las que quedan por debajo el 25, 50 y 75% de los datos.

Estas pruebas permiten, al igual que la prueba t de student, comparar la media de una población con un valor asumido como verdadero, o bien dos medias de dos muestras diferentes. Estas pruebas, simples pero afectadas por valores anómalos porque se sustituye la desviación estándar por el recorrido, se basan en el cálculo del parámetro T (ecuaciones 2.28 ó 2.29).

$T_i = |X - \mu|/w$ (Ec. 2.28)

μ: valor asumido verdadero.

O bien, al comparar dos grupos de valores 1 y 2:

$T_d = 2|X_1 - X_2|/(w_1 + w_2)$ (Ec. 2.29)

Al igual que la prueba t de student, T_i o T_d calculado se compara con el valor tabulado (tabla 13) para n muestras y un nivel de significación de p=0.05. Si es más pequeño que el tabulado no se puede rechazar H_0. El TTABULADO permite calcular límites de confianza → X ±(T*W).

*Tabla 13. Valores críticos (p=0.05) de la prueba Ti y Td**

n	Ti una cola	Ti dos colas	Td
2	3.175	6.353	3.43
3	0.885	1.304	1.27
4	0.529	0.717	0.81
5	0.388	0.507	0.61
6	0.312	0.399	0.50
7	0.263	0.333	0.43
8	0.230	0.288	0.37
9	0.205	0.255	0.33
10	0.186	0.230	0.30

**Tabla tomada de Lord E. Biometrika, 1947, 34, 66.*

5. VALORES ANÓMALOS "Q DE DIXON"

Esta prueba, usada de forma abusiva, permite eliminar valores extremos. En general se recomienda no aplicarla, sugiriendo siempre mitigar valores anómalos con la recogida de más datos. Cuando la recogida de más datos no es posible y se cree realmente que el valor anómalo es consecuencia de un error sistemático sospechoso, se aplica con precaución. Esta prueba es una variante de prueba de recorrido en la que se calcula el parámetro "Q"

Q= | $X_{sospechoso}$- $X_{más\ próximo}$ | / W (Ec. 2.30)

Si el valor Q calculado, para un determinado nivel de significación (generalmente p=0.05), supera el valor crítico tabulado (tabla 14), se rechaza el valor sospechoso. Una buena aproximación es que un resultado anómalo se puede rechazar si Q es mayor que √(2/n) (p=0.05) ó √(3/n) (p=0.01) siendo n el número de determinaciones, incluyendo el valor anómalo.

*Tabla 14**.

VALORES CRÍTICOS DE Q p= 0.05

Tamaño de muestra	Valor crítico
4	0.85
5	0.73
6	0.64
7	0.59
8	0.54
9	0.41
10	0.48

**Reproducido de W. J. Dixon, Ann Math Stat, 1951, 22, 68.*

6. PRUEBAS DE SIGNIFICACIÓN PARA DATOS ORDENADOS

6.a. Prueba de los signos de Wilcoxon

Se aplica para muestras que se distribuyen de forma simétrica (mediana y media coinciden), pero no se desea suponer que siguen una distribución normal. La distribución simétrica permite desarrollar pruebas más potentes que el test de los signos.

1) Se ordenan las diferencias con respecto a la mediana (o me dia) de mayor a menor valor absoluto, prescindiendo de los signos.

2) Se jerarquizan asignándoles números que indiquen su orden, pero se mantienen sus signos. A los empates se les asignan posiciones promedio.

3) Se suman los rangos positivos por un lado y los negativos por otro. El menor de estos dos sumatorios en valor absoluto, se toma como el estadístico de la prueba.

4) Si el estadístico es menor que el valor tabulado (tabla 15) para un tamaño de muestra determinado y un nivel de significación elegido (generalmente 95%), se puede rechazar la hipótesis nula de que las dos muestras provienen de la misma población con una probabilidad de error menor del 5% ($p < 0.05$).

Tabla 15. Prueba de rangos y signos de Wilcoxon.

