UTPL

La Universidad Católica de Loja

Modalidad Abierta y a Distancia

Sistemas de Conocimiento de Estadística Inferencial y su Didáctica

Guía didáctica

UTPL

Departamento de Ciencias de la Educación

Sección departamental de Pedagogía de las Ciencias Experimentales

Sistemas de Conocimiento de Estadística Inferencial y su Didáctica

Guía didáctica

Autora:

Granda Sivisapa Sonia Patricia

EDUC_3152

Asesoría virtual
www.utpl.edu.ec

Sistemas de Conocimiento de Estadística Inferencial y su Didáctica
Guía didáctica
Granda Sivisapa Sonia Patricia

Universidad Técnica Particular de Loja

Diagramación y diseño digital:

Ediloja Cía. Ltda.
Telefax: 593-7-2611418.
San Cayetano Alto s/n.
www.ediloja.com.ec
edilojainfo@ediloja.com.ec
Loja-Ecuador

ISBN digital - 978-9942-39-098-1

17 de marzo, 2021

Índice

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

1. Datos de información

1.1. Presentación de la asignatura

1.2. Competencias genéricas de la UTPL

- Vivencia de los valores universales del humanismo de Cristo.
- Pensamiento crítico y reflexivo.
- Compromiso e implicación social.
- Comportamiento ético.
- Orientación a la innovación y a la investigación.
- Comunicación oral y escrita.

1.3. Competencias específicas de la carrera

- Integra conocimientos pedagógicos, didácticos y curriculares a través del uso de herramientas tecnológicas pertinentes que permitan interdisciplinariamente la actualización de modelos y metodologías de aprendizaje e incorporación de saberes en matemáticas y física, basados en el desarrollo del pensamiento crítico, reflexivo, creativo, experiencial y pertinentes en relación con el desarrollo de la persona y su contexto.

- Implementa la comunicación dialógica como estrategia para la formación de la persona orientada a la consolidación de capacidades para la convivencia armónica en la sociedad, la participación ciudadana, el reconocimiento de la interculturalidad y la diversidad, y la creación de ambientes educativos inclusivos en el ámbito de las matemáticas y la física, a partir de la generación, organización y aplicación crítica y creativa del conocimiento abierto e integrado en relación a las características y requerimientos de desarrollo de los contextos.

- Organiza modelos curriculares y la gestión del aprendizaje relacionados con las matemáticas y la física, centrados en la experiencia de la persona que aprende, en interacción con los contextos institucionales, comunitarios y familiares, orientados al diseño de procesos de enseñanza- aprendizaje y de evaluación que integren la práctica de investigación acción hacia producción e innovación, la interculturalidad, la inclusión, la democracia, la flexibilidad metodológica para el aprendizaje personalizado, las interacciones virtuales, presenciales y las tutoriales.

- Potencia la formación integral de la persona desde los principios del humanismo de Cristo y del Buen Vivir, basado en el desarrollo de su proyecto de vida y profesional que amplíen perspectivas, visiones y horizontes de futuro en los contextos.

1.4. Problemática que aborda la asignatura

El análisis de las características del comportamiento de algún suceso o evento, es uno de los objetivos de la estadística como parte de la matemática, hoy en día no basta con solo reunir datos y calcularlos, lo que interesa en realidad son los procesos donde se utilizan datos de una muestra para realizar inferencias y tomar decisiones con respecto a la población de la cual se toma dicha muestra, ya que esto nos permite conocer la realidad y representarla, este es el fin mismo de la estadística inferencial, la que busca formular generalizaciones y a partir de esto facilitar la toma de decisiones dentro del campo de la investigación. La elaboración de inferencias a partir de pruebas aplicadas en los datos de una muestra se constituye en un reto para los investigadores, quienes deberán volverse expertos en la elección y aplicación de pruebas estadísticas paramétricas y no paramétricas, temas de aprendizaje de esta asignatura.

2. Metodología de aprendizaje

Con la finalidad de convertir al estudio de la asignatura Sistemas de Conocimientos de Estadística Inferencial y su Didáctica, en una experiencia de aprendizaje permanente que permita aplicar las técnicas de muestreo y los métodos de estimación paramétrica en casos de estudio, así como utilizar pruebas de hipótesis y pruebas estadísticas no paramétricas en el estudio de casos, se diseñan actividades y estrategias fundamentadas en una metodología activa, en donde los estudiantes acceden a los aspectos iniciales de la conceptualización teórica observando micro videos y accediendo a recursos planteados en el EVA, mientras que el trabajo en contacto con el docente se convierte en un taller de aplicación, dando el enfoque experimental en esta segunda parte, de tal manera que se consolidan los aprendizajes de manera experimental, apoyado en videoconferencias, semanales que serán grabadas para aquellos estudiantes que no pudieron participar en línea.

Aunque no existe un procedimiento único, lo esencial para el estudio de este curso es la utilización del texto básico, guía didáctica y recursos educativos abiertos, los que apoyados en otros recursos como los cuestionarios interactivos y las infografías; anticipan el estudio de los aspectos teóricos que serán consolidados luego, con la práctica experimental al momento de desarrollar los aportes de los estudiantes, principalmente los cuantificados, contando con la presencia virtual del profesor en algunos casos de forma síncrona y casi siempre a través de encuentros asincrónicos, fomentando el trabajo colaborativo en función de los distintos estilos de aprendizaje y enfocados a lograr resultados de aprendizaje que orientan el desarrollo de competencias.

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

3. Orientaciones didácticas por resultados de aprendizaje

Primer bimestre

Resultado de aprendizaje 1

Aplica las técnicas de muestreo y los métodos de estimación paramétrica en casos de estudio.

Para alcanzar el resultado de aprendizaje propuesto, se lleva a cabo una revisión sistemática de los conocimientos previos en matemática especialmente la estadística descriptiva, esto unido a la solución de problemas de aplicación, le permitirá a usted profesional en formación en un futuro administrar procesos de aprendizaje, fomentando cambios y provocando innovaciones en el desarrollo del pensamiento lógico, crítico y creativo de sus educandos.

Estimado estudiante, bienvenido al curso de: **Sistemas de Conocimiento de Estadística Inferencial y su Didáctica.**

A continuación, comparto con usted las orientaciones didácticas necesarias para lograr el resultado de aprendizaje planteado en este bimestre, con este propósito para cada semana se dará a conocer los contenidos a estudiar, los recursos, las actividades de aprendizaje recomendadas y evaluadas. Es importante que realice todas las actividades propuestas en el plan docente, con la ayuda de su texto base y la presente guía, esto lo capacitará para rendir sus evaluaciones de forma satisfactoria.

Contenidos, recursos y actividades de aprendizaje

Semana 1

Unidad 1. Distribuciones de probabilidad

Estimada(o) profesional en formación es momento de continuar sustentando el campo científico de su carrera a través del estudio de la distribución de probabilidad. Esta unidad corresponde al capítulo 6 de su texto base, donde se explica que, para continuar con el análisis estadístico de datos, luego de su detección y corrección de errores, se debe describir las distribuciones de las variables estudiadas, para esto contamos con la distribución normal, también llamada distribución Gaussiana, cuya importancia se fundamenta en la frecuencia con la que distintas variables asociadas a fenómenos naturales y cotidianos se le aproximan. Una forma práctica de acceder a estas definiciones es interactuando con la Máquina de Galton virtual dónde se simuló el lanzamiento de 250 bolas, lo invito

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

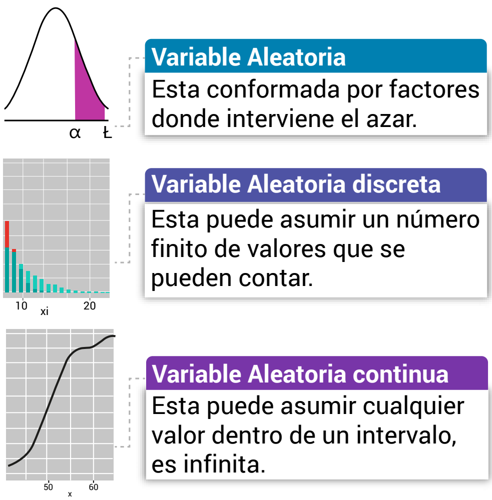
a seguir las indicciones de las leyendas de la aplicación, así acceder al histograma binomial teórico y la gráfica de la distribución normal Bienvenido!

1.1. Conceptos y aplicaciones teórico-prácticas de las distribuciones de probabilidad

Iniciamos considerando algunos conceptos claves que se unen a lo estudiado previamente en la estadística descriptiva en el ámbito de la recolección de datos, elaboración de tablas o cuadros, gráficos y la aplicación de ciertas medidas de posición y dispersión.

"Se dice que una ***distribución de probabilidad*** muestra los resultados esperados al realizar un experimento, junto a la probabilidad esperada para cada una de ellos" (Martínez, 2011, p. 213). A si mismo se requiere diferenciar entre los diferentes tipos de variables:

Figura 1.
Tipos de variables



Nota. La figura muestra los tipos de variables necesarias para el aprendizaje de la distribución de probabilidades. Fuente: Martínez. (2019). Elaboración: Granda S. (2020).

Luego de que usted haya analizado las páginas 218 y 219 de su texto base, donde se diferencia la distribución de frecuencias y la distribución de probabilidad daremos inicio al estudio de la variable aleatoria discreta, a la distribución binomial.

1.2. Variable aleatoria discreta: Distribución Binomial

Se denomina distribución binomial al método exacto que corresponde a una distribución de variable aleatoria discreta, tome en cuenta que sus áreas de aplicación están dentro de las inspecciones de calidad, ventas, mercadotecnia, etc. Para comprender este concepto resulta útil revisar el Binomio de Newton y el Triángulo de Pascal estudiados en asignatura anteriores.

Si consideramos *p=probabilidad de éxito* y *q=probabilidad de fracaso* la distribución binomial se calcula con la fórmula $P_x = \binom{n}{x} p^x q^{n-x}$, recuérdese que $(n\ x) = (n\ n-x) = \left(\frac{n!}{(n-x)!x!}\right)$ por lo que podemos modificar la expresión de que finalmente se expresara como $P_{(x)} = \left[\frac{n!}{(n-x)!x!}\right] p^x q^{n-x}$ cabe recalcar que esta distribución debe cumplir según Martínez (2013) con determinadas condiciones:

- Solo se tiene dos resultados; éxito o fracaso.
- El resultado que se pide deberá tomarse siempre como éxito.
- La probabilidad de éxito está representada por , por lo tanto, siempre será fijo o constante en un ensayo.
- La probabilidad de fracaso es *q=1 - p* ya que *p + q = 1*
- Se requiere que los ensayos sean independientes, es decir, que el resultado esperado en uno de ellos no afecte a los otros.
- Siempre se tendrá un número de ensayos repetidos.
- Los resultados serán establecidos mediante el conteo o recuento. (p. 215)

A continuación, procedo a explicarle los pasos para determinar una distribución binomial por medio de un problema de aplicación, le recomiendo analizar e interiorizar los procesos desarrollados.

PROBLEMA MODELO 1 Distribución de probabilidad

PROBLEMA MODELO # 1

Resolución de un problema modelo de Distribución Binomial.

En una facultad, la probabilidad de que un alumno apruebe el semestre es del 80%. Si consideramos 8 alumnos, ¿Cuáles la probabilidad de que:

a) dos ganen el semestre *b) dos pierdan el semestre*

Solución

a. Este procedimiento satisface los requisitos de una distribución binomial, como se muestra a continuación:

1. El número de ensayo (8) es fijo.
2. Los 8 ensayos son independientes, ya que la probabilidad de que cualquier estudiante apruebe el semestre, no se ve afectado por el resultado de cualquier otro estudiante.
3. Cada uno de los 8 ensayos tiene dos categorías de resultados: el estudiante puede aprobar o no aprobar el semestre.
4. Para cada estudiante, la probabilidad de que apruebe el semestre es 0.80 (80%), siendo la misma probabilidad para los 8 estudiantes.
5. Luego de haber concluido que el procedimiento da como resultado una distribución binomial, ahora procederemos a identificar los valores de ***n, x, p y q***
6. Con 8 estudiantes, tenemos ***n*= 8**
7. Queremos conocer la probabilidad de diferentes números de éxitos ***x***, según la condición del problema. (variable aleatoria)
8. La probabilidad de éxito (aprobar el semestre) en un ensayo es de 0.80, de modo que ***p* = 0.80**
9. La probabilidad de fracaso (no aprobar el semestre) es 0.20 de modo que ***q* = 0.20**.

NOTA: Es muy importante asegurarse de que ***x*** y ***p*** se refieren al mismo concepto de "éxito". En este ejemplo, se utiliza ***x*** para contar el número de aprobados, de modo que ***p*** debe ser la probabilidad de que un estudiante apruebe el semestre. Por consiguiente, ***x*** y ***p*** utilizan el mismo concepto de éxito (aprobar el semestre).

Para calcular las probabilidades correspondientes a la variable aleatoria ***x*** en una distribución binomial contamos con tres métodos:

1. Realizar cálculos mediante la fórmula *de probabilidad binomial* (base de los siguientes dos métodos)
2. Uso de un programa de computo o de una calculadora
3. El uso de la tabla del texto base TABLA III

En esta ocasión aplicaremos el primer método:

a. Dos aprueben → ***n* = 8**, ***x* = 2**, (variable aleatoria, aprobar)

p = 0.8 (aprobar)
q = 0.2 (no aprobar)

Anotemos la fórmula de la probabilidad binomial:

$$P_{(x)} = \left[\frac{n!}{(n-x)!\,x!}\right] p^x q^{n-x}$$

Procedemos a determinar $P_{(x=2)}$, entonces tenemos:

$$P_{(x=2)} = \left[\frac{8!}{(8-2)!\,2!}\right] 0.8^2 0.2^{8-2}$$

$$P_{(x=2)} = \left[\frac{8!}{6!\,2!}\right] 0.8^2 0.2^6$$

$$P_{(x=2)} = \left[\frac{8!}{6!\,2!}\right] 0.8^2 0.2^6$$

$$P_{(x=2)} = \left[\frac{8 \times 7 \times 6!}{6!\,2!}\right] (0.64)(6.4 \times 10^{-5})$$

Se simplifica **6!**

$$P_{(x=2)} = \left[\frac{8 \times 7}{2!}\right] (0.64)(6.4 \times 10^{-5})$$

$$P_{(x=2)} = \left[\frac{8 \times 7}{2 \times 1}\right] (0.64)(6.4 \times 10^{-5})$$

$$P_{(x=2)} = \left[\frac{56}{2}\right] (0.64)(6.4 \times 10^{-5})$$

$$P_{(x=2)} = [28](0.64)(6.4 \times 10^{-5})$$

$$P_{(x=2)} = 0.001147$$

Multiplicamos por 100%

$$P_{(x=2)} = 0.11\%$$

Solución:

La probabilidad de que dos estudiantes aprueben el semestre es del 0.11%

b. Dos no aprueben → ***n* = 8**, ***x* = 2** ,(variable aleatoria, no aprobar)

***p* = 0.2** (no aprobar)
***q* = 0.8** (aprobar)

Anotemos la fórmula de la probabilidad binomial:

$$P_{(x)} = \left[\frac{n!}{(n-x)!\,x!}\right] p^x q^{n-x}$$

Procedemos a determinar $P_{(x=2)}$, entonces tenemos:

$$P_{(x=2)} = \left[\frac{8!}{(8-2)!\,2!}\right] 0.2^2 0.8^{8-2}$$

$$P_{(x=2)} = \left[\frac{8!}{6!\,2!}\right] (0.04)(0.8)^6$$

$$P_{(x=2)} = \left[\frac{8 \times 7 \times 6!}{6!\,2!}\right] (0.04)(0.8)^6$$

$$P_{(x=2)} = \left[\frac{8 \times 7}{2 \times 1}\right] (0.04)(0.262144)$$

$$P_{(x=2)} = [28](0.04)(0.262144)$$

$$P_{(x=2)} = 29.36\%$$

La probabilidad de que dos estudiantes no aprueben el semestre es del 29.36%

c. Por lo menos dos no aprueben → ***n* = 8**, ***x* = 2, 3, 4, 5, 6, 7, 8,** (variable aleatoria, no aprobar)

***p* = 0.2** Probabilidad del éxito (no aprobar), porque esto se necesita en este caso.
***q* = 0.8** Probabilidad del fracaso (aprobar)

El valor de ***x*** está en el intervalo **2 ≤ x ≤ 8** o ***P* (2 ≤ x)** es decir de 2 hacia arriba sin pasarnos de 8 y empezando en 2.

Para determinar la probabilidad en este caso tenemos dos posibilidades:

- Distribución binomial acumulada

***P* (2 ≤ x) = *P* (x = 2) + *P* (x = 3) + *P* (x =4) + *P* (x = 5) + *P* (x = 6) *P* (x = 7) + *P* (x = 8)**

- Complemento

1 – [*P* (x = 0) + *P* (x = 1)]

Con las dos posibilidades se obtendrá el mismo resultado, lo haremos con la segunda opción por ser más corta.

$$P_{(2\leq x)} = 1 - \left\{\left[\frac{8!}{(8-0)!\,0!}\right]0.2^0 0.8^{8-0} + \left[\frac{8!}{(8-1)!\,1!}\right](0.2^1)(0.8^{8-1})\right\}$$
$$P_{(2\leq x)} = 1 - \left\{\left[\frac{8!}{8!\,0!}\right]0.2^0 0.8^8 + \left[\frac{8!}{(7)!\,1!}\right]0.2 \times 0.8^7\right\}$$
$$P_{(2\leq x)} = 1 - \left\{1 \times 0.8^8 + \left[\frac{8 \times 7!}{(7)!\,1!}\right](0.2)(0.2097152)\right\}$$
$$P_{(2\leq x)} = 1 - \{0.16777216 + [8](0.2)(0.2097152)\}$$
$$P_{(2\leq x)} = 1 - \{0.16777216 + 0.33554432\}$$
$$P_{(2\leq x)} = 1 - \{0.50331648\}$$
$$P_{(2\leq x)} = 1 - \{0.50331648\}$$
$$P_{(2\leq x)} = 0.49668352$$
$$P_{(2\leq x)} = 49.67\%$$

La probabilidad de que por los menos dos estudiantes no aprueben el semestre es del 49.67%

Nota. La información muestra el proceso de aplicación de la distribución binomial al determinar el porcentaje de ocurrencia de aprobación o no aprobación de un semestre en un grupo de estudiantes de una determinada facultad. Fuente: Martínez. (2019). Elaboración: Granda S. (2020)

Actividades de aprendizaje recomendadas

Al concluir el trabajo de la semana, donde se dio inició al análisis de distribución de probabilidad y distribución binomial, recuerde revisar en su texto base los conceptos y aplicaciones teórico prácticas de estas distribuciones y proceda a realizar las siguientes actividades:

1. Observar el recurso audiovisual ¿Qué es la distribución normal?

2. Ingresar a Estadística (distribución de probabilidad 1) e ingresar los datos que le permiten inferir el comportamiento de una distribución de probabilidad.

3. Interactuar con la aplicación Distribución binomial. Aquí usted podrá explorar distintas situaciones para los parámetros n y p, le invito a analizar el gráfico de la distribución binomial, cuando el valor de n no es muy grande el valor de p se acerca a 0.5.

4. Continuar con la resolución del problema modelo #1 en dos situaciones adicionales descritas en el siguiente enunciado:

> ***En una facultad, la probabilidad de que un alumno apruebe el semestre es del 80%. Si consideramos 8 alumnos, ¿Cuáles la probabilidad de que: ¿cómo máximo 6 ganen?***

RETROALIMENTACIÓN

Los recursos audiovisuales nos muestran la definición y relación entre la distribución de probabilidad, normal y binomial, con ejemplos prácticos que le permitirá acceder a un registro de representación de conceptos, que le serán de mucha utilidad para su aprendizaje.

Por otro lado con la solución del problema propuesto, usted pudo evidenciar su aprendizaje acerca de la distribución binomial en un casos adicional al explicado en el problema modelo donde para resolverlo utilizamos:$n = 8$, $x = 0,1,2,3,4,5,6$, $p = 0.8$, $q = 0,2$ y $P\,(x \leq 6)$ $\boldsymbol{P}\,(0 \leq x \leq 6)$ y la probabilidad se termina utilizando el complemento $P\,(x \leq 6) = 1 - [P\,(x = 7) + P\,(x = 8)]$,de haber existido alguna dificultad es necesario que revise una vez más su texto base y así lograr los resultados de aprendizaje esperados.

Con todo el contenido precedente damos por culminada la actividad planificada para la primera semana, se agradece su entusiasmo y participación en cada una de ellas y se le invita a continuar con el mismo ahínco en la siguiente semana.

¡Felicitaciones por su trabajo oportuno y efectivo en las actividades planteadas!

> *Recuerde que:*
>
> *A través del chat de tutoría y consultas puede aclarar sus dudas e inquietudes con su docente tutor.*
>
> *El **éxito** en la vida no se mide por lo que logras sino por los **obstáculos** que superas.*

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Semana 2

Estimado estudiante, en la semana anterior conocimos las distribuciones de probabilidad binomial, y en esta semana nos aprestamos a estudiar la distribución probabilística discreta denominada de Poisson, que según explica Miller (2010) "Expresa la probabilidad de un número K de eventos ocurriendo en un tiempo fijo si estos eventos ocurren con una frecuencia media conocida y son independientes del tiempo discurrido desde el último evento" (p.3). esta definición se puede apreciar claramente, por medio de la interacción con la aplicación Poisson Approximation to Binomial Distribution. Luego de esta revisión lo invito a que continúe con el aprendizaje de este interesante tema.