VALORES CRÍTICOS (p=0.05)

n	Prueba 1 cola	Prueba 2 colas
5	0	
6	2	0
7	3	2
8	5	3
9	8	5
10	10	8
11	13	10
12	17	13
13	21	17
14	25	21
15	30	25

Tabla tomada de HR Neave. Elementary Statistics Tables. George Allen & Unwin Ltd.

Ejemplo 8: se encontró que los niveles de kaliemia para 7 niños eran de 4.1, 5.0, 4.5, 4.7, 5.1, 3.8. 4.0. ¿podrían proceder los datos de una población simétrica de mediana 4.5?.

1. Las diferencias son -0.4, 0.5, 0, 0.2, 0.6, -0.7, -0.5.
2. Se ordenan de mayor a menor valor absoluto: -0.7, 0.6, 0.5, -0.5, -0.4, 0.2, 0.
3. Se jerarquizan manteniendo el signo: -1, 2, 3.5, -3.5, -5, 6. El cero no se incluye.
4. La suma de los positivos es 11.5 y la de los negativos 9.5. Se toma la menor de estas dos cifras como el estadístico de la prueba (-9.5).

Como 9.5 es mayor que 2, el valor tabulado para una prueba de dos colas con n=7 y significación del 95% (tabla 15), no puede rechazarse H0, porque el error sería mayor del 95%.

Esta prueba se puede usar como alternativa no paramétrica a la prueba t de datos apareados. El ensayo es muy sencillo. Se restan los valores por parejas y con las diferencias se realiza la prueba de los signos como se indica en los 4pasos anteriores. Si el estadístico obtenido es menor que el tabulado, H_0 se puede rechazar con una probabilidad determinada.

6.b Prueba de suma de rangos de Wilcoxon

Es equivalente a la prueba U de Mann-Whitney. Se aplica para comparar dos muestras independientes que no pueden reducirse a un único conjunto de datos y por lo tanto no se puede aplicar la prueba anterior de los signos. Así por ejemplo, se utiliza para muestra, simétricas no normales y de distinto tamaño.

1) Al igual que la prueba de los signos de Wilcoxon, los datos se ordenan y jerarquizan, con un número que indique su posición, de mayor a menor valor absoluto. Sin embargo se prescinde de los signos y los valores de una muestra se distinguen de la otra por una marca (Ej. subrayado). Si las mediciones de las dos muestras fueran indistinguibles, se esperaría que las posiciones subrayadas "1" y las no subrayadas "2" estuvieran mezcladas al azar en la lista.

2) Se calcula el parámetro T_1 y T_2 con las ecuaciones 2.31 y 2.32.

$T_1 = \Sigma_1 - n_1(n_1+1)/2$ (Ec. 2.31)

$T_2 = \Sigma_2 - n_2(n_2+1)/2$ (Ec. 2.32)

Σ_1: sumatorio de las posiciones subrayadas, Σ2: sumatorio de las posiciones no subrayadas
n_1 y n_2: el tamaño de las dos muestras.
$n_1(n_1+1)/2$ y $n_2(n_2+1)/2$: valores esperados para el sumatorio de posiciones de cada muestra

3) El menor de estos parámetros se compara con el valor tabulado para una prueba de una o dos colas, un nivel de significación determinado y tamaños de muestras n_1 y n_2 (tabla 16).

Tabla 16. Prueba de suma de rangos de Wilcoxon (Prueba U de Mann-Withney)

Valores críticos de U o el menor entre T1 y T2 (p=0.05)

n1	n2	Prueba 1 cola	Prueba 2 colas
3	3	0	
3	4	0	
3	5	1	0
3	6	2	1
4	4	1	0
4	5	2	1
4	6	3	2
4	7	4	3
5	5	4	2
5	6	5	3
5	7	6	5
6	6	7	5
6	7	8	6
7	7	11	8

Tabla tomada de HR Neave. Elementary Statistics Tables. George Allen & Unwin Ltd.