1.3. Distribución probabilística de Poisson

Otra distribución de probabilidad discreta importante es la distribución de Poisson, que se utiliza para describir comportamientos que ocurren en raras ocasiones (con posibilidades pequeñas), algunos ejemplos pueden ser el decaimiento radiactivo, la llegada de una persona a la fila, la reproducción de águilas en una región, los pacientes que llegan a las salas de emergencia, los usuarios de Internet que visitan un sitio Web, para estos casos y otros contamos con este tipo de distribución.

> *Definición de Distribución de Poisson*
>
> La distribución de Poisson es una distribución de probabilidad discreta que se aplica a las ocurrencias de algún evento durante un intervalo

> especifico. La variable aleatoria ***x*** es el número de veces que ocurre un evento en un intervalo. El intervalo puede ser tiempo, distancia, área, volumen o alguna unidad similar. (Triola,2013, p.235)

La probabilidad de que el evento ocurra *x* veces durante un intervalo está dada por la fórmula: $P(x) = \frac{\lambda^{x} \cdot e^{-\lambda}}{x!}$ Donde ***e* = 2.71828**

Tome en cuenta los requisitos de la distribución de Poisson.

I. La variable aleatoria *x* es el número de veces que ocurre un evento durante un intervalo.
II. Las ocurrencias deben ser aleatorias
III. Las ocurrencias deben ser independientes entre sí.
IV. Las ocurrencias deben estar uniformemente distribuidas dentro del intervalo considerado.

Finalmente, Triola (2013) sostiene que la distribución de Poisson difiere de una distribución binomial en estas formas fundamentales:

1. La distribución binomial se ve afectada por el tamaño de la muestra *n* y la probabilidad *p*, mientras que la distribución de Poisson solo se ve afectada por la media *λ*.

2. En una distribución binomial, los valores posibles de la variable aleatoria *x* son 0,1,…..*n*, pero los valores posibles de *x* de una distribución de Poisson son 0,1,2, .., sin límite superior.(p.235)

Luego de este preámbulo, procedo a explicarle los pasos para determinar una distribución de probabilidad de Poisson, por medio de un problema de aplicación, es necesario analizar e interiorizar los procesos desarrollados.

PROBLEMA MODELO # 2

Resolución de problema modelo de Distribución de Probabilidad de Poisson.

Si la probabilidad de que una persona adquiera la enfermedad como consecuencia de una vacuna contra la misma, es 0,0001 ¿Cuál es la probabilidad de que la adquieran exactamente 5 personas en una población de 10.000 vacunados?

Solución:

Determinamos las media: .λ **= 10.000 (0.001) = 1** y el número de casos favorables **x = 5**. Recuerde ***e*** **= 2.71828**

Tomamos la función de probabilidad

$$P(x) = \frac{\lambda^x \cdot e^{-\lambda}}{x!}$$

$$P(x=5) = \frac{1^5 \cdot e^{-1}}{5!}$$

$$P(x=5) = \frac{1 \cdot (0.367879)}{5 \times 4 \times 3 \times 2 \times 1}$$

$$P(x=5) = \frac{0.367879}{120}$$

$$P(x=5) = 3.06566201 \times 10^{-3}$$

Expresado como porcentaje:

$$P(x=5) = 0.31\%$$

La probabilidad de que la adquieran exactamente 5 personas es 0.31%

Nota. Se muestra el proceso de aplicación de la distribución de probabilidad de Poisson al determinar la probabilidad de contagio de una enfermedad de 5 personas como consecuencia de la aplicación de una vacuna. Fuente: Martínez. (2019). Elaboración: Granda S. (2020)

Para una explicación más detallada de cómo aplicar la distribución de probabilidad de Poisson, se procede a explicar el proceso de resolución de un problema más de aplicación.

PROBLEMA MODELO # 3

Resolución de problema modelo de Distribución de Probabilidad de Poisson.

Los clientes de una cafetería llegan a razón de nueve personas, en un período de 30 minutos. Calcule la probabilidad, a) de que en la primera media hora por lo menos lleguen 4 personas. b) de que en los 10 primeros minutos no llegue ningún cliente.

Solución:

Solución:

a. Determinamos las media: ***λ = 9*** en 30 min y el número de casos favorables ***x ≥ 4***. Recuerde ***e = 2.71828***

Tomamos la función de probabilidad tomando el complemento de las posibilidades dadas:

$$P(x \geq 4) = 1 - [P(x=0) + P(x=1) + P(x=2) + P(x=3)]$$

$$P(x \geq 4) = 1 - \left[\frac{9^0 \cdot e^{-9}}{0!} + \frac{9^1 \cdot e^{-9}}{1!} + \frac{\lambda^2 \cdot e^{-9}}{2!} + \frac{9^3 \cdot e^{-9}}{3!}\right]$$

$$P(x \geq 4) = 1 - (0.000123 + 0.000111 + 0.000499 + 0.015)$$

$$P(x \geq 4) = 1 - (0.01622)$$

$$P(x \geq 4) = 0.983$$

Expresado como porcentaje:

$$P(x \geq 4) = 98.3\%$$

La probabilidad de que en la primera media hora por lo menos lleguen 4 personas es del 98.3%.

---

b. De que en los 10 primeros minutos no llegue ningún cliente.

En este caso determinamos el intervalo de tiempo $\frac{\mathbf{30\ minutos}}{\mathbf{10\ minutos}} = \mathbf{3}$ y el valor de la media se ve afectada por dicho intervalo $\boldsymbol{\lambda} = \frac{\mathbf{9}}{\mathbf{3}} = \mathbf{3}$.

$$\boldsymbol{P(x) = \frac{\lambda^{x} \cdot e^{-\lambda}}{x!}}$$

$$\boldsymbol{P(x=0) = \frac{3^{0} \cdot e^{-3}}{0!} = 0,0497 = 4.97\%}$$

En los primeros 10 minutos existe la probabilidad de 4.97% de que no llegue ningún cliente.

Nota. Se muestra el proceso de aplicación de la distribución de probabilidad de Poisson al determinar la probabilidad de concurrencia a una cafetería. Fuente: Martínez. (2019). Elaboración: Granda S. (2020)

Actividades de aprendizaje recomendadas

Al terminar la segunda semana se ha abarcado la definición, aplicación e importancia de la distribución de Poisson, para afianzar estos conocimientos recuerde revisar en su texto base, los conceptos y aplicaciones teórico prácticas de la misma, revise además la explicación para realizar cálculos aplicando la tecnología, por medio del programa EXCEL y luego proceda a realizar las siguientes actividades:

1. Contestar la siguiente pregunta ¿Todas las distribuciones de probabilidades discretas son binomiales o de Poisson? ¿Por qué?

2. Revisar en el texto base en la página 237 la aplicación del EXCEL.

3. Observar el recurso audiovisual Distribución de Poisson ¿Cuándo usarla? Ejercicio resuelto

4. Resolver los problemas # 56, 60 y 64 de la página 241 de su texto base.

5. Revisar la Aplicación del Excel y uso de la calculadora de su texto base, en las páginas 237-240.

RETROALIMENTACIÓN

Al contestar esta pregunta usted pondrá en juego su habilidad de análisis entre los diferentes tipos de distribuciones y como se afectan por el tamaño de la muestra, y las variables. Así mismo podrá desarrollar operaciones de forma rápida y confiable utilizando la tecnología. Finalmente, en el recurso audiovisual se explica el proceso para desarrollar la distribución de Poisson, cuando es conveniente aplicarla y si los resultados obtenidos son adecuados.

Al finalizar esta semana de aprendizaje de la distribución de Poisson, le invito a continuar por la ruta del aprendizaje de las distribuciones de probabilidad.

¡Siga perseverando en su desempeño, lo está haciendo excelentemente, felicitaciones!

> Recuerde que:
>
> A través del chat de tutoría y consultas puede aclarar sus dudas e inquietudes con su docente tutor.
>
> Creé en ti y todo será posible

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Semana 3

Estimado estudiante en la semana anterior conocimos las distribuciones de probabilidad binomial, en la presente semana nos aprestamos a estudiar la variable aleatoria continua y su distribución normal, es importante destacar que:

> La distribución normal fue reconocida por primera vez por el francés Abraham de Moivre (1667-1754). Posteriormente, Carl Friedrrich Gauss (1777-1855) elaboró desarrollos más profundos y formuló la ecuación de la curva; de ahí que también se la conozca, más comúnmente, como la "**campana de Gauss**". (Peretegas y Pita ,2001, p.1).

La campana de Gauss puede ser reconocida por usted al interactuar con la aplicación Normal Distribution donde usted puede variar los parámetros estadísticos y observar la variabilidad del área bajo la curva. Todo este contenido será analizado a continuación, por lo que invito a que juntos exploremos este interesante tema.

1.4. Variable aleatoria continua: Distribución normal

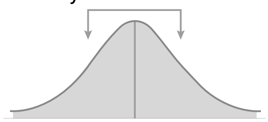
Las distribuciones normales ocurren con gran frecuencia en las aplicaciones reales y desempeñan un papel fundamental en los métodos de estadística inferencial.

A continuación, se expone la definición de la distribución normal, sin embargo, recomiendo acudir al texto base, donde se explica detalladamente la definición y elementos de la distribución normal en la página 247.

Figura 2.
Definición de Distribución Normal

Definición

Si una variable aleatoria continua tiene una distribución con una gráfica simétrica



y en forma de campana, y puede expresarse por medio de la fórmula $Y = \frac{N}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2\sigma^2}}$, decimos que tiene una ***distribución normal***

Nota. La figura explica la definición de la distribución normal para una variable aleatoria continua. Fuente: Triola. (2013). Elaboración: Granda S. (2020)

La distribución normal estándar tiene las siguientes tres propiedades:

1. Su gráfica tiene forma de campana.
2. Posee una media igual a 0 (es decir, $\mu = 0$)
3. Tiene una desviación estándar igual a 1 (es decir, $\sigma = 1$)

(Triola, 2013, p. 251)

Nosotros estudiaremos la curva normal probabilística por lo que es necesario revisar con detenimiento la página 249 del texto base, donde se explica las condiciones que debe reunir dicha curva. Una vez que usted reviso las mencionadas condiciones lo invito a practicar por medio de la resolución de los siguientes problemas modelo.

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

PROBLEMA MODELO # 4

Resolución de ejercicios del área bajo la curva normal.

Hallar el área bajo curva normal: a) Z = — 1,20 y Z = 2,40 b) Z = 1,23 y Z = 1.87; c) Z = —2.35 y Z = —0.50.

Solución:
Para resolver estos problemas es necesario saber cómo usar la tabla de distribución normal para esto debe ingresar al enlace dado y observar detenidamente la explicación del uso de una tabla de distribución normal como la tabla 6.9 de su texto base en la página 252.
Es necesario comprender que la probabilidad es igual al área bajo la curva y que al usar la tabla 6.9 correspondiente a Áreas bajo la curva normal, ésta tomará a valores del centro a la derecha y del centro hacia la izquierda. $\boldsymbol{P_{(x)} = A}$

a. Z = – 1,20 y Z = 2,40

Revisando en la tabla para Z tenemos:

$\boldsymbol{P\,(-1.20 \leq Z \leq 2.40) = 0.3849 + 0.4918 = 0.8767}$

Tabla 6.9 áreas de una distribución normal ordinaria.

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0754
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2258	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2518	0,2549
0,7	0,2580	0,2612	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2996	0,3023	0,3051	0,3079	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4430	0,4441
1,6	0,4452	0,4463	0,4474	0,4485	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4700	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4762	0,4767
2,0	0,4773	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857

b. Z = 1,23 y Z = 1.87
Revisando en la tabla para Z tenemos:
P **(1.23 ≤ *Z* ≤ 1.87) = 0.4693 + 0.3907 = 0.078**
c. Z=-2.35 y Z=-0.50.
Revisando en la tabla para Z tenemos:
P **(—2.35 ≤ *Z* ≤ 0.50) = 0.4906 — 0.1915 = 0.299**

Nota. Se muestra el proceso para determinar el área bajo la curva normal, en diferentes casos del valor de ***Z***. Fuente: Martínez. (2019). Elaboración: Granda S. (2020).

PROBLEMA MODELO # 5

Resolución de un problema de aplicación del área bajo la curva normal.

Suponiendo que las estaturas (X) de varones de un colegio se encuentran distribuidas normalmente con media igual a 169 cm. y desviación estándar igual a 3 cm. (Emplear la tabla de áreas bajo la curva para calcular la probabilidad). a) ¿Cuál es la probabilidad de que un estudiante tenga una estatura inferior a 165 cm? b) ¿Qué porcentaje de alumnos tendrá una estatura entre 1,65 y 1,70?

Solución:
a. Tomamos los datos e incógnitas del problema ***μ = 169***, ***σ = 3*** y ***P (x < 165) = ?***
Ahora procedemos a calcular la puntuación $z = \frac{x-\mu}{\sigma}$:

$$z = \frac{165 - 169}{3} = -1.33$$

Con este valor determinamos el área bajo la curva en la tabla 6.9 de su texto base donde se observa que al valor de ***z* = —1.33** le corresponde el área de 0.4082, pero recuerde que este valor corresponde al área bajo la curva desde ***μ* = 0** hacia la izquierda por el signo negativo , pero el valor del área es el mismo que a la derecha por la simetría de la campana de Gauss, este valor lo tomo de la tabla, pero este aún no es la solución al problema, como la pregunta del problema menciona que la probabilidad buscada está por debajo de 165 cm, en este caso procedemos de la siguiente manera:

P **(x < 165) = 0.5000 + 0.4082 = 0.0918**

(*x* < 165) = 0.0918 x 100% = 9.18%

Concluimos que la probabilidad de que un estudiante tenga una estatura inferior a 165cm es de 9.18%.

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

b. ¿Qué porcentaje de alumnos tendrán una estatura entre1?65 y 1.70?

P (165< x < 170) = ?

En este caso determinamos dos puntuaciones z

$$z = \frac{165 - 169}{3} = -1.33 \rightarrow 0.4082$$

$$z = \frac{170 - 169}{3} = 0.33 \rightarrow 0.1293$$

En este caso como las áreas encontradas están al a derecha e izquierda de la **μ = 0**, debemos sumar los valores encontrados para determinar el área bajo la curva ***A* = 0.4082 + 0.1293 = 0.5375**, y finalmente la probabilidad solicitada es: ***P* (165< x < 170) = 0.5375 x 100% = 53.75%**

El porcentaje de alumnos que tienen una estatura entre 1.65 y 1.70 es del 53.75%

Nota. Se muestra la resolución de un problema de aplicación del área bajo la curva normal en la determinación del porcentaje de estaturas en un grupo de estudiantes de un colegio. Fuente: Martínez. (2019). Elaboración: Granda S. (2020)

Actividades de aprendizaje recomendadas

Hemos culminamos el estudio de la semana número tres, donde se revisó las condiciones que debe reunir la curva normal o probabilística para el análisis de la distribución normal de datos de una determinada población. Recuerde analizar su texto base, ahí encontrará las explicaciones pertinentes para la comprensión del tema de esta semana, además ejemplos de problemas resueltos y propuestos, así como la indicación de cómo aplicar el EXCEL para simplificar procesos de cálculo, finalmente luego de todo esto es necesario que proceda a realizar las siguientes actividades:

1. Analizar el recurso audiovisual denominado Distribución Normal de Probabilidades-Campana de Gauss.
2. Resolver los problemas #86 y 89 de su texto base, página 258.

3. Revisar la sección Aplicación del Excel de su texto base en las páginas 253-257.
4. Revisar la definición y características de la Distribución Normal y proceder a contestar el siguiente cuestionario según el valor de verdad que corresponda.

Condiciones de la curva normal o probabilística	V/F
La curva es simétrica	
El área bajo la curva es igual al 50%	
X toma valores de mayor a menor, es decir de derecha a izquierda	
Al estandarizar, convertir los valores de x en los valores de z, esta tendrá una media de $\boldsymbol{\mu_z = 0}$ y $\boldsymbol{\sigma_z = 1}$	

RETROALIMENTACIÓN

En el recurso audiovisual se explica las características más de la Distribución Normal de Probabilidades y demás conceptos necesarios para resolver problemas de aplicación de esta distribución que le permitirán resolver los problemas planteados y con la revisión de la aplicación del Excel podrá simplificar los procesos de cálculo.

Al contestar este cuestionario usted pondrá autoevaluar sus conocimientos sobre las condiciones de la distribución de probabilidad normal. Estas condiciones son fundamentales para resolver problemas de aplicación de la distribución normal.

Hemos finalizado el estudio de la distribución normal, ahora estamos listos para continuar con el aprendizaje.

¡Aplausos por su trabajo activo y responsable, siga adelante!

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

> ***Recuerde que:***
>
> *A través del chat de tutoría y consultas puede aclarar sus dudas e inquietudes con su docente tutor.*
>
> *Trabajar duro te llevará a la cima, disfrutar el camino te llevará más lejos.*

AUTOEVALUACIÓN PRIMERA UNIDAD

Estimado estudiante, luego de culminar con el estudio de la primera unidad **Distribuciones de Probabilidad** es recomendable que usted realice la autoevaluación de lo aprendido, con la finalidad de retroalimentar su aprendizaje.

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Autoevaluación 1

Instrucción: Marque la alternativa correcta.

1. Dentro de los criterios para satisfacer una experiencia binomial, que significa que las pruebas deben ser mutuamente excluyentes:

   a. El resultado de un ensayo no afecta el resultado del otro.
   b. Que cada una de las pruebas tengan dos resultados, favorable o desfavorable.
   c. Que la probabilidad de éxito (y la de fracaso) de un acontecimiento es fijo.

2. Al lanzar cuatro monedas, ¿cuál es la probabilidad de obtener exactamente dos caras?

   a. 50%
   b. 37.5%
   c. 75%

3. En una facultad, la probabilidad de un alumno apruebe el semestre es del 80%. La probabilidad de que como máximo 6 ganen es de:

   a. $P_{(x \leq 6)} = 49.67\%$
   b. $P_{(x \leq 6)} = 29.36\%$
   c. $P_{(x \leq 6)} = 0.1146\%$

4. Si el 3% de las bombillas fabricadas por una compañía son defectuosas, hallar la probabilidad de que, en una muestra de 100 bombillas, 2 sean defectuosas:

   a. 14.94%
   b. 22.40%
   c. 10.08%

5. Cuando utilizamos la función POISSON.DIST() en Excel para el cálculo de la Distribución de Poisson, existe un argumento de función que se denomina "Acumulado". Cuando éste toma un valor VERDADERO, la función devolverá:

   a. La probabilidad de que ocurra cuando X sea igual al parámetro establecido.
   b. La probabilidad de obtener como máximo.
   c. La probabilidad de obtener como mínimo.

6. El número de ahogados en un accidente, por año, en un país es de tres por cada 100.000 habitantes. La probabilidad de que en una ciudad cuya población es de 200.000 haya menos de 3 ahogados por año es de:

   a. 69.60%
   b. 10.33%
   c. 6.20%

7. Hallar el área bajo la curva normal entre Z=0.90 y Z=—1.85. Use la tabla de áreas bajo la curva normal de probabilidad.

   a. 0.5536
   b. 0.7837
   c. 1

8. La distribución normal presenta la siguiente característica:

   a. Simétrica.
   b. Es asintótica.
   c. El área bajo la curva, es aproximadamente del 100%.
   d. La media, se localiza en el centro de la campana.
   e. Se puede aplicar en los ejercicios de distribución Binomial, siendo su resultado un valor aproximado
   f. Todas las anteriores

9. El tiempo necesario para atender un automóvil en una estación de gasolina, tiene una distribución normal, con media de 4,5 minutos y desviación estándar de 72 segundos. ¿Cuál es la probabilidad de que un automóvil seleccionado en forma aleatoria requiera, como máximo, 3?5 minutos para el servicio?

   a. 20.33%
   b. 35.50%
   c. 17.00%

10. Relaciones los tipos de variables con su definición:

i. Variable Aleatoria	a. Esta puede asumir cualquier valor dentro de un intervalo, es infinita
ii. Variable Aleatoria discreta	b. Está conformada por factores donde interviene el azar.
iii. Variable Aleatoria continua	c. Esta puede asumir un número finito de valores que se pueden contar.

   a. i-b, ii-c, iii-a
   b. i-a, ii-b, iii-c
   c. i-c, ii-c, iii-b

Una vez terminada la autoevaluación de la primera unidad, le recomiendo verificar sus respuestas con el solucionario que se encuentra al final de esta guía didáctica, esto le permitirá validar su desempeño. Si surgen dudas en una o más preguntas vuelva al leer el contenido científico para que identifique la validez de su respuesta.

Ir al solucionario

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Le recuerdo estimado estudiante que usted dispone del horario de tutoría, donde su profesor tutor se encuentra presto a atender sus dudas e inquietudes, él ayudará a resolver las dificultades que se le presenten en el transcurso de la semana de trabajo y a validar su desempeño. Finalmente le motivo a participar de las actividades síncronas planificadas a lo largo de la asignatura, ya que estas son la oportunidad de reducir obstáculos como el tiempo y el espacio y obtener así mejores resultados en su aprendizaje.