Ejemplo 9. Sobre un suero se realizan 10 mediciones de creatinina, 6 por un método A y 4 por un método B. Los resultados se observan en la tabla 17, ordenados de menor a mayor, subrayados los del método A para distinguirlos de los del otro método. Se ha decidido aplicar la prueba de la suma de rangos de Wilcoxon.

Tabla 17

Datos	<u>1.2</u>	1,3	<u>1.4</u>	<u>1.4</u>	<u>1.4</u>	<u>1.5</u>	1.5	<u>1.6</u>	1.7	1.9
Posiciones	<u>1</u>	2	<u>4</u>	<u>4</u>	<u>4</u>	<u>6.5</u>	6.5	<u>8</u>	9	10

T_1=27.5-21=6.5

T_2=27.5-10=17.5

El menor de los T (T_1= 6.5), es mayor que 2, el valor crítico tabulado para tamaños de muestra n_1=6 y n_2=4, siendo una prueba de dos colas con una significación del 95%. H0 no puede ser rechazada porque la probabilidad de error sería mayor del 5% ($p > 0.05$).

Nota: esta prueba se ve afectada cuando los valores promedios de las dos muestras son muy diferentes porque los valores de T1 y T2 serán parecidos, independientemente de la dispersión de ambas. Este problema se puede resolver estimando las medias de las dos muestras, y añadiendo el valor de su diferencia a cada valor de la muestra más pequeña, con lo que se consigue que ambas muestras tengan valores aproximadamente similares.

6.c Test de la z para datos ordenados

Si la muestra tiene más de 10 individuos, la estadística paramétrica renace de nuevo porque se puede utilizar el test de la z con buena aproximación. Sea N el tamaño total de la muestra y m el número de individuos del grupo con mayor parámetro T (ecuación 2.31 ó 2.32). El valor esperado de la suma de rangos es $m(N+1)/2$ y el valor de z calculado es $[T+0.5-(m(N+1)/2)]/\sqrt{(mn(N+1)/12)}$, valor que se compara con la tabla de valores z.

6.d Prueba de Siegel y Tukey

Es una variante de la prueba de suma de rangos de Wilcoxon. Esta prueba, que compara la dispersión de dos conjuntos de resultados, es una alternativa no paramétrica de la prueba F.

1. Se disponen por orden creciente los datos de los dos grupos a comparar, subrayando uno de los conjuntos de resultados para diferenciarlo del otro.

2. Se ordenan de forma que se asigna posiciones bajas a los resultados pequeños y altos y posiciones altas a los resultados centrales: a la medición más pequeña se le asigna la posición 1, a la más grande la posición 2, la medición inmediatamente inferior a la más grande se le asigna la posición 3 y a la posterior de la más pequeña la posición 4, la medición posterior a las dos más pequeñas se le asigna la posición 5 y así sucesivamente. Si el número total de mediciones es impar se ignora el valor central.

3. Se calcula los parámetros T_1 y T_2 con las ecuaciones 2.31 y 2.32. **Interpretación**: si uno de los grupos tuviera una dispersión significativamente mayor que el otro, la suma de sus posiciones debería ser mucho menor, mientras que si las sumas de sus posiciones fueran similares los grupos tendría una dispersión parecida.

4. El menor de estos estadísticos se usa para la prueba y se compara con el valor crítico tabulado, generalmente de una prueba de dos colas, a un p=0.05 (tabla 16). Si el estadístico es mayor que el valor tabulado, la hipótesis nula, es decir, que la dispersión de los dos grupos de datos es similar, debe ser aceptada.

Ejemplo 10. Los resultados del ejemplo 9 se observan ordenados en la tabla 18, subrayados los del método A. Se ha decidido aplicar la prueba de Siegel y Tukey para evaluar si por los dos métodos se obtienen dispersiones diferentes.