Semana 4

En la presente semana nos aprestamos a introducirnos en el estudio de la estadística inferencial, para esto partiremos del análisis del sustento teórico fundamental para el desarrollo de las distribuciones muestrales, capítulo siete de su texto base. Cabe recalcar que la formación estadística hoy en día exige del razonamiento sobre el muestreo, ya que según señala Inzunza (2019):

> El razonamiento sobre muestreo ha adquirido especial importancia para la educación estadística en los años recientes, como consecuencia del incremento en el uso de datos provenientes de muestras y experimentos aleatorios en investigaciones de interés en las profesiones, la ciencia y la vida cotidiana. (p.204)

Esta actual exigencia nos obliga a profundizar en el estudio de las distribuciones muestrales por lo que lo invito a observar y analizar la aplicación Poblaciones, muestras y variabilidad muestral donde observará la relación que existe entre una población, las muestras que se seleccionan de ella y el tamaño de muestra. ¡Bienvenido!

Unidad 2. Distribuciones muestrales

2.1. Algunos conceptos sobre el muestreo

En esta unidad nos ocuparemos de la inferencia estadística o método inductivo, considerando que dicha inferencia es el conjunto de métodos que permite inducir a través de una muestra estadística el comportamiento de una determinada población, lo que nos permitirá extraer conclusiones sobre los parámetros de la población y el grado de fiabilidad de los resultados extraídos.

Iniciamos recordando algunos conceptos básicos, usted debe ampliar estas definiciones con la lectura comprensiva de su texto básico en las páginas 274 a 277 y de las 658-669, en estas secciones encontrará el fundamento teórico de esta unidad. Este fundamento teórico procedo a sintetizarlo en los esquemas conceptuales siguientes.

Figura 3.
Definiciones básicas.

Población o Universo: conjunto de unidades y elementos que presetan unas características en común. Estas unidades se denominan variables continuas y discretas.

Unidad: son cada uno de los valores que se han obtenido al realizar un estudio estadístico.
Elemento: es la persona o cosa a la cual pertenece (o que contiene) la información requerida por la investigación.

Muestra: es un subconjunto representativo de unidades de análisis de una población dada.

Nota. La figura explica los conceptos básicos de población, unidad, elemento y muestra fundamentales en el estudio de las distribuciones muestrales. Fuente: Martínez. (2019). Elaboración: Granda S. (2020).

> Muestreo aleatorio**. -** Es el proceso de seleccionar un conjunto de individuos de una población con el fin de estudiarlos y poder caracterizar el total de población. El muestreo es útil gracias a que podemos acompañarlo de un proceso inverso, que llamamos generalización ***de resultados***. Es decir, para conocer un universo lo que hacemos es: extraer una muestra del mismo, medir un dato u opinión, proyectar en el universo el resultado observado en la muestra. (Ochoa,2015)

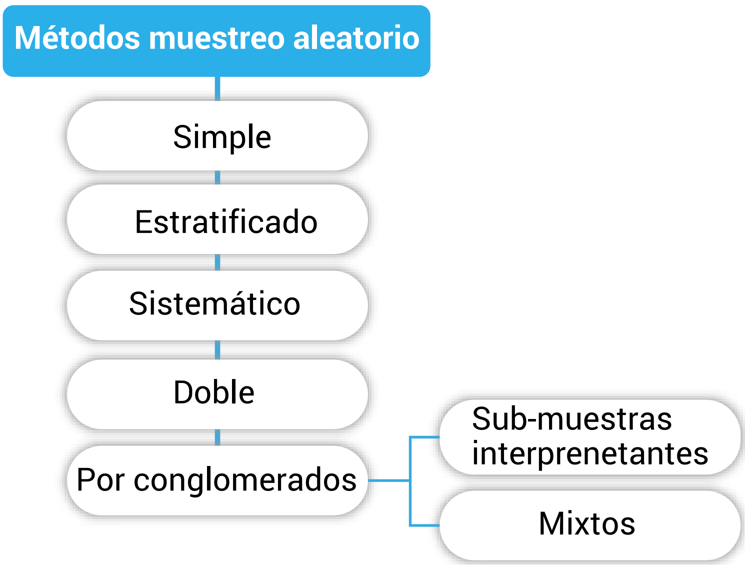
Para entender las teorías del muestreo es necesario que acceda en su texto base al Capítulo 13, en las páginas 663-666 donde se define el muestreo aleatorio, el mismo que es condensado en la siguiente figura.

Métodos del muestreo aleatorio

Para entender las teorías del muestreo es necesario que acceda en su texto base al Capítulo 13, en las páginas 663-666 donde se define el muestreo aleatorio, el mismo que es condensado en la siguiente figura. Esta figura expone los métodos de muestreo aleatorio, el mismo que sirve para generalizar conclusiones obtenidas a través de una muestra.

Métodos de muestreo aleatorio

Figura 4.
Métodos del muestreo aleatorio.



Nota. La figura señala los métodos del muestro aleatorio que sirve para generalizar conclusiones obtenidas a través de una muestra. Fuente: Martínez. (2019). Elaboración: Granda S. (2020)

Es importante considerar que también existe el ***muestreo no aleatorio*** circunstancial o errático, estos métodos proporcionan resultados o estimaciones no confiables, por estar cargados de subjetividad del investigador. Entre estos métodos tenemos:

- Muestreo a juicio, intencional u opinático
- Muestreo por conveniencia
- Muestreo voluntario
- Muestreo por cuotas

Una vez revisado el contenido teórico es momento de dar paso al análisis de la distribución muestral, tema de la siguiente sección.

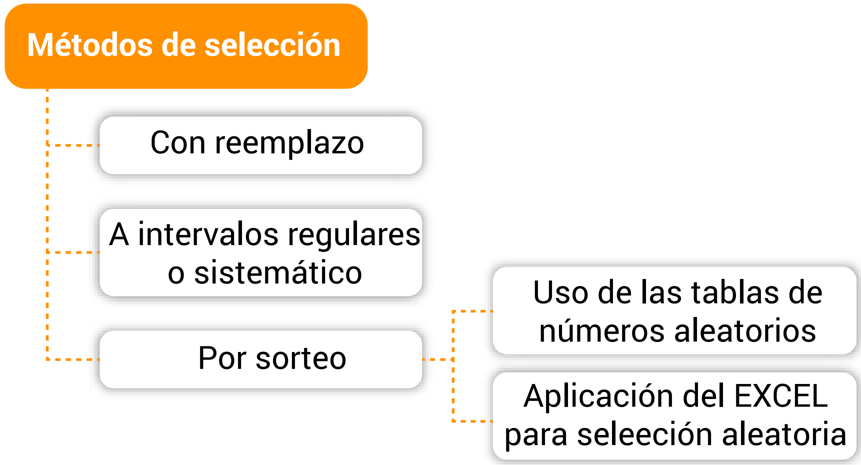
2.2. Distribución muestral

Las distribuciones muestrales se consideran importantes en el trabajo del profesional, porque proporcionan herramientas metodológicas para analizar la variabilidad, determinar relaciones entre variables, diseñar de forma óptima los experimentos y mejorar las predicciones y la toma de decisiones en situaciones de incertidumbre.

Triola (2013) afirma que "La distribución muestral de un estadístico (como una media muestral o una proporción muestral) es la distribución de todos los valores del estadístico cuando se obtienen todas las muestras posibles del mismo tamaño *n* de la misma población" (p.276).

Continuando con la explicación del tema debemos revisar los métodos de selección de unidades, según se explica en el capítulo 13 del texto base, tenemos:

Figura 5.
Métodos de selección de unidades.



Nota. La figura señala los métodos de selección de unidades al azar, esto como parte del estudio de la distribución muestral. Fuente: Martínez. (2019). Elaboración: Granda S. (2020).

Ahora nos encaminamos al estudio de las distribuciones de muestras, a fin de lograr formular conclusiones de la población de donde tomamos dichas muestras, le resultara orientador observar el objeto Distribución de la media muestral y luego analizar las secciones siguientes para continuar con el proceso de aprendizaje.

2.3. Distribución de medias muestrales: $\bar{x}$

> ***Definición***
>
> "La distribución muestral de la media es la distribución de medias muestrales, donde todas las muestras tienen el mismo tamaño ***n*** y se obtiene de la misma población" (Triola,2013, p.276).

Para le resolución de los problemas modelo de la distribución muestral de la media utilizaremos la siguiente simbología.

Figura 6.
Simbología estadística.

Medidas	Población	Muestra	
Media aritmética	μ	x	$\mu_{\bar{x}}$:medias de todas las medias muestrales.
Varianza	σ^2	S^2	$\sigma_{\bar{x}}$: desviación típica de todas las medias muestrales.
Desviación típica	σ	S	M: Número de muestras posibles
Tamaño	N	n	**Selección** Sin reposición: $M=\binom{N}{n}=\frac{N!}{(N-n)!n!}$ Con reposición: $M=N^n$

Nota. La figura muestra la simbología a emplearse en la resolución de problemas de aplicación de la distribución de medias muestrales. Fuente: Martínez. (2019). Elaboración: Granda S. (2020).

Completamos el fundamento teórico de esta unidad con la definición de la ***Teoría del límite central.***

> ***Definición***
>
> “El teorema del límite central, el cual plantea que, para una población con cualquier distribución, la distribución de las medias muestrales se aproxima a una distribución normal conforme aumenta el tamaño de la muestra” (Triola,2013, p.287).

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

El teorema o la teoría del límite central es útil para resolver problemas prácticos que requieren la utilización de dos propiedades que las explica Triola (2013):

- Valor individual: Cuando trabaje con un valor individual de la población distribuida normalmente utilice $z = \frac{x - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$
- Muestra de valores: cuando trabaje con una media de alguna muestra (o grupo) asegúrese de utilizar el valor de $\frac{\sigma}{\sqrt{n}}$ para la desviación estándar de las medias muestrales. Utilice $z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$ (p.290).

Finalmente debemos tomar en cuéntalo a Martínez (2019) quien señala que en los casos de poblaciones finitas se puede aplicar el ***factor de corrección***, el que se presenta de muchas formas que se resumen en la expresión dada a continuación, esta se empleará cuando se tenga información del tamaño poblacional y el tamaño de la muestra dada sea mayor al 5% de la población.

$$z = \frac{\bar{x} - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)\sqrt{\frac{N - n}{N - 1}}}$$

Con toda la revisión que precede vamos a desarrollar algunos problemas modelo, donde se aplica la distribución muestral.

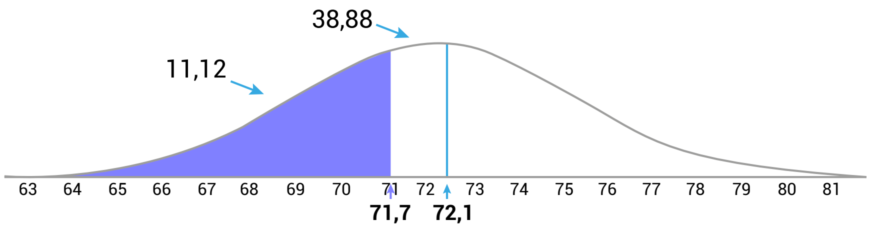
<u>PROBLEMA MODELO # 6</u>

Solución de un problema de distribución de medias muestrales.

En una población normal, con media 72,1 y desviación estándar 3,1, encuentre la probabilidad de que, en una muestra de 90 observaciones, la media sea menor que 71,7.

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Solución	
Anotamos los datos e incógnitas del problema:	Calculamos la probabilidad solicitada.
$\bar{x}$ = 71.7 **σ = 3.1** **μ = 72.1** **n = 90** **$P(\bar{x} < 71.7) = ?$**	$z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{71.7-72-1}{\frac{3.1}{\sqrt{90}}} = \frac{-0.4}{\frac{3.1}{9.48}} = \frac{-0.4}{0.33} = -1.22$ Con el valor de z = — 1,22 buscamos el área bajo la curva por medio de la Tabla 6.9 Áreas de una distribución normal ordinaria. z = —1,22 $\rightarrow A(0.3888)$ Determinamos la probabilidad requerida: P($\bar{x}$<71.7) = A(0.5000) — A(0.3888) = 0.1112 P($\bar{x}$<71.7) = 0.1112 x 100% = 11.12%



Respuesta: La probabilidad de que, en una muestra de 90 observaciones, la media sea menor que 71,7 es de 11,12%	

Nota. Se explica el proceso para determinar una probabilidad empleando distribución de medias muestrales. Fuente: Martínez. (2019). Elaboración: Granda S. (2020)

Luego de haber culminado el estudio de los temas propuestos, lo invito a desarrollar las siguientes actividades de aprendizaje.

Actividades de aprendizaje recomendadas

Al culminar la cuarta semana de trabajo y dar inicio al estudio de las distribuciones muestrales, empezando por la revisión de los conceptos básicos sobre el muestreo, la definición de la distribución muestral y la distribución de medias muestrales, es necesario recomendarle analizar su texto base en los capítulos 7 y 13, ahí encontrará las explicaciones pertinentes para la comprensión del tema de esta semana, además ejemplos de problemas resueltos y propuestos, así como la indicación de cómo aplicar el EXCEL para simplificar procesos de cálculo, todo esto como antesala a proceder a realizar las siguientes actividades:

1. Habilitar el hipervínculo y mirar atentamente el video titulado Inferencia Estadística/Distribución muestral de medias y proporciones.

2. Contestar el siguiente cuestionario según el valor de verdad que corresponda, para responder tomar como base la presente guía didáctica y su texto base capítulos 7 y 13.

Cuestionario de Conceptos distribución muestral.	V/F
Población o universo es un conjunto de unidades o elementos que presentan características diferentes.	
La unidad hace referencia a una persona, una familia, una vivienda, una manzana, un barrio, etc.	
Las características de una unidad o elemento de investigación se clasifican en cualitativas y cuantitativas,	
La investigación exhaustiva debe observar una parte de la población.	
Una muestra es no aleatoria cuando los elementos son elegidos por medio de métodos no aleatorios.	
El marco de referencia o marco muestral correspondiente a la población objetivo al total de unidades o elementos que integran la población a investigar.	

RETROALIMENTACIÓN

En este video se explican las definiciones de inferencia estadística, con algunos ejercicios que muestran la aplicación de los conceptos de distribución muestral de media. Así también al contestar el cuestionario usted pondrá autoevaluar sus conocimientos sobre la fundamentación teórica, explicada en su texto base Capítulo 7 y 13. Estas definiciones son fundamentales para resolver problemas de aplicación de distribuciones muestrales que serán el tema de estudio de las siguientes semanas.

Hemos culminado la semana de trabajo y empezamos a estudiar las distribuciones muestrales llegando hasta la distribución de medias muestrales, ahora estamos listos para continuar con el aprendizaje.

¡Felicitaciones por todo el trabajo desarrollado hasta la presente semana, esto fortalece su aprendizaje!

> ***Recuerde que:***
>
> *A través del chat de tutoría y consultas puede aclarar sus dudas e inquietudes con su docente tutor.*
>
> *Da siempre los mejor de ti. Lo que plantes ahora, lo cosecharás más arde.*
>
> *OG. Mandino*

Semana 5

En esta semana continuaremos con el estudio de las distribuciones muestrales, específicamente aplicaremos la distribución muestral de una proporción *p*, para esto continuaremos trabajando en el capítulo siete de su texto base. Debo señalar lo importante que resulta para nuestro estudio que el estudiante comprenda que al incrementar el tamaño de la muestra se acerca a estimaciones más precisas de la probabilidad y de características poblacionales, que las mediciones repetidas de un mismo fenómeno, es un proceso aleatorio que conduce a resultados distintos y que las inferencias están influenciadas por la muestra seleccionada. Inzunza (2019).

Con estas consideraciones le invito a analizar de forma crítica el desarrollo del contenido siguiente. ¡Bienvenido!

2.4. Distribución muestral de una proporción: *p*

Esta distribución se genera al igual que la distribución muestral de medias, a excepción que al extraer las muestras de la población se calcula el estadístico proporción en lugar del estadístico media. Recordemos que una proporción es la fracción, proporción relativa o porcentaje que expresa la parte de la población o muestra que tiene un atributo particular de interés.

En esta sección Martínez (2019) explica "en el análisis de una característica cualitativa o atributo, se emplea la proporción de éxitos y no el número de éxitos como en la distribución binomial" (p.287). Esto se explica en el objeto de estudio Distribución muestral de proporciones , obsérvelo detenidamente como trabajo previo al estudio de esta sección.

La proporción que emplearemos está dada por el número de elementos con el atributo en la muestra (*a*) dividido para el tamaño del a muestra (*n*)

$$p = \frac{\sum a_i}{n} = \frac{Número\ de\ \ éxitos}{Tamaño\ de\ la\ muestra}$$

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Un concepto adicional que se requiere es la ***variante estadística*** ya que se puede utilizar la distribución normal para evaluar la distribución muestral de proporciones:

$$Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} = \frac{p - \mu_p}{\sigma_{\bar{p}}}$$

Utilizaremos la siguiente simbología

$p = \frac{a}{n} = \frac{\sum a_i}{n}$	Proporción muestral ↔$\underline{p}$
$P = \frac{A}{N}$	Proporción poblacional ↔$\mu_p = \underline{P}$
$\sigma_p^2 = PQ$	Varianza de una proporción en la población
$\sigma_p = \sqrt{PQ}$	Desviación proporcional en la población
$s_p^2 = pq$	Varianza de una proporción muestral
$s_p = \sqrt{pq}$	Desviación típica proporcional en la muestra
$\sigma_{\underline{p}} = \sqrt{\frac{PQ}{n}}$	Error estándar de una proporción

Es necesario revisar su texto base en las páginas 287-289 aquí encontrará la simbología y demostración de las expresiones necesarias para determinar una distribución muestral de una proporción *p*. Así mismo le invito a revisar la resolución de los siguiente problemas modelo, donde se expone la aplicación del tema de esta semana.

PROBLEMA MODELO # 7

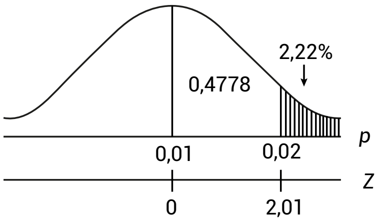
Solución de un problema de distribución muestral de una proporción.

Un fabricante de desodorantes recibe cada semana lotes de 10.000 válvulas para los tarros rociadores. Para aceptar o rechazar dichos lotes, selecciona al azar 400 válvulas de cada lote; si el 2% o más

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

resultan defectuosos, se rechaza el lote. En caso contrario se acepta el lote. ¿Cuál es la probabilidad de rechazar un lote que contenga el 1% de válvulas defectuosas?

Solución	
Anotamos los datos e incógnitas del problema:	Calculamos la probabilidad solicitada.
$\boldsymbol{P} = 0{,}01$ $n = 400$ $P_{(p>0{,}02)} = ?$	$$Z = \frac{p-P}{\sqrt{\frac{PQ}{n}}} = \frac{0{,}02 - 0{,}01}{\sqrt{\frac{0{,}01(0{,}99)}{400}}} = 2{,}01$$ Con el valor de $Z = 2{,}01$ buscamos el área bajo la curva por medio de la Tabla 6.9 Áreas de una distribución normal ordinaria. $Z = 2{,}01 \rightarrow A\ (0.4778)$ Determinamos la probabilidad requerida: $P = 0{,}5000 - 0{,}4778 = 0{,}0222$ $P_{(p>0{,}02)} = 0{,}0222 \times 100 = 2{,}22\%$



Respuesta:
La probabilidad de rechazar un lote que contenga el 1% de válvulas defectuosas es de 2,22%.

Nota. Se explica el proceso para determinar la probabilidad de rechazar un lote que contenga el 1% de válvulas defectuosas empleando distribución muestral de una proporción. Fuente: Martínez. (2019). Elaboración: Granda S. (2020).

PROBLEMA MODELO # 8

Solución de un problema de distribución muestral de una proporción.

Según registros del Departamento de Circulación y Transito, el 25% de los heridos en accidentes de tránsito quedan con alguna discapacidad de por vida. En un mes cualquiera se registran 150 personas que sufrieron lesiones. ¿Cuál es la probabilidad de que 42 o más víctimas queden con alguna incapacidad?

Solución

Anotamos los datos e incógnitas del problema:

$\boldsymbol{P} = 0{,}25$

$\boldsymbol{n} = 150$

$$p = \frac{42}{150} = 0.28$$

$P_{(p>0{,}28)} = ?$

Calculamos la probabilidad solicitada.

$$Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} = \frac{0{,}28 - 0{,}25}{\sqrt{\frac{0{,}25(0{,}75)}{150}}} = 0{,}85$$

Con el valor de $Z = 0.85$ buscamos el área bajo la curva por medio de la Tabla 6.9 Áreas de una distribución normal ordinaria.

$z = 0.85 \rightarrow A(0.3023)$

Tabla 6.9 áreas de una distribución normal ordinaria.

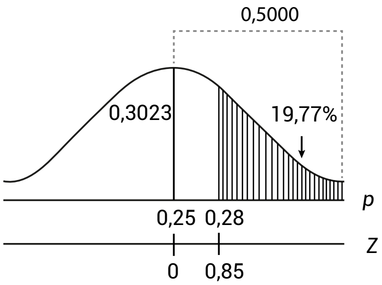
z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0754
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2258	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2518	0,2549
0,7	0,2580	0,2612	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2996	0,3023	0,3051	0,3079	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389

Determinamos la probabilidad requerida:

$$P = 0{,}5000 - 0{,}3023 = 0{,}1977$$

$$P_{(p>0{,}28)} = = 0{,}1977 \times 100 = 19{,}77\%$$

Solución



Respuesta:
La probabilidad de que 42 o más víctimas queden con alguna incapacidad es de 19,77%.