Tabla 18

Datos	1.2	1,3	1.4	1.4	1.4	1.5	1.5	1.6	1.7	1.9
Posiciones	1	4	6	6	6	9.5	9.5	8	3	2

T_1= 36.5 21= 15.5

T_2= 18.5 - 10 = 8.5

El menor de los T (T_2= 8.5), es mayor que 2, el valor crítico tabulado para tamaños de muestra n_1=6 y n_2=4, siendo una prueba de dos colas con una significación del 95%. H0 no puede ser rechazada, los dos métodos no son significativamente diferentes (p> 0.05).

6.e Prueba de Kruskal-Wallis

Extensión del procedimiento de suma de rangos de Wilcoxon para comparar medianas de más de 2 muestras sin ningún tipo de emparejamiento.

6.f Prueba de Friedman

Se aplica con 3 o más conjuntos de resultados emparejados. Fue ideada por el famoso economista americano Friedman y utiliza de nuevo el estadístico X^2. Es similar en la práctica al método ANOVA aunque no tiene su capacidad.

B.7. MEDIDAS DE RELACIÓN PARA DATOS ORDENADOS. PRUEBA DE CORRELACIÓN ORDINAL DE SPEARMAN

La prueba de correlación de Spearman es la más sencilla y empleada. Se calcula un coeficiente de correlación ordinal de Spearman "ρ", alternativa simple no paramétrica a la correlación lineal de Pearson. Se aplica cuando los datos están correlacionados linealmente pero no se distribuyen normalmente.

Se calcula "ρ" (ecuación 2.33).

$\rho = 1 \ \ [6\ \Sigma d_i^2/(N^3\text{-}N)]$ Ec. 2.33

N: tamaño de la muestra.

d_i: diferencia de rangos entre la dos medidas para el caso i

Al igual que para el coeficiente de Pearson, el test de significación del coeficiente de Spearman se basa en un t-test (ecuación 2.34).

$t= \rho/\sqrt{[(1\text{-}\rho^2)/(N\text{-}2)]}$ Ec. 2.34

III) *Interpretación de las pruebas de significación*

Hemos visto numerosos tipos de pruebas de significación, la mayoría de los cuales se incluyen en la tabla 19 que nos puede orientar en la prueba a aplicar según el tipo de variable dependiente e independiente.

Tabla 19

Variable Independiente	Dependiente	Cualitativa 2 Categorías	Cualitativa >2Categorías	Ordinal	Cuantitativa
Cualitativa 2 categorías	Independientes	X^2 ó Fisher	X^2	U de Mann-W	t Student
	Apareados	$X^2{}_M$ ó Fisher	Q de Cochran	Signos de W	t Student
Cualitativa >2 Categorías	Independientes	X^2	X^2	K-W	
	Apareados	Q de Cochran	Q de Cochran	Friedman	Anova
Cuantitativa		t Student	Anova	Spearman Tau de Kendall	Pearson Regresión lineal

X^2: prueba de la chi cuadrado (ap. II.B.1.a), $X^2{}_M$: prueba de chi cuadrado de McNemar (ap. II.B.1.b), U de Mann-W: prueba de la suma de rangos de Wilcoxon o prueba U de Mann-Whitney (ap. II.B.6.b), Fisher: test exacto de Fisher (ap. II.B.1.d), Signos de W: prueba de los signos de Wilcoxon (ap II.B.6a), K-W: prueba de Kruskal-Wallis (ap II.B.6.e), Friedman: prueba de Friedman (ap II.B.6.f), Spearman: prueba de correlación ordinal de Spearman (ap II.B.7), Tau de Kendall: prueba de correlación de datos ordenados al igual que la prueba de Spearman, Pearson: coeficiente de correlación de Pearson (ap 5.B.1), Anova: análisis de variancia (ap II.A.4), t Student: prueba t de Student-Fisher (ap II.A.1).

Nota: cuando las variables cuantitativas no cumplen la ley de distribución normal, se transforman las variables en cualitativas ordinales con varias categorías y se aplica el test no paramétrico correspondiente.