Nota. Se explica el proceso para determinar la probabilidad de que 42 o más víctimas queden con alguna incapacidad es de 19,77%. Fuente: Martínez. (2019). Elaboración: Granda S. (2020).

Una vez revisado el fundamento teórico de esta semana y analizado cada uno de los pasos de resolución de los problemas modelos dados, se propone el desarrollo de las siguientes actividades.

Actividades de aprendizaje recomendadas

Un vez terminada la semana de estudio número cinco, donde se analizó la distribución muestral de una proporción, tanto en sus características como su aplicación, recuerde analizar su texto base, donde encontrará las explicaciones pertinentes para la comprensión del tema de esta semana, además ejemplos de problemas resueltos y propuestos, así como la indicación de cómo aplicar el EXCEL para simplificar procesos de cálculo, finalmente luego de esto es necesario que proceda a realizar las siguientes actividades:

- Habilitar el hipervínculo y mirar atentamente el video titulado Inferencia Estadística/Distribución muestral de medias y proporciones.

- Estimado estudiante le recomiendo leer el texto básico correspondiente al contenido del Capítulo 7 Distribuciones muestrales. Muestreo Aleatorio, las páginas 287-290 y elabore un resumen de su contenido.

- Resuelva los problemas propuestos de la página 291 los numerales 40, 45 y50.

RETROALIMENTACIÓN

En este video se explican definiciones de inferencia estadística, con algunos ejercicios que muestran la aplicación de los conceptos de distribución muestral de proporciones. Así mismo, en el texto básico usted puede acceder a los fundamentos teóricos necesarios para la resolución de todos los ejercicios propuestos al final de la unidad.

Una vez desarrolladas las actividades recomendadas, seguro reafirmo sus conocimientos, que lo han preparado para continuar con su proceso de aprendizaje.

Felicitaciones por el arduo trabajo desarrollado durante la presente semana. ¡Adelante con el siguiente reto!

> ***Recuerde que:***
>
> *A través del chat de tutoría y consultas puede aclarar sus dudas e inquietudes con su docente tutor.*
>
> *La vida te pondrá obstáculos, pero los límites los pones tú.*

Semana 6

Continuando con la ruta del aprendizaje, en esta semana nos aprestamos a seguir profundizando en las distribuciones muestrales, en este punto es necesario detenernos a reflexionar en ¿para qué sirve la distribución muestral de medias? Pues bien, cuando se investiga lo que interesa es inferir si los hallazgos de un determinado grupo, son similares a los de la población en general, o a los de otro grupo, o bien si se trata de valores distintos, es aquí donde radica la importancia del trabajo con la distribución muestral de medias. Como señala la Triola (2013)

> Cuando en una población se toma una muestra y se mide una variable continua, se obtiene un conjunto de mediciones que puede resumirse en un valor de media. Si se toma otra muestra de la misma medición se obtendrá otra media. Puede intuirse entonces que podemos tomar infinitas muestras y obtener por lo tanto infinitas medias. Esas medias por lo tanto constituyen a su vez una variable continua, que como toda variable continua tiene determinada distribución de probabilidades.
>
> El teorema del límite central nos dice que todas las medias de una variable se distribuyen alrededor de la media de la población: la media de todas las medias es la media poblacional. Notemos que no estamos hablando ahora de datos individuales en torno de la media de una muestra: estamos hablando de medias de muestras en torno de la media poblacional (p.277).

Por lo que en esta sección nos ocuparemos de las muestras mayores que treinta donde las medias se distribuyen alrededor de la media poblacional en forma gaussiana como lo puede observar en el recurso audiovisual Diferencia de la Media en Distribuciones muestrales, en base a los antecedentes indicados estimado estudiante, lo invito a seguir explorando este interesante tema. ¡Bienvenido!

2.5. Distribución de diferencias entre las medias muestrales: $\bar{x} - \bar{y}$

Ahora analizaremos el caso donde se tiene dos poblaciones normales e independientes, que las vamos a identificar con X **y** Y con tamaños N_1 y N_2 respectivamente, con medias μ_x y μ_y con sus correspondientes desviaciones típicas σ_x y σ_y, de esto se obtendrá un número (M) de pares de muestras posibles.

A partir de esto se determina:

- La media aritmética de la diferencia de las medias.

$$\mu_{\bar{x}-\bar{y}} = \frac{\sum \bar{x}_i}{M} - \frac{\sum \bar{y}_i}{M}$$

- La desviación típica de las diferencias entre los pares de medias muestrales.

$$\sigma_{\bar{x}-\bar{y}} = \sqrt{\frac{\sum \left[(\bar{x}_i - \bar{y}_i) - (\mu_x - \mu_y)\right]^2}{M}}$$

- Error estándar de las diferencias entre las medias muestrales.

$$\sigma_{\bar{x}-\bar{y}} = \sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}}$$

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

- La variable estadística Z si la diferencia entre las medias muestrales se asemeja a la distribución normal.

$$Z = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}}}$$

Para completar esta información es necesario revisar las páginas 293-295 del texto base, donde encontrará la simbología que se requiere para determinar una distribución de diferencias entre dos medias muestrales:$x - y$. Así mismo le invito a revisar la resolución del siguiente problema modelo, donde se expone la aplicación de la distribución muestral.

PROBLEMA MODELO #9

Solución de un problema de distribución de diferencias entre dos medias muestrales.

En promedio, los estudiantes de la Universidad A se levantan 50 minutos después de la salida del sol, con una desviación estándar de 15 minutos. Los estudiantes de la Universidad B se levantan 60 minutos después de la salida del sol, con una desviación estándar de 18 minutos. Un grupo de 25 estudiantes de la Universidad A realiza un viaje junto con 20 alumnos de la B. Encontrar la probabilidad de que la hora media de levantada del grupo de la Universidad B sea más temprana que la del grupo de la Universidad A.

Solución	
Anotamos los datos e incógnitas del problema:	Calculamos la probabilidad solicitada.

Solución		
$\mu_x = 50$	$\mu_y = 60$	Vamos a determinar la variable estadística Z $$Z = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}}}$$
$\sigma_x = 15$	$\sigma_y = 18$	
$n_1 = 25$	$n_2 = 20$	**Medias de muestras**
$P_{(x-y<0)} = ?$		$(\bar{x} - \bar{y}) = 0 \rightarrow$ ya que no nos dan un valor especifico, nos indican únicamente que la hora media de levantada "sea más temprana" en este caso tomamos el cero. $(\bar{x} - \bar{y}) < 0$ es menor debido a que el grupo B tiende a ser menor en tiempo, pero más temprano. **Medias poblacionales** $(\mu_x - \mu_y) = 50 - 60 = -10$ Entonces $$Z = \frac{(0) - (-10)}{\sqrt{\frac{15^2}{25} + \frac{18^2}{20}}} = 1.99$$ Con el valor de Z = 2,01 buscamos el área bajo la curva por medio de la Tabla 6.9 Áreas de una distribución normal ordinaria. $z = 1.99 \rightarrow A(0.4767)$ Determinamos la probabilidad requerida: P = 0,5000 – 0,4767 = 0,0233 $P_{(x-y<0)}$ = 0,0233 x 100% = 2.33%
Respuesta: La probabilidad de que la hora media de levantada del grupo de la Universidad B sea más temprana que la del grupo de la Universidad A es de 2.33%		

Nota. Se explica el proceso para determinar probabilidad de que la hora media de levantada del grupo de la Universidad B sea más temprana que la del grupo de la Universidad A empleando distribución de diferencias entre dos medias muestrales. Fuente: Martínez. (2019). Elaboración: Granda S. (2020).

Seguimos estudiando la distribución muestral específicamente la distribución de diferencias entre dos proporciones muestrales por lo que lo luego de revisar atentamente su texto base en las páginas 299-302, observar el objeto Distribución de la diferencia

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

de proporciones muestrales donde se explica en que ocasiones las distribuciones muestrales se pueden aproximar por normales y dada la diferencia de proporciones calcular la distribución muestral, todo este contenido se complementa con el desarrollo de la siguiente sección. ¡Bienvenido!

2.6. Distribución de diferencias entre dos proporciones muestrales $P_1 - P_2$

Muchas aplicaciones involucran poblaciones de datos cualitativos que deben compararse utilizando proporciones o porcentajes, si estas poblaciones son independientes, de diferente tamaño y distribuidas binomialmente se puede determinar su distribución muestral empleando la siguiente ecuación:

$$Z = \frac{(p_1 - p_2) - (\mu_{p_1} - \mu_{p_2})}{\sqrt{\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}}}$$

Donde:

Z = Área de la distribución normal

$(p_1 - p_2)$ = diferencia deseada

P_1 = Probabilidad de éxito de la muestra 1

Q_1 = Probabilidad de fracaso de la muestra 1

P_2 = Probabilidad de éxito de la muestra 2

Q_2 = Probabilidad de fracaso de la muestra 2

n_1 = Tamaño de la muestra 1

n_2 = Tamaño de la muestra 2

Estimado estudiante es necesario que revise su texto base en las páginas 299-200 donde se explica cómo determinar el error estándar de la diferencia entre dos medias proporcionales. Luego de esta revisión es necesario resolver un problema modelo donde se expone la aplicación de la distribución de diferencias entre dos proporciones muestrales.

PROBLEMA MODELO #10

Solución de un problema de distribución de diferencias entre dos proporciones muestrales.

Ciertas encuestas realizadas en una ciudad de la costa, revelan que el 25% de los hombres y el 33% de las mujeres escuchan cierto programa radial. ¿Cuál es la probabilidad de que en dos muestras de 150 hombres y 100 mujeres respectivamente, domiciliados en dicha ciudad, se encuentre que la proporción de mujeres que escuchan el programa sea menor o igual a la proporción de hombres?

Solución	
Anotamos los datos e incógnitas del problema:	Calculamos la probabilidad solicitada.

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Solución		
P_1=25%	P_2=33%	Vamos a determinar la variable estadística Z
Q_1=75%	Q_2=67%	$$Z = \frac{(p_1 - p_2) - (\mu_{p_1} - \mu_{p_2})}{\sqrt{\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}}}$$
n_1=150	n_2=100	
hombres	mujeres	
$P_{(P_1 - P_2 > 0)} = ?$		**Proporciones muestrales** $(p_1 - p_2) = 0 \rightarrow$ ya que no se da una diferencia específica, únicamente se solicita encontrar la proporción de mujeres que escuchan el programa sea menor o igual a la proporción de los hombres en este caso tomamos el cero. $(P_1 - P_2) < 0$ según lo indica el problema **Medias muestrales** $(P_1 - P_2) = 0.25 - 0.33 = -0.08$ Entonces $Z = \frac{(0) - (-0.08)}{\sqrt{\frac{(0.25)(0.75)}{150} + \frac{(0.33)(0.67)}{100}}} = 1.36$ Con el valor de Z = 1.36 buscamos el área bajo la curva por medio de la Tabla 6.9 Áreas de una distribución normal ordinaria. $z = 1.36 \rightarrow A\ (0.4131)$ Determinamos la probabilidad requerida: $P = 0.5000 - 0.4131 = 0.0869$ $P_{(P_1 - P_2)} > 0 = 0{,}0869 \times 100\% = 8.69\%$
Respuesta: la proporción de mujeres que escuchan el programa sea menor o igual a la proporción de hombres de 8.69%		

Nota. Se explica el proceso para determinar la proporción de mujeres que escuchan el programa sea menor o igual a la proporción de hombres empleando distribución de diferencias entre dos proporciones muestrales. Fuente: Martínez. (2019). Elaboración: Granda S. (2020)

Una vez revisado el fundamento teórico de esta semana y analizado paso a paso los problemas modelo dados, es momento de desarrollar las siguientes actividades.

Actividades de aprendizaje recomendadas

Una vez terminada la semana de estudio número seis, donde se analizó la distribución de diferencias entre dos medias muestrales y dos proporciones muestrales, en su aplicación práctica, recuerde analizar su texto base, donde encontrará las explicaciones pertinentes para la comprensión del tema de esta semana, además ejemplos de problemas resueltos y propuestos de cada tema tratado y los ejercicios misceláneos con la síntesis respectivo del capítulo.

- Habilitar el hipervínculo y mirar atentamente el video Distribución muestral de diferencias de medias.
- Habilitar el hipervínculo y mirar atentamente el video Distribución de diferencia entre dos proporciones.
- Leer su texto básico en las secciones correspondientes al contenido del Capítulo 7 Distribuciones muestrales: Distribución de diferencias entre las medias muestrales y diferencias entre dos proporciones muestrales.
- Elaborar un cuestionario que involucre preguntas relacionadas con el contenido teórico de esta semana y la resolución de los problemas propuestos los numerales 82-85 de la página 302 de su texto base.

RETROALIMENTACIÓN

En los recursos audiovisuales se explican las definiciones de inferencia estadística estudiadas en esta semana, con algunos ejercicios que exponen los conceptos de distribución de diferencias entre las medias muestrales y distribución de diferencias entre dos proporciones muestrales. Así mismo, en el texto básico usted puede acceder a los fundamentos teóricos necesarios para la resolución de todos los problemas propuestos al final de la unidad.

Una vez desarrolladas las actividades recomendadas, usted está preparado para continuar con su proceso de aprendizaje no sin antes realizar su autoevaluación del aprendizaje de la unidad culminada. Recuerde la unidad de estudio comprende las distribuciones muestrales.

¡En hora buena, todo su trabajo responsable le ha permito culminar la sexta semana, continúe con el mismo ahínco!

> ***Recuerde que:***
>
> *A través del chat de tutoría y consultas puede aclarar sus dudas e inquietudes con su docente tutor.*
>
> *El 80% del éxito se basa simplemente en insistir.*
>
> *Woody Allen*

AUTOEVALUACIÓN SEGUNDA UNIDAD

Estimado estudiante, luego de culminar con el estudio de la segunda unidad **Distribuciones Muestrales**, es recomendable realice la autoevaluación de lo aprendido, con la finalidad de retroalimentar sus inquietudes. Recuerde que la unidad dos de la presente guía corresponde al capítulo siete de su texto base.

Autoevaluación 2

Instrucción: Marque la alternativa correcta.

1. Relacionar las definiciones dadas

I.	Población o Universo	a.	Es un subconjunto representativo de unidades de análisis de una población dada
II.	Unidad	b.	Conjunto de unidades y elementos que presentan unas características comunes. Estas unidades se denominan variables continuas y discretas.
III.	Muestra	c.	Son cada uno de los valores que se han obtenido al realizar un estudio estadístico.

2. Relacionar las definiciones dadas

I. En este muestreo se da igual oportunidad de selección a cada elemento o unidad dentro de la población.

II. Asignación igual, proporcional y óptimo, garantiza la representatividad, reduciendo el error de la muestra al formar grupos o subpoblaciones más o menos homogéneas, en cuanto a su composición interna y heterogénea cuando se comparan entre sí.

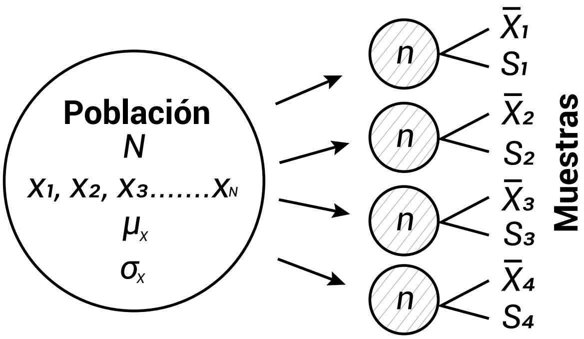
III. La selección de las unidades se hace a intervalos regulares, en un orden sistemático.

a. Muestreo aleatorio estratificado

b. Muestreo sistemático

c. Muestreo aleatorio simple o muestreo aleatorio irrestricto.

3. Dada una población, si extraemos todas las muestras posibles de un mismo tamaño, entonces la media de la distribución de todas las medias muestrales posibles, será igual a la:



A. Media poblacional.
B. Media muestral
C. Media proporcional

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

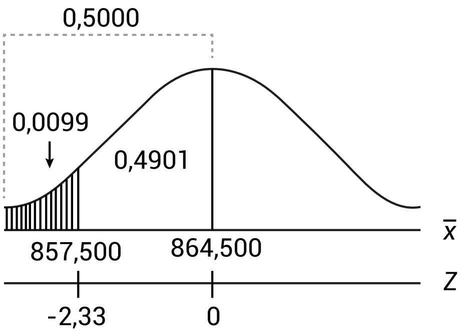
4. La teoría del límite central se cumple cuando independientemente de la población origen, la distribución de las medias aleatorias se aproximan a una distribución normal a medida que el tamaño de la muestra crece. ES decir que, si las muestras provienen de una población que no es normal, es de importancia tener en cuenta el tamaño de la muestra, si el tamaño muestral es pequeño, la distribución obtenida con sus medias muestrales tendrán un comportamiento similar al de la población de donde se extrajeron. Por el contrario:

   A. Si el tamaño muestral es grande, el comportamiento de estas medias muestrales será igual al de una distribución normal, independientemente de la población de donde fueron extraídas.

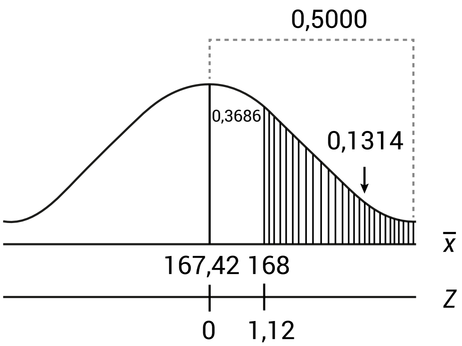
   B. Si el tamaño muestral es pequeño, el comportamiento de estas medias muestrales será igual al de una distribución normal, independientemente de la población de donde fueron extraídas.

   C. Si el tamaño muestral es grande, el comportamiento de estas medias muestrales será diferente al de una distribución normal, independientemente de la población de donde fueron extraídas.

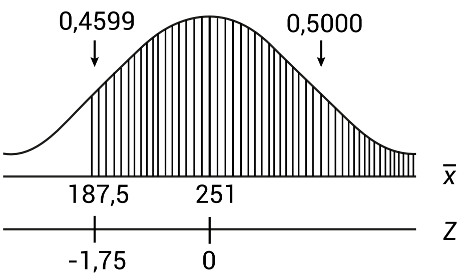
5. En cierta región los salarios diarios de los mineros del carbón están distribuidos normalmente con una media de $864.500 y una desviación estándar de $15.000. Indicar el área que determina la probabilidad de que una muestra representativa de 25 mineros, tenga un promedio diario inferior a $857.500



A.



B.



C.

6. Se sabe que el 25% de los estudiantes de un colegio usan anteojos. ¿Cuál es la probabilidad de que 8 o menos usen anteojos en una muestra de 36 estudiantes?

A. $P_{(p < 0,22)} = 33,36\%$
B. $P_{(p > 0,22)} = 33,36\%$
C. $P_{(p = 0,22)} = 33,36\%$

7. Para determinar la distribución de diferencias entre las medias muestrales $\bar{x} - \bar{y}$ se requiere determinar la desviación típica de las diferencias entre los pares de medias muestrales, para este efecto se utilizará la expresión:

A. $\sigma_{\bar{x}-\bar{y}} = \sqrt{\frac{\sum \left[(\bar{x}_i - \bar{y}_i) - (\mu_x - \mu_y)\right]^2}{M}}$

B. $\sigma_{\bar{x}-\bar{y}} = \sqrt{\frac{\sum \left[(\bar{x}_i - \bar{y}_i) + (\mu_x - \mu_y)\right]^2}{M}}$

C. $\sigma_{\bar{x}-\bar{y}} = \sqrt{\frac{\sum \left[(\bar{x}_i - \bar{y}_i)(\mu_x - \mu_y)\right]^2}{M}}$

8. Muchas aplicaciones involucran poblaciones de datos cualitativos que deben compararse utilizando proporciones o porcentajes, si estas poblaciones son independientes, de diferente tamaño y distribuidas binomialmente se puede determinar su distribución muestral empleando el estadístico Z:

A. $Z = \frac{(p_1 - p_2) + (\mu_{p_1} - \mu_{p_2})}{\sqrt{\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}}}$

B. $Z = \frac{(p_1 - p_2) - (\mu_{p_1} - \mu_{p_2})}{\sqrt{\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}}}$

C. $Z = \frac{(p_1 - p_2)(\mu_{p_1} - \mu_{p_2})}{\sqrt{\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}}}$

9. El nivel de confianza tiene relación directa con el tamaño de la muestra, por lo tanto, se dirá que a mayor nivel de confianza más grande debe ser el tamaño de la muestra así mismo los valores de Z se obtienen mediante:

    A. El uso de tablas
    B. Aleatoriamente
    C. La experiencia del investigador

10. La fórmula que se utiliza para el cálculo de *n* en poblaciones infinitas en la variable es:

    A. $n = \left(\frac{Z\sigma}{E}\right)^2$
    B. $n = \left(\frac{Z+\sigma}{E}\right)^2$
    C. $n = \left(\frac{Z\sigma}{E}\right)$

Una vez terminada la autoevaluación de la segunda unidad, le recomiendo verificar sus respuestas con el solucionario que se encuentra al final de esta guía didáctica, esto le permitirá validar su desempeño. Si surgen dudas en una o más preguntas vuelva al leer el contenido científico para que identifique la validez de su respuesta.

Ir al solucionario

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Le recuerdo estimado estudiante que usted dispone del horario de tutoría, donde su profesor tutor se encuentra presto a atender sus dudas e inquietudes, él ayudará a resolver las dificultades que se le presenten en el transcurso de la semana de trabajo y a validar su desempeño. Finalmente le motivo a participar de las actividades síncronas planificadas a lo largo de la asignatura, ya que estas son la oportunidad de reducir obstáculos como el tiempo y el espacio y obtener así mejores resultados en su aprendizaje.

Resultado de aprendizaje 2

Utiliza las pruebas de hipótesis y las pruebas estadísticas no paramétricas en el estudio de casos.

Semana 7

Luego de estudiar las distribuciones muestrales, debemos continuar con las pruebas de hipótesis ya que otra manera de hacer inferencia es haciendo una afirmación acerca del valor que el parámetro de la población bajo estudio puede tomar, es este el fundamento del segundo resultado de aprendizaje de nuestra asignatura, el que busca que usted logre utilizar las pruebas de hipótesis y las pruebas estadísticas no paramétricas en el estudio de casos. Ya que como bien lo explican Inzunza y Vidal (2013).

Las pruebas de hipótesis se constituyen en uno de los principales métodos de inferencia estadística, según explica una prueba de hipótesis, es un método que permite verificar una aseveración acerca del valor de un parámetro poblacional; dado que los datos son proporcionados por la muestra, los resultados pueden estar sujetos a variaciones aleatorias, por lo que una prueba de hipótesis

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

permite decidir si pequeñas desviaciones observadas respecto al resultado que idealmente debería haber ocurrido según nuestra hipótesis, son atribuibles al azar o efectivamente los resultados no se corresponden con la hipótesis que se ha planeado sobre el valor del parámetro. (p.181)

El primer paso para comprender esta unidad es la interiorización de los conceptos que fundamentan esta teoría, por lo que le invito a escuchar el objeto CONCEPTOS que le permitirá comprender argumentos estadísticos de moderada complejidad en resultados de diversos tipos de estudios, así mismo le presento una explicación de un Test de hipótesis que lo invito es escuchar comprensivamente y luego a seguir estudiando con ahínco las secciones siguientes. ¡Bienvenido!

Unidad 3. Pruebas de hipótesis

3.1. Conceptos generales

Para la comprensión de este tema revisemos las definiciones de **hipótesis y prueba de hipótesis** a las que Triola (2013) explica que "En estadística, una hipótesis es una afirmación o aseveración acerca de una propiedad de una población. Una prueba de hipótesis (o prueba de significación) es un procedimiento para someter a prueba una afirmación acerca de una propiedad de una población" (p.392).

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Le recomiendo revisar el contenido de la página 324-326 del texto base donde se da una explicación clara y detallada con ejemplos sobre la definición de la prueba de hipótesis y su clasificación la que se explicará tomando en cuenta lo que explica Triola (2013)

> La **hipótesis nula** (denotada con H_0) es la afirmación de que el valor de un parámetro poblacional (como una proporción, media o desviación estándar) es igual a un valor establecido.
>
> La **hipótesis alternativa** (denotada con H_a) es la afirmación de que el parámetro tiene un valor que, de alguna manera, difiere de la hipótesis nula (p.395).

Así mismo es necesario explicar el tipo de error que se puede dar al decidir el aceptar o rechazar una hipótesis.

- Aceptar la hipótesis nula cuando se debió rechazar.
- Rechazar la hipótesis nula cuando se debió aceptar.

Considere las decisiones en cuanto a los tipos de error que señala Triola (2019)

- Si se acepta una hipótesis verdadera, la decisión es correcta.
- Si se acepta una hipótesis falsa, cometemos el error de rechazar la hipótesis nula cuando se debió aceptar.
- Si rechazamos una hipótesis verdadera, cometemos el error de aceptar la hipótesis nula cundo se debió rechazar.
- Si rechazamos una hipótesis falsa, la decisión es correcta.

3.2. Prueba unilateral y bilateral

En su texto base usted encontrará la explicación de hipótesis unilateral, bilateral, nivel de significación y puntos críticos, debe revisar la página 328 donde Martínez (2019) expone que:

> La prueba de **hipótesis unilateral** es aquella en la cual la zona de rechazo o zona crítica está completamente comprendida en uno de los extremos de la distribución.
>
> En el caso de que la prueba comprenda áreas o zonas de rechazo en ambos extremos de la distribución, se dice que la **prueba es bilateral** o sea que la hipótesis alternativa es diferente.
>
> Se entiende por **nivel de significación**, la máxima probabilidad de que se especifique con el fin de hacer mínimo el primer tipo de error.
>
> El valor del **nivel de significación** corresponde a aún área bajo la curva de probabilidad o normal, denominada región crítica o zona de rechazo.

Finalmente, en la página 329 de su texto base se indica el procedimiento a seguir en las pruebas de hipótesis por medio de una síntesis de los pasos a seguir, los que tienen su ampliación respectiva para la comprensión de dicho proceso, el que será consolidado con los respectivos ejemplos.

3.3. Distribución de medias muestrales ($\overline{x}$)

A continuación, se expondrá la resolución de problemas modelo de una distribución de medias muestrales, estos ejemplos consideran el conocimiento o desconocimiento de la varianza poblacional. Recuerde los pasos del procedimiento a seguir en este caso son:

1. Formular la hipótesis nula y la alternativa.
2. Seleccionar el nivel de significación
3. Conocer la varianza.
4. Determinar la técnica y la prueba estadística.
5. Determinar los valores críticos y sus regiones de rechazo.
6. Calcular los datos muestrales, utilizando las fórmulas correspondientes.
7. Tomar la decisión estadística, de aceptar o rechazar.

A continuación, vamos a practicar lo aprendido en la resolución del siguiente problema modelo, por lo que le invito a seguir paso a paso el desarrollo del mismo.

PROBLEMA MODELO #11

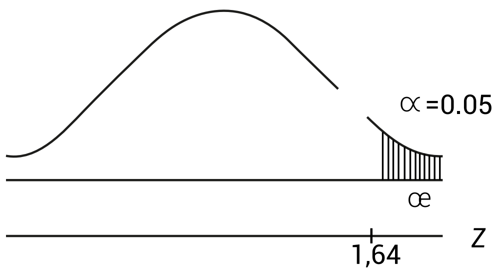
Solución de un problema de prueba de hipótesis.

Muchos años de experiencia en un examen de ingreso a la universidad en inglés, arroja una calificación promedio de 64, con una desviación estándar de 8. Todos los estudiantes de cierta ciudad, en la cual existen 64, han obtenido una calificación promedio de 68. ¿Puede tenerse la certeza de que los estudiantes de esta ciudad son superiores en inglés?

Solución	
Anotamos los datos e incógnitas del problema:	Calculamos la probabilidad solicitada.

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Solución	
$\mu = 64$	Hipótesis nula $H_0: \mu = 64$
$\sigma = 8$	Hipótesis alternativa $H_0: \mu > 64$ Dócima unilateral a la derecha
$n = 64$	Nivel de significación α=0.05
$\bar{x} = 68$	Variante estadística
	Distribución de medias muestrales $z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$
	$z = \frac{68-64}{\frac{8}{\sqrt{64}}} = 4$
	La región crítica de la dócima unilateral a la derecha Z_s<1,64



Finalmente Z=4 se ubica en la zona de rechazo por lo tanto puede tenerse la certeza, con un nivel de significación del 5% que los estudiantes de esta ciudad son superiores en inglés.

Nota. Se explica el proceso para probar que los estudiantes de la ciudad son superiores en inglés. Fuente: Martínez. (2019). Elaboración: Granda S. (2020).

Actividades de aprendizaje recomendadas

Una vez terminada la semana de estudio número siete, donde se dio inicio al análisis de las pruebas de hipótesis solicito analizar su texto básico, esta guía y demás recursos propuestos, en vista de la necesidad de que usted como estudiante y futuro profesional comprenda la integración y relación que guardan entre sí en el proceso de prueba los conceptos estadísticos. Para esto se recomienda apoyar el proceso con el desarrollo de las siguientes actividades.

- Observar cuidadosamente el recurso audiovisual Prueba de Hipótesis para la media.
- Interactuar con la aplicación Contraste de Hipótesis para la media.
- Leer su texto básico en las secciones correspondientes al contenido del Capítulo 8 Pruebas de hipótesis, aquí usted encontrará el fundamento teórico y el desarrollo de algunos problemas resueltos que le guiaran en su aprendizaje y práctica de esta unidad.
- Resolver los ejercicios 1 y 2 de la página 336.

RETROALIMENTACIÓN

En el recurso audiovisual recomendado se explica la prueba de hipótesis para la media, con una síntesis de los parámetros a tomar en cuenta para poder aplicar el contenido de la unidad. En la aplicación de GeoGebra usted podrá interactuar al variar los parámetros de un contraste de hipótesis. Así mismo en el texto básico usted puede acceder a los fundamentos teóricos necesarios para la resolución de todos los problemas propuestos al final de la unidad.

Una vez desarrolladas las actividades recomendadas, usted está preparado para culminar sus actividades del bimestre.

¡Felicitaciones por su entrega a cada una de las actividades planteadas, todo esto lo ha ubicado, en este importante punto de la ruta del aprendizaje!

> ***Recuerde que:***
>
> *A través del chat de tutoría y consultas puede aclarar sus dudas e inquietudes con su docente tutor.*
>
> *Todos nuestros sueños se pueden hacer realidad si tenemos el coraje de perseguirlos*
>
> *Walt Disney*

Actividades finales del bimestre

Semana 8

Le recomiendo llevar a cabo las siguientes actividades, como trabajo previo a rendir la evaluación presencial del primer bimestre.

Actividad 1. Recupere las actividades de aprendizaje calificadas desarrolladas durante el primer bimestre.

Actividad 2. Organice un documento que contenga dicha información cuidando la presentación; si lo hizo a mano, puede fotografiar y organizar un documento en PDF.

Actividad 3. Revise los contenidos del primer bimestre y participe de la evaluación presencial.

Para ello considere:

- Su diario de notas.
- Actividades de aprendizaje recomendadas.

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

- Actividades de aprendizaje calificadas.
- Actividades de aprendizaje interactivas; y,
- Evaluaciones parciales.

Actividades de aprendizaje recomendadas

Además de las actividades anteriores, para consolidar sus aprendizajes del primer bimestre, y prepararse a la prueba presencial en físico o virtual, sugiero:

- Revise cada uno de los conceptos estudiados en las dos unidades planificadas y desarrolladas en este primer bimestre.

- Realice suficientes ejercicios y problemas de aplicación de los diferentes conceptos, propiedades y leyes, de cada una de las unidades estudiadas, desarrollando los problemas propuestos al final de cada unidad del texto básico.

- Para cada una de las unidades, es valedero que sistematice el conocimiento aprendido, a través de la construcción de un mapa conceptual o algún otro organizador gráfico que estime conveniente.

> ***Recuerde que:***
>
> *A través del chat de tutoría y consultas puede aclarar sus dudas e inquietudes con su docente tutor.*
>
> *Nuestra mayor debilidad radica en renunciar. La forma más segura de tener éxito es intentarlo una vez más,*
>
> *Thomas Edison*

Segundo bimestre

Resultado de aprendizaje 2

Utiliza las pruebas de hipótesis y las pruebas estadísticas no paramétricas en el estudio de casos.

Para alcanzar el resultado de aprendizaje propuesto, es necesario tomar en cuenta que uno de los propósitos de la estadística es proporcionar un proceso metódico para obtener conclusiones válidas de una muestras con respecto a la población, proceso que verificaremos por medio de la solución y análisis de problemas de aplicación, que le permitirá a usted profesional en formación en un futuro administrar procesos de aprendizaje, fomentando cambios por medio de la innovación en el desarrollo del pensamiento lógico, crítico y creativo de sus educandos.

Contenidos, recursos y actividades de aprendizaje

Semana 9

Haciendo un paréntesis en el estudio de la prueba de hipótesis, en la presenta semana se hará una introducción al estudio del paquete estadístico SPSS, como una herramienta para el manejo de datos estadísticos y sus procesos analíticos para su análisis y la toma de decisiones pertinentes, motivo por el cual le motivo a continuar estudiando con ahínco y perseverancia.

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

3.4. Paquete estadístico SPSS Statistical Package for the Social Sciences / Paquete estadístico para ciencias sociales

Esta unidad se propone invitarle a usted estimado estudiante a conocer y utilizar el programa estadístico SPSS, que le servirá como una herramienta didáctica-pedagógica en su formación profesional, de modo que pueda experimentar y realizar análisis adecuados para la investigación educativa con diversas técnicas estadísticas.

Esta propuesta es básica ya que se confía en el ímpetu indagador de cada estudiante para continuar en el estudio avanzado del SPSS.

3.4.1. Introducción al SPSS

El SPSS es uno de los programas estadísticos con mayor demanda en los EE UU de Norteamérica y América Latina, ofrece diversas posibilidades para crear vínculos con otros programas comunes tales como Microsoft Word, Microsoft Excel, y Microsoft Power Point. Finalmente, SPSS permite manejar bancos de datos de gran magnitud y también efectuar análisis estadísticos muy complejos.

Dentro de los usos potenciales del programa SPSS están:

- Desarrollar un archivo de datos en una forma estructurada y también organizar una base de datos que puede ser analizada con diversas técnicas estadísticas.
- Sujetar y analizar los datos sin necesidad de depender de otros programas.
- Transformar un banco de datos creado en Microsoft Excel en una base de datos SPSS.

El SPSS facilita análisis estadísticos básicos y avanzados según el requerimiento del usuario.

Finalmente es importante mencionar que si el usuario no tiene experiencia previa utilizando SPSS o si sus conocimientos de estadística no están actualizados la navegación por el programa podría dificultarse pese a su aparente facilidad. Por otro lado, está el nivel excesivo de información que arrojan los reportes de resultados que a veces pueden confundir al usuario.

> Para comprender su uso le recomiendo revisar el CURSO DE SPSS STATISTICS-COMPLETO, donde se explica las principales funcionalidades y la modalidad de trabajo del programa, en su versión 24.
>
> Luego de esta familiarización con el programa SPSS usted tendrá acceso al Laboratorio Virtual de la UTPL donde podrá acceder al programa y empezar a conocerlo y a utilizarlo, para esto recibirá un Manual de acceso a los Laboratorios Virtuales de la UTPL.

La información de acceso será ubicada en los anuncios académicos de la asignatura con la pertinencia de la planificación y llegará a usted por medio de la mensajería de parte de los responsables de dicho proceso.

Ahora se muestra un ejemplo de resolución de un ejercicio con la aplicación del SPSS.

EJEMPLO RESUELTO

EJERCICIO 1. Coparemos dos muestras aleatorias de 10 hombres y 10 mujeres en un test que mide su autoestima (escala cuantitativa de 0 a 10 puntos).

a. ¿Podemos afirmar que ambas muestras difieren significativamente en autoestima?

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

b. ¿Podemos afirmar que la autoestima de los hombres es significativamente mayor que la de las mujeres?
c. Resuelve la pregunta a) por medio de la prueba no paramétrica adecuada.

HOMBRES	8	7	6	8	7	5	6	4	9	9
MUJERES	8	6	5	6	5	4	4	4	6	4

SPSS: Analizar –comparar medias-Prueba T para muestras independientes

a. Se trata de comparar las medias de hombres y de mujeres (6.9 y 5.2, respectivamente) con una prueba t para muestras independientes (contraste bilateral o de dos colas): el SPSS nos da t (18) =2.53, p=0.021, luego la respuesta es sí.

b. Igual que en a) sólo cambia aquí que el hecho de que el contraste es ahora unilateral (una cola). En este caso sólo hay que comprobar que las medias van en la dirección que establece la pregunta y dividir la sig. que nos da el programa por 2. Luego quedaría así: t (18) =2.53, p=0.0105, siendo la respuesta también que sí.

c. (SPSS: Analizar > Pruebas no paramétricas > Cuadros de diálogo antiguos > 2 muestras independientes). Debemos aplicar la prueba de Mann-Whitney que nos da z =2,235, p= 0.029, luego los resultados no cambian.

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

RESULTADOS SPSS

Prueba t

Estadísticos de grupo

i	N	Media	Desviación típ.	Error típ. de la media
Vd	10	6.90	1,663	0,526
Varones	10	5.20	1,317	0,416
mujeres				

Prueba de muestras independientes

	Prueba de Leven para la igualdad de varianzas		Prueba T para igualdad de medias		
	F	Sig.	t	gl	Sig. (bilateral)
Vd	0,576	0,458	2,534	18	0,021
Se han asumido varianzas iguales.			2,534	17,098	0,021
No se han asumido varianzas iguales					

Actividades de aprendizaje recomendadas

- Leer comprensivamente el Manual de acceso a los Laboratorios Virtuales de la UTPL.
- Ingresar al laboratorio virtual de la UTPL y acceder al programa SPSS.
- Revisar los ejemplos que se encuentran dentro del programa
- Resolver el siguiente problema:

Medimos la capacidad lectoescritura de 10 niños disléxicos a través de un cuestionario (escala 0 a 100 puntos) antes y después de recibir una terapia. Sus resultados fueron:

ANTES	70	72	80	75	77	80	74	81	76	73
DESPUES	74	73	84	75	84	95	88	86	80	79

a. ¿Ha aumentado la capacidad lectoescritura de los niños tras el tratamiento?

b. Resuelve la pregunta anterior por medio de la prueba no paramétrica adecuada.

RETROALIMENTACIÓN

El cumplimiento de las actividades recomendadas en esta semana le permitirá empezar su aprendizaje del uso del programa SPSS, para esto contará con el acceso al laboratorio virtual de UTPL, así mismo la resolución del problema planteado le permitirá practicar el manejo del programa en cálculos estadísticos, así como la revisión de los problemas que están dentro del mismo programa.

¡El cumplimiento de las actividades planteadas hasta este punto, es digno de felicitar!

> ***Recuerde que:***
>
> *A través del chat de tutoría y consultas puede aclarar sus dudas e inquietudes con su docente tutor.*
>
> *Buenos amigos, buenos libros y una conciencia tranquila: esa es la vida ideal.*
>
> *Mark Twain*

Semana 10

En esta semana continuamos con el estudio de las pruebas de hipótesis, específicamente la distribución de proporciones muestrales ya que según explica Suárez (2012)

Cuando el objeto 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 (p.171).

El estudio de la distribución de la proporción muestral requiere que el estudiante sepa discriminar entre población y muestra y calcular la media y desviación típica, para esto es necesario que revise el objeto de aprendizaje Distribución muestral de la proporción que lo preparará para iniciar el estudio de esta sección ¡Bienvenido!

3.5. Distribución de proporciones muestrales *p*

En determinados momentos el interés sobre el comportamiento de una población no está enfocada en la media sino más bien se requiere estimar la proporción (tanto por 1) o el porcentaje (tanto por ciento) de individuos que poseen cierta característica, en este caso es útil emplear la distribución muestral de proporciones.

En esta sección emplearemos los procedimientos indicados para las medias muestrales, con la diferencia de que la desviación típica y el error estándar de la proporción se determinan con los datos de la muestra. Emplearemos la fórmula de:

Distribución de ***proporciones muestrales:***

$$z = \frac{p - P}{\sqrt{\frac{pq}{n}}}$$

Donde:

Z = Área de la distribución normal

p = probabilidad deseada

p = Probabilidad de éxito

q = probabilidad de frecaso

n = Tamaño de la muestra

A continuación, se explicará este contenido en el siguiente problema modelo, le invito a seguir y reflexionar el proceso de resolución, así como los problemas resueltos de la página 340-341 de su texto base.

<u>PROBLEMA MODELO #12</u>

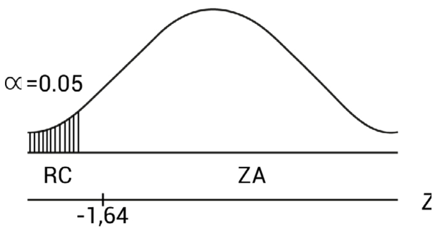
Solución de un problema de prueba de hipótesis.

La fracción de artículos defectuosos de cierto proceso supervisado es 0.14. Un proveedor de materia prima ofrece un nuevo producto, asegurando que reduce la fracción de defectuosos. Con las muestras que el proveedor suministra, se hace un ensayo en la producción con el resultado de 48 defectuosos de un total de 360.

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Contrastar si el proveedor tiene o no razón en la calidad de la nueva materia prima, con un 5% de significación.

Solución

Anotamos los datos del problema:	Calculamos la probabilidad solicitada.
$n = 360$ $a = 0.05$ $P = 0.14$	Hipótesis nula H_0: $P = 0.14$ Hipótesis alternativa H_a: $P < 0.14$ Dócima unilateral a la izquierda.
Cálculo de la probabilidad o proporción: $p = \frac{48}{360} = 0.13$ $q = 1 - 0.13 = 0.87$	Nivel de significación $a = 0.05$ Variante estadística Distribución de medias muestrales $z = \frac{p - P}{\sqrt{\frac{pq}{n}}}$ $z = \frac{0.13 - 0.14}{\sqrt{\frac{(0.13)(0.87)}{360}}} = -0.5642$ La región crítica de la dócima unilateral a la izquierda $Z_s < 1{,}64$  El valor de z= — 0.5642 caen en zona de aceptación a la izquierda por lo tanto se acepta H_0

Concluimos que el proveedor tiene razón en que la calidad de la nueva materia prima reduce la fracción de artículos defectuosos.