La relación causa-efecto entre dos variables de una investigación no puede confirmarse sólo por la aplicación de una prueba de significación. Hay unas condiciones básicas que deben cumplirse para que una relación entre variables pueda ser interpretada en sentido causal:

- **Temporalidad**: la causa debe preceder al efecto.

- **Asociación**: obtener probabilidades significativas al aplicar pruebas estadísticas.
- **Ausencia de espuriedad**: que no haya una 3ª variable causa de ambas.

Es imprescindible para medir la relación causal entre las variables investigadas, que el diseño de la investigación sea el correcto. Aplicar un test de significación es sencillo y con ligeros conocimientos en estadística también es fácil elegir el correcto entre la batería de los disponibles, que no son todos los que en este resumen se han descrito. Hay otros problemas en estadística mucho más importantes que no se consideran, por desconocimiento o por dificultad pero que marcan de forma irreversible la calidad de toda la investigación y la fiabilidad de los datos y conclusiones logradas.

Multitud de laboratorios toman muestras de una población (paciente) sin valorar si son representativas, gastan enormes recursos en su análisis y los resultados obtenidos se creen asociados al paciente, cuando en realidad, casi siempre, están asociados solamente a esa muestra. Y lo que es peor, se toman decisiones clínicas sobre la base de resultados obtenidos de una sola muestra, con un solo análisis de la misma, eso sin tener en cuenta los errores sistemáticos graves que se pueden cometer por mala calibración y control del método analítico, ni los errores por confusión entre muestras y/o informes que se cometen al tratar conjuntamente tantas muestras, como se tratan en los laboratorios de análisis clínico.

En investigación los errores graves cometidos sobre muestras individuales se mitigan, pero aparecen muchos defectos que invalidan los resultados obtenidos:

* Población a estudiar no idónea y/o sesgada.

* Existencia de sesgos en la muestra escogida de la población a estudiar (muestra no representativa).

* Obviar que las diferencias entre los grupos escogidos pueden ser debidas a errores sistemáticos y/o aleatorios. Querer ver diferencias demasiado pequeñas como apreciables.

* Mala elección del grupo control de comparación, o elección sesgada del mismo.

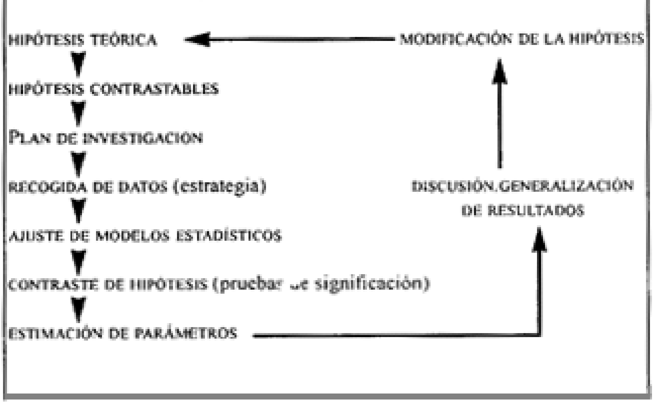
* Nulo control en las variables que modifican los resultados obtenidos (tiempo de toma de la muestra, origen y condiciones de conservación y procesamiento de la misma, ...).

* Falta de diseño previo en la investigación o diseño erróneo.

* Tratamiento subjetivo de los resultados obtenidos (descarte de datos anómalos o que no interesan, falsedad en el tamaño de la muestra analizada, falta de comprobación de la normalidad de la muestra, aplicación de tests inapropiados por interés, ...).

El control de calidad y diseño de investigación pueden ser temas para otro trabajo, ahora sólo indicar de forma esquemática los pasos a seguir en la investigación de la población (Figura.5).

Figura 5. Esquema básico del diseño de investigación



Bibliografía básica y complementaria:

Devore, Jay L. Probabilidad y estadística para ingeniería y ciencias. Internacional Thompson Hildebrand, David K. & Ott, Lyman R. Estadística aplicada a la administración y la economía. Addison-Wesley Iberoamericana