Nota. Se explica el proceso para aceptar o rechazar la hipótesis aplicando la distribución de proporciones muestrales. Fuente: Martínez. (2019). Elaboración: Granda S., (2020) Continuando con el análisis de la prueba de hipótesis, y en esta ocasión nos ocuparemos del caso de la distribución de diferencias entre dos medias muestrales. Estas pruebas se utilizan para decidir si las medias de dos poblaciones son iguales. Recuerde es necesario contar con dos muestras independientes, de las dos poblaciones estudiadas. En referencia al uso de este tipo de prueba Suarez (2012) explica “Con frecuencia se utilizan pruebas de dos muestras para comparar dos métodos de enseñanza, dos marcas, dos ciudades, dos distritos escolares y otras cosas semejantes” (p.155). Esto se explica en el objeto de estudio PRUEBA DE HIPÓTESIS PARA DIFERENCIAS DE MEDIAS.MUESTRAS GRANDES Y VARANZAS CONOCIDAS , lo invito a revisarlo minuciosamente y luego continuar con el estudio de la siguiente sección. ¡Bienvenido!

3.6. Distribución de diferencias entre dos medias muestrales $\bar{y} - \bar{x}$

Ahora analizaremos el caso de dos poblaciones independientes donde se requiere conocer la diferencia entre sus dos medias muestrales es significativa o si una media es mayor o menor que la otra para esto le recomiendo estimado estudiante revisar y analizar minuciosamente la página 343 -344 de su texto base.

El proceso para realizar la prueba de diferencias con las dos muestras varía según:

- Las muestras son mayores a 30 (muestras grandes)
- Se conocen las desviaciones típicas poblacionales.

Recuerde que variante estadística *Z* depende de las condiciones dadas y podemos utilizar cualquiera de las siguientes identidades.

$$Z = \frac{\bar{x}-\bar{y}}{\sqrt{\frac{\sigma_x^2}{n_1}+\frac{\sigma_y^2}{n_2}}} \quad Z = \frac{\bar{x}-\bar{y}}{\sqrt{\frac{S_x^2}{n_1}+\frac{S_y^2}{n_2}}} \quad Z = \frac{\bar{x}_1-\bar{x}_2}{\sqrt{\frac{S_x^2}{n_1}+\frac{S_y^2}{n_2}}}$$

A continuación, se explicará este contenido en el siguiente problema modelo, le invito a seguir y reflexionar en el proceso de resolución:

PROBLEMA MODELO #13

Solución de un problema de prueba de hipótesis.

Se requiere comparar el nivel salarial de los empleados de dos empresas. La primera reporta que, en una muestra aleatoria de 46 empleados, su salario promedio fue de $818.000, con una desviación estándar de $32.000. Se elige una muestra aleatoria de 60 empleados de la segunda empresa obteniéndose un salario promedio de $842.000 y desviación estándar de $41.000 ¿Con los anteriores resultados podemos concluir que los salarios en la primera empresa son inferiores? (Nivel del 1%).

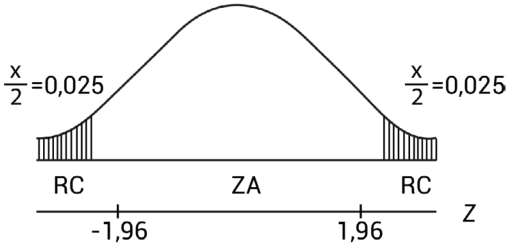
Solución	
Anotamos los datos del problema.	Calculamos la probabilidad solicitada.

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Solución

$n_1 = 100$
$n_2 = 90$
$a = 0.01$
$\bar{x} = 107$
$\bar{y} = 103$
$S_x = 17$
$S_y = 16$

Hipótesis nula $H_0 : \mu_x = \mu_y$
Hipótesis alternativa $H_\alpha : \mu_x \neq \mu_y$
Dócima bilateral



Nivel de significación $a = 0.05$

Cálculo de la varianza

$$S_{\bar{x}-\bar{y}} = \sqrt{\frac{289}{100} + \frac{256}{90}} = \sqrt{2.89 + 2.840} = 2{,}3947$$

Variante estadística

Distribución de medias muestrales $Z = \dfrac{\bar{x}-\bar{y}}{\sqrt{\frac{S_x^2}{n_1} + \frac{S_y^2}{n_2}}}$

$$Z = \frac{107 - 103}{2.3947} = 1.67$$

El valor de Z = 1.67 cae en zona de aceptación de H_0

Concluimos que al nivel del 5% no existe diferencia significativa entre las medias de los dos productos.

Nota. Se explica el proceso para aceptar o rechazar la hipótesis nula del problema, aplicando la distribución de diferencias entre dos medias muestrales. Fuente: Martínez. (2019). Elaboración: Granda S. (2020)

Actividades de aprendizaje recomendadas

Al terminar la semana número diez donde estudiamos las distribuciones de proporciones muestrales y la de diferencias entre dos medias muestrales por medio de su aplicación en diferentes tipos de problemas, le recomiendo desarrollar las siguientes actividades.

- Observar cuidadosamente el recurso audiovisual Prueba de hipótesis **.**
- Revisar minuciosamente el objeto de estudio PRUEBA DE HIPOTESIS PARA DIFERENCIAS DE MEDIAS CON MUESTRAS GRANDES .
- Leer su texto básico en las secciones correspondientes al contenido del Capítulo 8 Pruebas de hipótesis, específicamente los temas de distribución de proporciones muestrales y de diferencias entre medias muestrales, aquí usted encontrará el fundamento teórico y el desarrollo de algunos problemas resueltos que le guiaran en su aprendizaje y práctica de esta unidad.
- Resolver los problemas número 47,49 y 50 de la página 345 de su texto base.

RETROALIMENTACIÓN

En los recursos audiovisuales dados se explican la teoría básica de esta semana y los pasos resumidos para validar una hipótesis de medias con un ejemplo resuelto paso a paso. Así mismo, en el texto básico usted puede acceder a los fundamentos teóricos necesarios para la resolución de todos los problemas propuestos al final de la unidad.

Hemos culminado esta sección de las pruebas de hipótesis, continuaremos en la próxima semana con más casos de distribuciones.

¡Continúe trabajando, su perseverancia y esmero son dignos de felicitar!

> ***Recuerde que:***
>
> *A través del chat de tutoría y consultas puede aclarar sus dudas e inquietudes con su docente tutor.*
>
> *¡No te rindas! Cada logro por más pequeño que sea te aproxima cada día a tu objetivo.*

Semana 11

En esta semana continuamos estudiando las pruebas de hipótesis, esta vez la distribución de diferencias entre dos proporciones muestrales y la distribución **t de Student,** recuerde que el objetivo de las mismas es someter a prueba una afirmación acerca de una propiedad de una población.

> 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 es realmente verdadera (Suarez,2012, p.170).

Por otro lado, recuerde la prueba t de Student para dos muestras independientes se utiliza para determinar si hay una diferencia significativa entre las medias de dichas poblaciones, En este tipo de pruebas debemos contar con una variable dependiente y otra independiente, esta prueba se explica con detenimiento en la sección correspondiente sin embargo le invito a observar el objeto de estudio Estadísticas Z vs estadísticas T , esto nos permitirá comprender el cambio de valor de estos estadísticos, dentro de las pruebas a estudiar en esta semana. ¡Adelante!

3.7. Distribución de diferencias entre dos proporciones muestrales: $P_1 - P_2$

En esta sección utilizaremos proporciones como medidas aplicadas a características cualitativas, tomando en cuenta que la prueba de hipótesis involucra el uso de la distribución normal que nos permitirá saber si existe alguna diferencia entre dos proporciones de poblaciones independientes.

En este caso la hipótesis nula H_0 establece que las dos proporciones de la población son iguales y el parámetro estadístico Z se determinará por medio de la expresión para muestras grandes $n > 30$

$$Z = \frac{p_1 - p_2}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}$$

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Desviación típica proporcional en la muestra $S_{P_1} = \sqrt{p_1 q_1}$ y $S_{P_2} = \sqrt{p_2 q_2}$

A continuación, se explicará este contenido en el siguiente problema modelo, le invito a seguir y reflexionar el proceso de resolución:

PROBLEMA MODELO #14

Solución de un problema de prueba de hipótesis.

Un estudio del consumo de café en el trabajo, por sexo, mostró en una muestra aleatoria de 200 mujeres que 128 lo toman durante su trabajo, mientras que una muestra de 150 hombres reveló que 106 lo toman. ¿Hay alguna diferencia entre la proporción de los dos grupos, en cuanto al hábito de tomar café en el trabajo?

Solución	
Anotamos los datos del problema.	Calculamos la probabilidad solicitada.

Solución

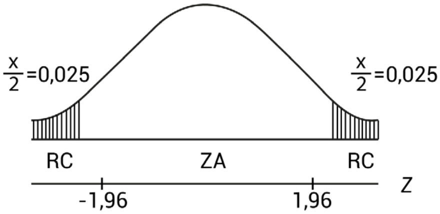
$n_1 = 200$
$n_2 = 150$
$a = 0.05$

Hipótesis nula $H_0: P_1 = P_2$
Hipótesis alternativa $H_a: P_1 \neq P_2$
Dócima bilateral

Determinamos:

$$p_1 = \frac{128}{200} = 0.64$$

$$p_2 = \frac{106}{150} = 0.71$$



Nivel de significación a = 0.05
Cálculo de la varianza

$$S_{P_1} = \sqrt{(0.64)(0.36)} = 0.288$$

$$S_{P_2} = \sqrt{(0.71)(0.29)} = 0.2059$$

Variante estadística

$$Z = \frac{p_1 - p_2}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}$$

$$Z = \frac{0.64 - 0.71}{\sqrt{\frac{0.64(0.36)}{200} + \frac{0.71(0.29)}{150}}} = -1.39$$

El valor de z = 1.67 cae en zona de aceptación de H_0

Concluimos que al nivel del 5% no existe diferencia en cuanto a los hábitos de tomar café.

Nota. Se explica el proceso para aceptar o rechazar la hipótesis nula del problema, aplicando la distribución de diferencias entre dos proporciones. Fuente: Martínez. (2019). Elaboración: Granda S. (2020).

3.8. Distribución t Student

La distribución **t Student** es una prueba de hipótesis de medias en la cual se usa la distribución t, para esto se necesita un espacio muestral menor de 30 es decir una nuestra pequeña, una población que sea considerada normal y se puede contar o no con la desviación típica poblacional.

Para ampliar la definición de distribución t Student necesario que revise detalladamente las páginas 351-353 de su texto base. Recuerde la función de la distribución "t" es:

$$Y = C\left(1 + \frac{t^2}{v}\right)^{-\frac{v+1}{2}}$$

Para comprender esta expresión es importante revisar la definición de grados **de libertad,** que según Martínez (2019) "están dados por el número de valores que pueden ser asignados arbitrariamente, antes de que el resto de las variables queden completamente determinadas" (p.351). Los grados de libertad se pueden simbolizar por *v* y es igual a V = n — 1

Finalmente, querido estudiante es importante que entienda el tratamiento que debe dar a la varianza o a la desviación estándar muestral para las pruebas de hipótesis con muestras pequeñas en las distintas distribuciones. Para esto debe utilizar la Tabla 8.3 **t" de Student** de la página 354 de su texto básico, esta tabla toma en cuenta los grados de libertad y el nivel de significación para pruebas de una y dos colas.

Recuerde que si la dócima es bilateral se tomará el total y si es unilateral, se tomará el doble del nivel de significación asignado. Observar gráfica del texto base página 352 del ejemplo 1 resuelto.

3.9. Distribución de medias muestrales

Para comprender esta distribución debe tomar en cuenta que, si en el problema a resolver se indica la desviación típica muestral y la muestra es pequeña, se considerara que la desviación está sin corregir y teniendo que corregirla para la aplicación en la variante estadística "t".

Se utilizará los siguientes símbolos:

$$\widehat{S} = desviación\ típica\ sin\ sin\ corregir$$

$$S = desviación\ típica\ corregida$$

$$S = \hat{S}\sqrt{\frac{n}{n-1}}$$

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

$$t = \frac{\bar{x} - \mu}{\frac{\hat{S}}{\sqrt{n-1}}}$$

A continuación, se explicará el contenido dado al solucionar el problema modelo, le invito a seguir y analizar dicho proceso.

PROBLEMA MODELO #15

Solución de un problema de prueba de hipótesis.

Un pescador decide que necesita un sedal que resista más de 10 libras si ha de capturar el tamaño de ´pescado que desea. Prueba 16 piezas de sedal de la marca G y halla una media muestral de 10.4 libras. Si en la muestra se obtiene que la desviación típica es de 0.5

libras, ¿qué conclusión se puede sacar del sedal de la marca G.? (Nivel de significación del 5%)

Solución

Anotamos los datos del problema.

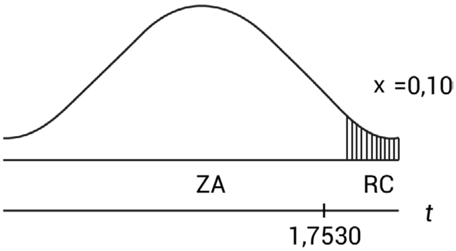
$\mu = 10$
$n = 16$
$\bar{x} = 10.4$
$\hat{S} = 0{,}5$
$\alpha = 0.05$

Determinamos los grados de libertad.

$v = 16 - 1 = -15$
$v = 15$

Calculamos la probabilidad solicitada.

Hipótesis nula $H_0: \mu = 10$
Hipótesis alternativa $H_a: \mu > 10$
Dócima unilateral
Por ser unilateral se toma el doble del nivel de significación en la RC (zona de rechazo)



Nivel de significación $\alpha = 0.10$
Valor del punto crítico obtenido de la tabla 8.3 = 1.7530 como se indica en la figura.
Cálculo de la variante estadística "t".

$$t = \frac{\bar{x} - \mu}{\frac{\hat{S}}{\sqrt{n-1}}}$$

$$t = \frac{10.4 - 10}{\frac{0.5}{\sqrt{15}}} = \frac{0.4\sqrt{15}}{0.5} = \frac{0.4(3.87)}{0.5} = 3.10$$

Concluimos que al nivel del 5% el sedal de la marca G, ofrece garantía de resistencia superior a 10 libras.

Nota. Se explica el proceso para aceptar o rechazar la hipótesis nula del problema, aplicando la distribución de medias muestrales. Fuente: Martínez. (2019). Elaboración: Granda S. (2020).

Actividades de aprendizaje recomendadas

Se ha culminado el trabajo de la semana once, donde profundizamos en la definición, características y aplicación de la distribución de diferencias entre dos proporciones muestrales, distribución t Student y distribución de medias muestrales, luego de esto le recomiendo desarrollar las siguientes actividades.

- Observar cuidadosamente el recurso audiovisual: Uso de la Tabla T-Student.
- Interactúe con la aplicación Student t Distribution cambiando los grados de libertad para que compare el área bajo la curva de una distribución normal con una distribución t.
- Leer en su texto básico el Capítulo 8 Pruebas de hipótesis, específicamente los temas de distribución **t de Student** y distribución de medias muestrales para muestras pequeñas, aquí a parte del fundamento teórico encontrará el desarrollo de algunos problemas resueltos, un resumen sobre el trato a la varianza para muestras grandes y pequeñas y la tabla de distribución “t” Student, todo esto le ayudará en el aprendizaje y práctica de esta unidad.
- Resolver de la página 349 el problema # 58 y de la página 357 el problema #72.

RETROALIMENTACIÓN

Los recursos audiovisuales explican el uso de la tabla t **de Student**, su origen, definición y aplicación, la comparación de áreas de las diferentes distribuciones y en el texto básico usted puede acceder a los fundamentos teóricos necesarios para la resolución de todos los problemas propuestos al final de la unidad.

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Hemos finalizado el estudio de esta semana, en la que hemos profundizado en un grupo más de distribuciones muestrales dentro de las pruebas de hipótesis, está listo para continuar avanzando con el contenido de la siguiente semana.

¡Adelante con su trabajo eficiente en el desarrollo de las actividades planteadas!

> ***Recuerde que:***
>
> *A través del chat de tutoría y consultas puede aclarar sus dudas e inquietudes con su docente tutor.*
>
> *El futuro les pertenece a quienes creen en la belleza de sus sueños.*
>
> *Eleanor Roosevelt*

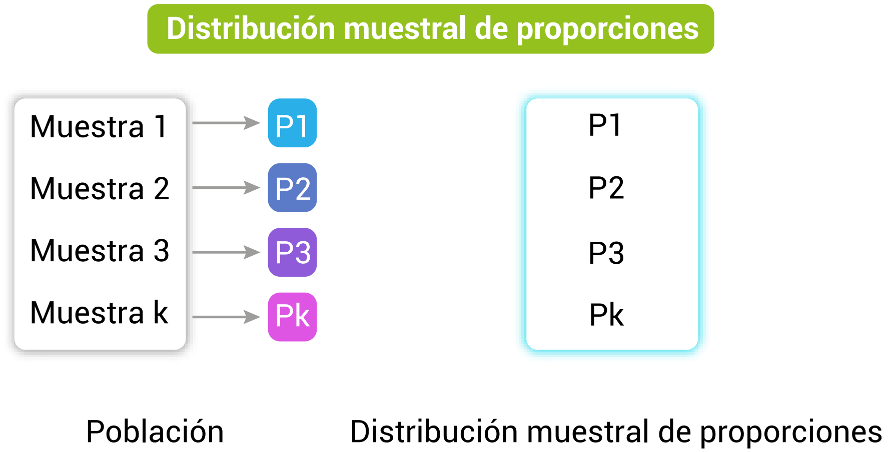
Semana 12

Seguimos trabajando y en esta semana nos corresponde estudiar la distribución de una proporción muestral y de diferencias entre dos medias muestrales, recuerde que se calcula un estadístico de prueba (*t*) que tiene una distribución muestral conocida (t-Student o normal por ejemplo), luego se compara el valor obtenido con los valores críticos $t_{\frac{\alpha}{2}}$ de la distribución para un nivel de significancia establecido. En este momento considero necesario recordar los conceptos esenciales para avanzar en el estudio de las pruebas de hipótesis, por lo que lo invito a observar el objeto de estudio Contraste de hipótesis: definición y conceptos básicos luego de este repaso lo invito a seguir profundizando en las pruebas o contrastes de hipótesis ¡Adelante!

3.10. Distribución de una proporción muestral

Cuando se requiere investigar la proporción de algún atributo en una muestra la distribución muestral de una proporción es la adecuada para dar respuesta a dichas situaciones.

Figura 7.
Distribución muestral de proporciones



Nota. La figura muestra el conjunto de una población conformada por un número k muestras de dónde se obtienen proporciones y el conjunto de la distribución muestral formado por k proporciones. Fuente: Facultad de Educación a Distancia (FAEDIS). (2018).

Si la muestra tomada es igual o menor que 30 que se consideran muestras pequeñas, debe elegirse la distribución t, por lo que recordaremos los pasos generales para probar una hipótesis.

1. Establecer la hipótesis nula y la hipótesis alternativa.
2. Seleccionar un nivel de significancia para la prueba.
3. Identificar el estadístico de prueba.
4. Se formula una regla para tomar decisiones.
5. Se toma una muestra y se llega a una decisión: se acepta o se rechaza la hipótesis.

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

En las pruebas de hipótesis se debe comprobar con base en los resultados obtenidos en una muestra, si el valor verdadero de una proporción es igual a una constante determinada.

En el caso de una sola proporción P las hipótesis son:

Hipótesis nula	Hipótesis alternativa
$H_0 : P = p_o$ Donde p_o es la constante determinada.	Se tiene las siguientes tres posibilidades: $H_1 : P > p_o$ $H_1 : P < p_o$ $H_1 : P \neq p_o$

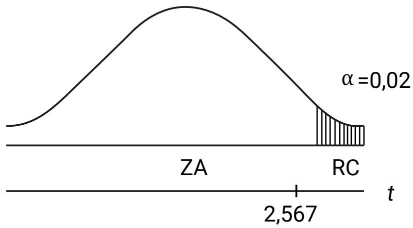
Así mismo para determinar el estadístico de prueba emplearemos $t = \frac{p-P}{\sqrt{\frac{pq}{n-1}}}$ expresión que aplicaremos en la resolución del siguiente problema modelo, lo invito a revisar el proceso desarrollado y a consultar en su texto base los problemas ahí resueltos.

<u>PROBLEMA MODELO #16</u>

Solución de un problema de prueba de hipótesis.

Un fabricante de automóviles afirma que sus autos de tipo familiar, en el 86% de los casos pueden resistir un choque de frente a una velocidad inferior a los 70 k/h, si utilizan cierto equipo. Se toma una muestra de 18 vehículos que tienen este equipo; se encuentran que 16 autos resisten un choque de frente. ¿Se puede decir, al nivel del 1% que el equipo es mucho más efectivo que la afirmación del fabricante?

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Solución	
Anotamos los datos del problema.	Calculamos la probabilidad solicitada.
***P* = 0.86** **n = 18** $p = \frac{16}{18} = 0.89$ ***α* = 0.01** Determinamos los grados de libertad. *v* = 18 — 1 = 17 *v* = 17	Hipótesis nula H_o : *P* = 0.86 Hipótesis alternativa H_a : *P* > 0.86 Dócima unilateral derecha Por ser unilateral se toma el doble del nivel de significación en la RC (zona de rechazo)  Nivel de significación *α* = 0.02 Valor del punto crítico obtenido de la tabla 8.3 z= 2.567 como se indica en la figura. Cálculo de la variante estadística "t". $t = \frac{p-P}{\sqrt{\frac{pq}{n-1}}} = \frac{0.89-0.86}{\sqrt{\frac{0.89(0.11)}{18-1}}} = 0.40$
Cálculo de la desviación típica. $S_p = \sqrt{pq}$ $S_p = 0.89(0.11)$	
El equipo no es mucho más efectivo que el señalado por el fabricante, al nivel del 1%.	

Nota. Se explica el proceso para aceptar o rechazar la hipótesis nula del problema, aplicando la distribución de medias muestrales. Fuente: Martínez. (2019).
Elaboración: Granda S. (2020).

Ahora continuamos con la segunda parte de esta semana, la distribución de diferencias entre dos medias muestrales, recuerde que cuando n_1 o n_2 o ambas, son menores de 30 y se desconocen las varianzas poblacionales, se usa el estadístico t siempre que se pueda suponer razonablemente que las poblaciones son normales y que $\sigma_1 = \sigma_2 = \sigma$.

3.11. Distribución de diferencias entre dos medias muestrales $\bar{x} - \bar{y}$

Para tamaños muestrales menores o iguales a 30 utilizaremos la distribución "t" de Student, a través de diferentes fórmulas según el tipo de información dada, para esto le recomiendo revisar su texto base en las páginas 360-361 donde se explica los casos donde se considera las muestras de dos poblaciones que tienen varianzas poblacionales iguales ($\sigma_x^2 = \sigma_y^2$), condición que se cumple en la mayoría de los casos y que implica calcular el error estándar de las diferencias entre las dos medias muestrales. Así tenemos que:

1. Con la información de cada observación muestral se debe calcular una varianza común para las dos muestras.

$$S^2 = \frac{\sum(x_i - \bar{x})^2 + \sum(y_i - \bar{y})^2}{n_1 + n_2 - 2} \quad \text{o} \quad S^2 = \frac{\left[\sum x_i^2 - n_1\bar{x}^2\right] + \left[\sum y_i^2 - n_2\bar{y}^2\right]}{n_1 + n_2 - 2}$$

Con este resultado debemos calcular el error estándar para la diferencia entre las medias muestrales.

$S_{\bar{x}-\bar{y}} = \sqrt{\frac{S^2}{n_1} + \frac{S^2}{n_2}}$ para finalmente obtener la variante estadística

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{S^2}{n_1} + \frac{S^2}{n_2}}}$$

2. Si las dos varianzas muestrales se dan ya calculadas (supuestamente corregidas) se aplica las siguientes formulas:

$$S_{\bar{x}}^2 = \frac{\sum(x_i - \bar{x})^2}{n_1 - 1} = \frac{\sum x_i^2 - n_1\bar{x}^2}{n_1 - 1} \text{ y } S_{\bar{y}}^2 = \frac{\sum(y_i - \bar{y})^2}{n_1 - 1} = \frac{\sum y_i^2 - n_1\bar{y}^2}{n_1 - 1}$$

Con este resultado debemos calcular el error estándar para la diferencia entre las medias muestrales.

$$S_{\bar{x}-\bar{y}} = \sqrt{\frac{(n_1-1)S_{\bar{x}}^2 + (n_2-1)S_{\bar{y}}^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

para finalmente obtener la variante estadística

$$t = \frac{\bar{x}-\bar{y}}{\sqrt{\frac{(n_1-1)S_{\bar{x}}^2 + (n_2-1)S_{\bar{y}}^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

Finalmente cabe recalcar que en los casos en que las varianzas poblacionales son desiguales o su igualdad no es probable, no se debe utilizar los procesos indicados anteriormente sino más bien una posibilidad sería la distribución "t" de Student, con los grados de libertad que resulten de la expresión.

$$\nu = \frac{\left(\frac{S_{\bar{x}}^2}{n_1}\right)+\left(\frac{S_{\bar{y}}^2}{n_2}\right)}{\frac{\left(\frac{S_{\bar{x}}^2}{n_1}\right)}{n_1-1}+\frac{\left(\frac{S_{\bar{y}}^2}{n_2}\right)}{n_2-1}}$$

A continuación, se explicará el contenido explicado anteriormente en la solución del problema modelo, le invito a seguir y analizar dicho proceso, así como los problemas resueltos de su texto base.

PROBLEMA MODELO # 17 Prueba de hipotesis

PROBLEMA MODELO #17

Solución de un problema de prueba de hipótesis.

Se propone un nuevo método para fabricar concreto. Para docimar (someter a prueba) si el nuevo método ha aumentado la resistencia a la comprensión, se elaboran cinco bloques con cada método.
La resistencia a la comprensión, en libras por pulgada cuadrada, resultan se las siguientes:

NUEVO METODO:	154	143	132	147	139
METODO USUAL	144	131	155	126	134

¿Diría usted que el nuevo método ha aumentado la resistencia a la comprensión?

Solución

Determinamos las medias respectivas $\bar{x}$, y y organizamos los datos en una tabla para obtener los datos necesarios para el cálculo de la varianza común de las dos muestras.

$$\bar{x} = \frac{715}{5} = 143 \quad \bar{y} = \frac{690}{5} = 138$$

x_i	y_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$y_i - y$	$(y_i - \bar{y})^2$
154	144	11	121	6	36
143	131	0	0	-7	49
132	155	-11	121	17	289
147	126	4	16	-12	144
139	134	-4	16	-4	16
715	690	0	274	0	534

$$S^2 = \frac{\sum(x_i - \bar{x})^2 + \sum(y_i - \bar{y})^2}{n_1 + n_2 - 2} = \frac{274 + 534}{5 + 5 - 2} = \frac{808}{8} = 101$$

Cálculo del error estándar para la diferencia entre las medias muestrales.

$$S_{\bar{x}-\bar{y}} = \sqrt{\frac{S^2}{n_1} + \frac{S^2}{n_2}} = \sqrt{\frac{101}{5} + \frac{101}{5}} = \sqrt{\frac{202}{5}} = \sqrt{40{,}4} = 6{,}36$$

Solución

Hipótesis nula $H_0 : \mu_x = \mu_y$

Hipótesis alternativa $H_a : \mu_x > \mu_y$

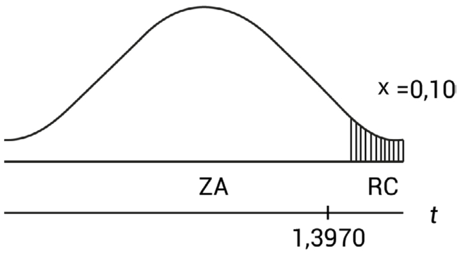
Dócima unilateral a la derecha.

$$\alpha = 0.05$$

Por ser unilateral se toma el doble del nivel de significación en la RC (zona de rechazo) $\alpha = 0.10$

$$v = n_1 + n_2 - 2 = 5 + 5 - 2 = 8$$

El valor del punto crítico que se obtiene de la tabla 8.3 con $\alpha = 0.10$ y $v = 8$ es 1.397 como corresponde en la figura dada a continuación.



Cálculo de la variante estadística "t".

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{S^2}{n_1} + \frac{S^2}{n_2}}} = \frac{143 - 138}{6.36} = 0,79$$

Se puede concluir que el nuevo método no ha aumentado la resistencia a la comprensión, al nivel 5%.

Nota. Se explica el proceso para aceptar o rechazar la hipótesis nula del problema, aplicando la distribución de dos medias muestrales. Fuente: Martínez. (2019). Elaboración: Granda S. (2020).

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Actividades de aprendizaje recomendadas

Al culminar la semana doce, donde se analizó la distribución de una proporción muestral y de diferencias entre dos medias muestrales para muestras pequeñas recuerde revisar su texto base los conceptos y aplicaciones teórico prácticas de estas distribuciones y proceda a realizar las siguientes actividades:

- Analizar la aplicación con EXCEL en la página 369-380 para la prueba de hipótesis, esto le permitirá utilizar la tecnología para simplificar los procesos de solución de los problemas estudiados en esta unidad.
- Analizar el objeto de estudio Prueba de hipótesis para diferencia de medias con muestras pequeñas aquí se desarrolla paso a paso un ejemplo.
- Observar el objeto de estudio tutorial pasos contraste de hipótesis

RETROALIMENTACIÓN

El texto explica los procesos a seguir para resolver problemas de distribución de una proporción muestral y de diferencias entre dos medias muestrales por medio de ejemplos desarrollados paso a paso usted podrá comprender los temas estudiados en esta unidad y así mismo logrará simplificar los procesos de solución de los problemas por medio del empleo del EXCEL. Los objetos de estudio le brindaran una síntesis de los tipos de pruebas de hipótesis estudiados, así como su aplicación.

¡El trabajo desarrollado hasta este punto merece felicitaciones!

> ***Recuerde que:***
>
> *A través del chat de tutoría y consultas puede aclarar sus dudas e inquietudes con su docente tutor.*
>
> *Aunque nadie puede volver atrás y hacer un nuevo comienzo, cualquiera puede comenzar a partir de ahora y crear un nuevo final.*
>
> *Carl Bard*

AUTOEVALUACIÓN TERCERA UNIDAD

Estimado estudiante, luego de culminar con el estudio de la tercera unidad **Pruebas de Hipótesis** es recomendable que usted realice la autoevaluación de lo aprendido, con la finalidad de retroalimentar aquellas inquietudes más importantes. Recuerde que la unidad dos de la presente guía corresponde al capítulo ocho de su texto base.

Autoevaluación 3

Instrucción: Marque la alternativa correcta.

PREGUNTAS

1. El termino Estadística se refiere al grupo de valiosos medios analíticos utilizados para recopilar, organizar, analizar e interpretar información numérica para tomar decisiones eficaces y adecuadas por lo que la inferencia estadística comprende dos partes principales:

   a. Estimación de parámetros y la prueba o docimasia de hipótesis.
   b. Organizar y resumir conjuntos de datos numéricos.
   c. Distribución de frecuencias y gráficas estadísticas.

2. La inferencia estadística está basada en el supuesto de tomar muchas muestras, todas con igual probabilidad de ser seleccionadas y a través de una de ellas sabremos algo acerca de la población, mediante el cálculo de estimadores, que nos permitan hacer aseveraciones, incorrectas algunas veces, estableciéndose la probabilidad de error.

   a. Este método se basa en la aplicación de técnicas de regresión.
   b. Este método se basa en la aplicación de técnicas de recuento.
   c. Este método se basa en la aplicación de técnicas de muestreo.

3. Las pruebas de hipótesis, denominadas también pruebas de significación, tienen como objeto principal evaluar suposiciones o afirmaciones acerca de los valores estadísticos de la población, denominados parámetros por lo que una hipótesis estadística, también puede considerarse como:

   a. La afirmación acerca de una característica ideal de un conglomerado sobre el cual hay inseguridad en el momento de formularla y que, a la vez, es expresada de tal forma que puede ser rechazada.

   b. La afirmación acerca de una característica ideal de una población sobre la cual hay inseguridad en el momento de formularla y que, a la vez, es expresada de tal forma que puede ser rechazada.

   c. La afirmación acerca de una característica ideal de un grupo de datos sobre la cual hay inseguridad en el momento de formularla y que, a la vez, es expresada de tal forma que puede ser rechazada.

4. Relacionar las definiciones correctas.

I.	Corresponde a la hipótesis alternativa o falsa, estableciendo que el parámetro puede ser mayor, menor o igual, de acuerdo con la propuesta hecha en la hipótesis nula.	a.	Error tipo II
II.	Es considerada como la hipótesis nula, ya que hace referencia al valor del parámetro que se quiere probar como verdadero.	b.	Error tipo I
III.	Aceptar la hipótesis nula (Ha) cuando se ha debido rechazar. En el ejercicio que estamos desarrollando, sería: "Aceptar la moneda como correcta, cuando en verdad no lo es".	c.	H_0
IV.	Rechazar la hipótesis nula (Ha) cuando se ha debido aceptar. "Rechazar la moneda como incorrecta, cuando en verdad está equilibrada".	d.	H_1

5. Tomando en cuenta que existen dos posibles decisiones: aceptar o rechazar la hipótesis que a la vez puede ser cierta o falsa y si tomamos en cuenta que los dos tipos de error son inherentes al proceso de la prueba de significación y la probabilidad de cometer error será igual al nivel de significación, entonces si tiene las decisiones:

		VERDADERA	FALSA
A	ACEPTAR	Decisión correcta	Error tipo II
B	ACEPTAR	Error Tipo I	Decisión correcta
C	RECHAZAR	Error Tipo I	Decisión correcta
D	RECHAZAR	Decisión correcta	Error tipo II

6. Suponga que estamos realizando una prueba de hipótesis de la afirmación de que un método de selección del género aumenta la probabilidad de una niña, de modo que la probabilidad de que nazca una niña es $p > 0{,}5$. Las hipótesis nula y alternativa serían

$$H_0 : p = 0{,}5$$

$$H_1 : p > 0{,}5$$

Entonces la afirmación que identifica un error de tipo II es:

a. El método de selección de género es eficaz cuando en realidad no tiene efecto.

b. El método de selección del género no tiene efecto cuando en realidad es eficaz para incrementar la probabilidad de que nazca una niña.

c. El método de selección de género no es eficaz cuando en realidad no tiene efecto.

7. Se ha dicho que una hipótesis estadística es un supuesto, concerniente a los parámetros o a la forma de la distribución de probabilidad, correspondiente a una o más poblaciones dadas. En otras palabras, se resume diciendo que corresponde a un enunciado acerca de un valor estadístico (parámetro) poblacional por lo que un ejemplo de hipótesis sería:

   a. La desintegración familiar de los padres provoca baja autoestima en los hijos.
   b. La exposición de los adolescentes a los videojuegos en Ecuador
   c. Los alumnos de la zona rural son más disciplinados que los de la zona urbana.
   d. La ansiedad en los alumnos en la práctica docente.

8. Considere la afirmación de que el peso medio de pasajeros de aeronaves (incluyendo al equipaje de mano) es, a lo sumo, de 195 libras (el valor que actualmente esa la Federal Aviation Administration). A partir de esto identifique la hipótesis nula y la hipótesis alternativa correcta:

   a. H_0 $\mu \leq 195$ libras y H_1 $\mu > 195$ libras
   b. H_0 $\mu > 195$ libras y H_1 $\mu >195$ libras
   c. H_0 $\mu = 195$ libras y H_1 $\mu >195$ libras

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

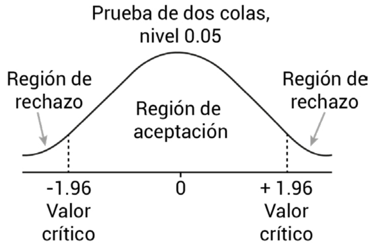
9. En los siguientes diagramas se muestran las áreas para una prueba de dos colas y una prueba de una cola para las que se aplicara la prueba con z. Enlace las definiciones con su gráfica correspondiente:

I. Tipos de error al aceptar o rechazar una hipótesis.

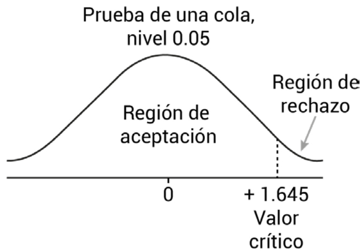
II. Esta prueba de hipótesis es aquélla en la cual la zona de rechazo o zona crítica está completamente comprendida en uno de los extremos de la distribución.

III. Esta prueba de hipótesis es aquella que comprende áreas o zonas de rechazo en ambos extremos de la distribución, ósea que la hipótesis alternativa es diferente.

a.



b.



c.

Decisiones		Verdadera	Falsa
	Aceptar	Decisión correcta	Error tipo II
	Rechazar	Error tipo I	Decisión correcta

10. Se entiende por nivel de significación, la máxima probabilidad de que se especifique con el fin de hacer mínimo el primer tipo de error. Generalmente, esta probabilidad se fija antes de escoger la muestra. Por lo que se puede afirmar que:

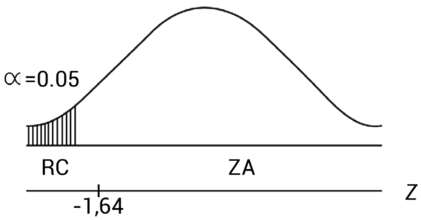
    a. Si se trabaja con un nivel del 10%, el resultado es significativo;
    b. Si se emplea el 1 %, el resultado es altamente significativo.
    c. Si es del 5%, se considera poco significativo.

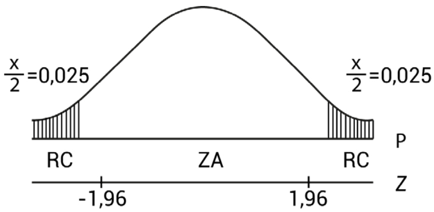
11. Al docimar la hipótesis: μ = 86 teniendo como datos $\bar{x}$ = 82, σ = 15 y n = 25 utilizamos la variante estadística:

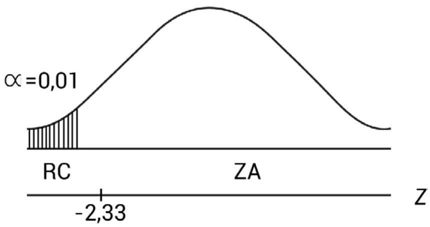
    a. $Z = \frac{X-\mu}{\sigma}$
    b. $Z = \frac{\bar{x}-\mu}{\sigma}$
    c. $Z = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

12. Si la fracción de artículos defectuosos de cierto proceso supervisado es 0,14. Y un proveedor de materia prima ofrece un nuevo producto, asegurando que reduce la fracción de defectuosos. Con las muestras que el proveedor suministra, se hace un ensayo en la producción con el resultado de 48 defectuosos de un total de 360. Al contrastar si el proveedor tiene o no razón en la calidad de la nueva materia prima, con un 5% de significación se verifica la región crítica de la dócima unilateral es:

a.


b.


c.


13. Una fábrica de salsa de tomate desea saber, si una nueva presentación de su producto aumentaría sus ventas. Una muestra de 40 tiendas reveló un promedio mensual de ventas de 310 botellas, con la nueva presentación y una desviación estándar de 20 botellas. Otra muestra realizada en 34 tiendas donde se expende el producto con el envase tradicional, revela unas ventas promedio de 292 botellas con desviación estándar de 26 botellas. ¿Qué dócima se debe utilizar para afirmar que al nivel del 10%, aumentaron las ventas?

    a. Dócima Unilateral Izquierda
    b. Dócima Unilateral Derecha
    c. Dócima Bilateral.

14. Un sociólogo cree que la proporción de hombres que pertenecen a un grupo socio-económico determinado (grupo A) y que ve regularmente deportes en televisión, es superado por un segundo grupo de hombres (grupo B) que también ve deportes. Muestras aleatorias simples de los dos grupos arrojan los siguientes resultados:

GRUPO	TAMAÑO	No. HOMBRES
A	$n_1 = 130$	80
B	$n_2 = 100$	96

Para apoyar la tesis del sociólogo, al nivel del 5% se debe utilizar una dócima unilateral izquierda, entonces el parámetro estadístico Z para este efecto tiene un valor de:

    a. z = 7, 25
    b. z = -7, 25
    c. z = 7,25

15. Dos muestras de 15 especímenes cada una, de dos tipos de género de lana fueron sometidos a una prueba de resistencia. La media de la primera muestra, fue de 131 libras por pulgada cuadrada y de 136 la de la segunda. Las respectivas desviaciones estándar fueron 6,25 y 4,65 libras por pulgada cuadrada.

    Para indicar que la segunda calidad de género es superior se requiere calcular el error estándar para la diferencia entre las medias muestrales, el que es igual a:

    a. $S_{\bar{x}-\bar{y}} = 2{,}01$
    b. $S_{\bar{x}-\bar{y}} = 1{,}01$
    c. $S_{\bar{x}-\bar{y}} = 0{,}01$

Una vez terminada la autoevaluación de la tercera unidad, le recomiendo verificar sus respuestas con el solucionario que se encuentra al final de esta guía didáctica, esto le permitirá validar su desempeño. Si surgen dudas en una o más preguntas vuelva al leer el contenido científico para que identifique la validez de su respuesta.

Ir al solucionario

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Le recuerdo estimado estudiante que usted dispone del horario de tutoría, donde su profesor tutor se encuentra presto a atender sus dudas e inquietudes, él ayudará a resolver las dificultades que se le presenten en el transcurso de la semana de trabajo y a validar su desempeño. Finalmente le motivo a participar de las actividades síncronas planificadas a lo largo de la asignatura, ya que estas son la oportunidad de reducir obstáculos como el tiempo y el espacio y obtener así mejores resultados en su aprendizaje.

Semana 13

En esta semana estudiaremos métodos estadísticos ***no paramétricos,*** que se constituyen en métodos que pueden ser usados con datos nominales y ordinales que pueden ser datos de intervalo o de relación cuando no cabe supuesto alguno sobre la distribución de probabilidad de la población. Por lo que se considera que, por tener requisitos menos restrictivos, sobre la naturaleza de los datos, y por la menor cantidad de supuestos sobre la distribución de la población estadística, de la que proviene la muestra; los métodos no paramétricos pueden ser aplicados en muchas más situaciones que los métodos paramétricos.

Unidad 4. Otras Pruebas de hipótesis

Las pruebas no paramétricas son aquellas que se encargan de analizar datos que no tienen una distribución particular y se basan en hipótesis, pero los datos no están organizados de forma normal. Aunque tiene algunas limitaciones, cuentan con resultados estadísticos ordenados que facilita su comprensión.

La definición de prueba no paramétrica nos la explica Triola (2013) “no requieren que las muestras provengan de poblaciones con distribuciones normales o con cualquier otro tipo particular de distribución. En consecuencia, las pruebas de hipótesis no paramétricas suelen llamarse **pruebas de distribución libre**” (p.662).

El estudio de esta sección le permitirá aplicar pruebas no paramétricas, por lo que lo invito hacerlo con detenimiento y logrará los mejores resultados. ¡Bienvenido!

4.1. Prueba de hipótesis de una varianza

En situaciones como control estadístico de calidad, de antemano se conocen los parámetros de referencia del proceso bajo control. La actividad para decidir si en un momento dado, el proceso está bajo control, es la confrontación permanente de los datos obtenidos con la hipótesis sobre la centralidad del proceso (media) sobre la magnitud de su variabilidad (varianza).

La varianza como medida de dispersión es importante dado que nos ofrece una mejor visión de dispersión de datos.

Así podremos determinar una franja de confianza, con la base en la cual podríamos tomar decisiones al respecto, esto se indica en el objeto Prueba de hipótesis para la varianza.

Para la prueba de una sola varianza obtenida de una muestra aleatoria suponemos que se tiene una población normal con media μ y varianza σ generalmente desconocida. Utilizaremos la siguiente simbología:

$$H_0 : \sigma^2 = \sigma_0^2 \leftrightarrow H_0 : \frac{\sigma^2}{\sigma_0^2} = 1$$

HIPÓTESIS NULA

$$H_a : \frac{\sigma^2}{\sigma_0^2} \neq 1$$

$$H_a : \frac{\sigma^2}{\sigma_0^2} > 1$$

HIPÓTESIS ALTERNATIVA

$$H_a : \frac{\sigma^2}{\sigma_0^2} < 1$$

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$$

VARIANTE ESTADÍSTICA

Chi-cuadrado con $n - 1$ grados de libertad

En su texto base usted encontrará el **proceso** a seguir en este tipo de prueba de hipótesis, así como el uso de la tabla para determinar los valores críticos, todo esto debe revisar y analizar en las páginas 404-406 aquí claramente se explica con un ejemplo los **límites de confianza** y se muestra la **Tabla 9.1 De percentiles de la distribución** χ^2

A continuación, se explicará el contenido explicado anteriormente en la solución del problema modelo, le invito a seguir y analizar dicho proceso, así como los problemas resueltos de su texto base.

PROBLEMA MODELO #18

Solución de un problema de prueba de hipótesis de una varianza

Dada x normalmente distribuida y los valores de la muestra n = 15 y $\hat{S}$ = 7, docimar la hipótesis de σ = 5.

DATOS: n = 15 y $\hat{S}$ = 7, σ = 5

Solución:

Vamos a establecer la hipótesis:

$$H_0: \frac{\sigma^2}{\sigma_0^2} = \frac{\sigma^2}{25} = 1$$

$$H_a: \frac{\sigma^2}{25} \neq 1$$

Solución:

Tomando en cuenta el nivel de significación $\alpha = 0.05$

Establecemos los límites de confianza

Tomando en cuenta que el nivel de confianza es del 95% por lo tanto se establecen dos valores que encierren este porcentaje de la población:

$\frac{\alpha}{2} = 0.025 \quad \frac{\alpha}{2} = 0.9755$

Establecemos el supuesto

La muestra es aleatoria

La población es normal

Formulamos la variante estadística

$$\frac{\chi^2}{\nu} = \frac{\hat{S}^2}{25}$$

Determinamos los valores críticos:

Se calcula los grados de libertad $v = n - 1 = 15 - 1 = 14$

$$\left(\frac{\chi^2}{\nu}\right)_{\frac{\alpha}{2};\nu} \rightarrow \left(\frac{\chi^2}{\nu}\right)_{0.025;14}$$

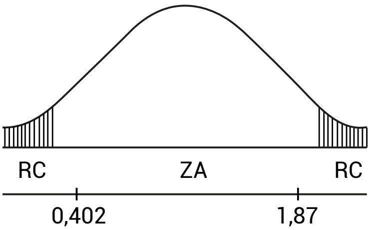
$$\left(\frac{\chi^2}{\nu}\right)_{1-\frac{\alpha}{2};\nu} \rightarrow \left(\frac{\chi^2}{\nu}\right)_{0.975;14}$$

Ahora debemos utilizar la tabla 9.1 de la página 406 de Percentiles de la distribución $\frac{\chi^2}{\nu}$ para determinar los valores críticos

Solución:

$$\left(\frac{\chi^2}{\nu}\right)_{0.025;14} = 0.402$$

$$\left(\frac{\chi^2}{\nu}\right)_{0.975;14} = 1.87$$



Finalmente calculamos el valor estadístico para analizarlo y adoptar una decisión.

La variante estadística debe estar entre los valores críticos encontrados.

$$0.402 > \frac{\chi^2}{\nu} > 1.87$$

Entonces

$$\frac{\widehat{S}^2}{\sigma^2} = \frac{49}{25} = 1.96$$

Se obtiene que la variante estadística esta fuera de la zona de aceptación por lo que se rechaza H_0 y se acepta $H_a: \frac{\sigma^2}{25} \neq 1$

Nota. Se explica el proceso para aceptar o rechazar la hipótesis nula del problema, aplicando la prueba de hipótesis de una varianza. Fuente: Martínez. (2019). Elaboración: Granda S. (2020).

Existe un procedimiento que se utiliza con mayor frecuencia en la prueba con una varianza, dicho proceso utiliza la tabla de Chi-cuadrado y se explica a continuación.

4.1.1. Procedimiento más utilizado en la prueba con una varianza

Otra alternativa para realizar una prueba de hipótesis sobre la varianza de una población es utilizar en el proceso la tabla de ***Chi-cuadrado.*** Aquí vamos a contrastar la hipótesis por medio del siguiente proceso

$H_\alpha : \sigma^2 = \sigma_0^2$

$H_\alpha : \sigma^2 \neq \sigma_0^2$

FORMULACIÓN DE HIPOTESIS

$\alpha = 0.05$ ***GRADO DE SIGNIFICACIÓN***

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$$

VARIANTE ESTADÍSTICA

Pudiendo realizarse pruebas bilaterales y unilaterales a la izquierda o la derecha, según el planteamiento del problema:

A continuación, se explicará este proceso en la solución del problema modelo, le invito a seguir y analizar este problema, así como los problemas resueltos de su texto base.

<u>PROBLEMA MODELO #19</u>

Solución de un problema de prueba de hipótesis de una varianza

Un vendedor experto que trabaja en un almacén, asegura que el tiempo dedicado a la atención de un cliente en promedio es de 25 minutos, con una desviación típica (σ) de 8 minutos. Se pregunta, si el grado de variación en la atención al cliente es diferente. Una muestra a 20 clientes da una desviación típica de 10 minutos. Pruebe la hipótesis al nivel del 10%.

DATOS: n= 20 y S = 10, σ = 8, $\bar{x}$ = 25 *minutos*, σ^2 = 64, α = 0.1

Solución:
Vamos a establecer la hipótesis: $H_0 : \sigma^2 = 64$ $H_a : \sigma^2 \neq 64$
Tomando en cuenta el nivel de significación $\alpha = 0.1$ Formulamos la variante estadística. Recuerde que para fijar los límites de confianza donde debe estar la varianza poblacional con 90% de confiabilidad en este caso, se emplea la fórmula: $$\frac{(n-1)S^2}{\chi_s^2} < \sigma^2 < \frac{(n-1)S^2}{\chi_i^2}$$ En la tabla de distribución de chi cuadrada (χ_s^2) con los grados de libertad $v = 19$ y $\alpha = 0.10$ se obtiene el valor inferior $\chi_i^2 = 11.651$ y con $v = 19$ y $\alpha = 0.90$ se obtiene el valor superior $\chi_s^2 = 27.204$ $$\frac{(20-1)64}{27.204} < \sigma^2 < \frac{(20-1)64}{11.651}$$ $$\frac{1216}{27.204} < \sigma^2 < \frac{1216}{11.651}$$ Por lo tanto los límites para la varianza poblacional son: $44.7 < \sigma^2 < 104.37$. La varianza poblacional, estará entre 44.7 y 104.37 con una confianza del 90% es decir se cumple con lo indicado por el vendedor experto.

Nota. Se explica el proceso para aceptar o rechazar la hipótesis nula del problema, aplicando la prueba de hipótesis de una varianza. Fuente: Martínez. (2019). Elaboración: Granda S. (2020).

Actividades de aprendizaje recomendadas

Al concluir el trabajo de la semana trece donde se estudió otras pruebas de hipótesis, pruebas no paramétricas, recuerde revisar en su texto base los conceptos y aplicaciones teórico prácticas de estas pruebas y proceda a realizar las siguientes actividades:

- Observar cuidadosamente los recursos audiovisuales siguientes:

  Como usar la Tabla Chi-cuadrado
  Prueba de hipótesis para la varianza.

- Leer en su texto básico el Capítulo 9. Otras Pruebas de hipótesis específicamente la prueba de hipótesis para la varianza, aquí a parte del fundamento teórico encontrará el desarrollo de algunos problemas resueltos esto le permitirá solucionar las actividades planteadas en esta unidad.

RETROALIMENTACIÓN

Los videos explican el proceso para desarrollar una prueba de hipótesis para la varianza y el texto muestra los procesos a seguir para resolver problemas de prueba de hipótesis para la varianza, por medio de ejemplos desarrollados paso a paso usted podrá comprender los temas estudiados en esta unidad.

¡Aplausos a todo el trabajo desarrollado en su proceso de aprendizaje!

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

> ***Recuerde que:***
>
> *A través del chat de tutoría y consultas puede aclarar sus dudas e inquietudes con su docente tutor.*
>
> *El éxito no se da de la noche a la mañana. Es cuando cada día eres un poco mejor que el día anterior.*
>
> *Todo suma.*
>
> *DWAYNE JOHNSON*

Semana 14

Seguimos analizando los métodos estadísticos ***no paramétricos*** entre ellos está la comparación entre varianzas de dos poblaciones; en esta sección nos corresponde estudiar la **Distribución F** que busca probar hipótesis a través de la comparación de varianzas poblacionales idénticas o diferentes y la prueba del coeficiente de correlación de Pearson.

La distribución F compara entre las varianzas de dos poblaciones normales, este método es importante porque aparecen en pruebas en las que queremos determinar si dos muestras provienen de poblaciones que tienen varianzas iguales, como se explica en el siguiente objeto de estudio Distribución F , le invito a observarlo detenidamente para luego iniciar el estudio de esta prueba en la siguiente sección. ¡Bienvenido!

4.2. Comparación entre varianzas de dos poblaciones

La prueba ***F*** llamada así en honor del especialista en estadística Sir Ronald Fisher, esta prueba requiere que las dos poblaciones tengan distribuciones normales, esta prueba es muy sensible a las desviaciones que se alejan de la distribución normal.

Para el análisis de las propiedades de la **Distribución *F*** es necesario estimado estudiante que revise en su texto base las páginas 410-414 aquí se presenta el fundamento teórico y una gama de ejemplos que nos permitirá estudiar esta prueba.

Los pasos a seguir en esta prueba son:

1. Plantear las hipótesis H_0 y H_α

$$H_{0:}\ \sigma_1^2 = \sigma_2^2$$

$$H_{\alpha:}\ \sigma_1^2 \neq \sigma_2^2$$

2. Establecer el nivel de significación a = …

3. Determinar el valor de ***F*** en la variante estadística.

$$F = \frac{S_1^2}{S_2^2} = \frac{Varianza\ mayor}{Varianza\ menor}$$

4. Decidir el aceptar o rechazar si las varianzas poblacionales son idénticas.

Finalmente recuerde el consultar la ***tabla F*** que se encuentre al final de su texto base, según el proceso indicado en la página 411, todo esto se explica a continuación a través del desarrollo del problema modelo, el que le invito a revisar y analizar previamente a la ejecución de las actividades planteadas en esta semana.

Índice

Primer bimestre

Segundo bimestre

Solucionario

Glosario

Referencias bibliográficas

PROBLEMA MODELO #20

Solución de un problema de Distribución F.

Se comparan dos métodos para realizar cierta operación. Supongamos que los resultados obtenidos en las dos muestras fueron: $\bar{x} = 725$; $S_x^2 = 61$ y $\bar{y} = 661$; $S_y^2 = 86$; y los tamaños muestrales son: 7 y 13 respectivamente. Docime o pruebe la hipótesis de que la segunda muestra *(y)* , presenta mayor variabilidad que *x*.

DATOS: $n_x = 7$, $n_y = 13$, $\bar{x} = 725$; $S^2_x = 61$ y $\bar{y} = 661$;

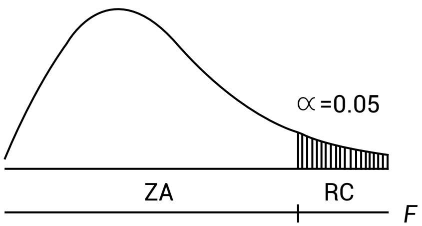
$S^2_x = 86$, $v_x = n - 1 = 7 - 1 = 6$, $v_y = n - 1 = 13 - 1 = 12$

Solución:

Vamos a establecer la hipótesis:

$$\boldsymbol{H_0: \sigma_1^2 = \sigma_2^2}$$

$$\boldsymbol{H_a: \sigma_2^2 < \sigma_1^2}$$

Solución:
Tomando en cuenta el nivel de significación *a* = 0.05 Formulamos la variante estadística. $F = \frac{S_y^2}{S_x^2} = \frac{86}{61} = 1.41$ El valor crítico de la distribución F en nivel de significancia de 0.05 tomada de la tabla de valores críticos, ubicada al final del texto base es 4 tomando como grados de libertad del numerador 12 y grados de libertad del denominador 6.    Estos resultados permiten aceptar H_0 es decir la primera prueba no presenta mayor variabilidad que la segunda al nivel del 5% de significancia.

Nota. Se explica el proceso para aceptar o rechazar la hipótesis nula del problema, aplicando la distribución F. Fuente: Martínez. (2019). Elaboración: Granda S., (2020)

4.3. Prueba del coeficiente de correlación de Pearson *r* = *R*

Para iniciar se requiere recordar en que el análisis de regresión consiste en la búsqueda de una función que exprese la forma en que se relaciona una variable dependiente (Y) con una o más variables independientes(x), esta relación por lo general se aproxima a una línea recta cuya trayectoria se visualiza por medio de un diagrama de dispersión o nube de puntos.

La correlación es una técnica estadística que nos indica si dos variables están relacionas o no, existiendo correlación positiva y negativa.

La **correlación positiva** se presenta cuando la relación entre una variable y otra es lineal y directa, de manera que un cambio en una variable predice el cambio en la otra variable. En ese caso, se dice que la correlación es positiva perfecta, es decir, ambas variables varían al mismo tiempo. Este tipo de correlación es directamente proporcional. Hay correlación positiva cuando las dos variables se correlacionan en sentido directo. Por lo que, a valores altos de una le corresponden valores altos de la otra e igualmente con los valores bajo.

La **correlación negativa** se presenta cuando una variable y otra es opuesta o inversa, es decir, cuando una variable cambia, la otra se modifica hacia lo contrario. Entonces, cuando una variable posee valores altos, la otra posee valores bajos y mientras este valor está más cerca de -1, más evidente será esta covariación. Se dice que hay correlación negativa perfecta cuando $r = -1$.

Con estos antecedentes nos aprestamos a realiza la prueba de coeficiente de correlación Pearson por lo que lo invito a observar el objeto de estudio Prueba de Hipótesis de Correlación

La prueba de coeficiente de **correlación Pearson** busca determinar si el coeficiente de correlación calculado en un análisis de regresión es o no producto del azar.

La prueba de hipótesis en correlación contempla los siguientes aspectos:

- Cuando la muestra es pequeña de utiliza la distribución “t” y “z”.

- Para n > 30 se aplica la fórmula $t = r\sqrt{\frac{n-2}{1-r^2}}$

  $r: coeficentete\ de\ corrrelación \qquad n: \#\ de\ pares\ observados$

- También: $t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} \qquad r = \frac{n\sum x_i y_j - (\sum x_i)(\sum y_j)}{\sqrt{[n\sum x_i^2 - (\sum x_i)^2][n\sum y_i^2 - (\sum y_i)^2]}}$

- Para estos casos se considera que la Hipótesis nula H_0 es igual, mayor o menor que 0.

A continuación, se explicará este proceso en la solución de un problema modelo, le invito a seguirlo y analizarlo, así como los problemas resueltos de su texto base.

PROBLEMA MODELO #21

Solución de un problema de prueba del coeficiente de correlación de Pearson.

En un estudio de correlación entre el número de accidentes infantiles en relación con las zonas verdes, se obtuvo un Coeficiente de correlación de -0.92 (r = —0.92) sobre la base de 18 barrios de Cali. Mediante una dócima bilateral y un nivel de significación del 5%, ¿qué se puede concluir respecto al Coeficiente de correlación?

DATOS: *r* = —0.092 y *n* = 18, *a* = 5%

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Solución:

Vamos a establecer la hipótesis:

$$H_0: R = 0 \rightarrow \textit{no existe correlación}$$

$$H_a: R \neq 0 \rightarrow \textit{el coeficiente de correlación es significativo}$$

O también para poblaciones tenemos: $H_0: p = 0$

$$H_1: p > 0$$

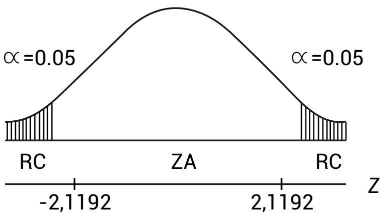
Tomamos en cuenta el nivel de significación $\propto = 0.05$

También los grados de libertad $v = n - 2 = 18 - 2 = 16$

Calculamos el parámetro estadístico $t = r\sqrt{\frac{n-2}{1-r^2}}$

$$\boldsymbol{t = -0.92\sqrt{\frac{16}{1-(-0.92)^2}} = -9.39}$$

Determinar el valor crítico se encuentra con la distribución t Student y debe ser una prueba bilateral, recuerde que para esta prueba $\propto = \frac{\mathbf{0.05}}{\mathbf{2}} = \mathbf{0.025}$



Finalmente hacemos la comparación $-9.39 < -2.1192$ y se nota que el parámetro estadístico t se ubica en la zona de rechazo. Por lo que se concluye que al nivel 5% el coeficiente de correlación es extremadamente significativo.

Nota. Fuente: se explica el proceso para aceptar o rechazar la hipótesis nula del estudio dado en el problema, aplicando la prueba del coeficiente de correlación de Pearson. Martínez. (2019). Elaboración: Granda S. (2020).

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Actividades de aprendizaje recomendadas

Al concluir el trabajo de la semana catorce, donde se analizó la comparación entre varianzas de dos poblaciones y la prueba del coeficiente de correlación de Pearson, recuerde revisar en su texto base los conceptos y aplicaciones teórico prácticas de estas pruebas de hipótesis y luego proceda a realizar las siguientes actividades:

- Observar cuidadosamente el recurso audiovisual siguiente: Prueba de hipótesis para la varianza de dos poblaciones.
- Leer en su texto básico el Capítulo 9. Otras Pruebas de hipótesis específicamente la distribución F y prueba del coeficiente de correlación de Pearson.
- Revisar el proceso para realizar la Prueba de hipótesis de la correlación
- Resolver los problemas 18 y 19 planteados en su texto base en la página 418 por medios analíticos y la tecnología.

RETROALIMENTACIÓN

Los recursos audiovisuales indican por medio de la resolución de problemas de aplicación de la prueba de hipótesis para la varianza de dos poblaciones (***Distribución F***) y la prueba de **hipótesis de la correlación**, además el texto explica los procesos a seguir para resolver problemas de aplicación del contenido estudiado en esta semana, por medio de ejemplos desarrollados paso a paso usted podrá desarrollar las actividades planteadas en esta unidad.

¡Adelante en su empeño de aprendizaje, ya que lo ha pensionado en la recta final, felicitaciones!

Recuerde que:

A través del chat de tutoría y consultas puede aclarar sus dudas e inquietudes con su docente tutor.

Siempre parece imposible hasta que se hace.

Nelson Mandela

Semana 15

En esta semana estudiaremos las pruebas con observaciones apareadas en poblaciones dependientes, a diferencia de lo estudiado anteriormente donde se trabajó con poblaciones independientes. Lo invito a culminar su proceso de formación en esta asignatura con la perseverancia que lo ha colocado en este sitio de la ruta del aprendizaje. ¡Bienvenidos!

4.4. Pruebas con observaciones apareadas

Este es un procedimiento estadístico que se aplica en grupos pareados (equivalentes), se caracteriza por medir a un mismo grupo en ciertas condición antes y después de aplicar una variable, nos permite identificar si existe diferencias estadísticamente significativas cuando comparamos dos grupos de datos, todo esto se explica en el objeto de estudio T pareada que le invito a observar cómo trabajo previo al estudio de la siguiente sección.

Para el estudio de estas pruebas según indica Martínez (2019) “en las observaciones apareadas, se debe tomar una muestra aleatoria de pares, de manera que cada observación esté asociada con alguna observación en particular" (p.420). Aquí se debe aclarar que las muestras de análisis son dependientes.

La simbología que se utilizará es la siguiente:

$d_i = x_i - y_i$	Diferencia para cada observación
$\bar{d} = \frac{\sum d_i}{n}$	Media aritmética de las medianas
n	Número total de pares que conforman la muestra
N	Numero de pares de observación en la población.
$a_{\bar{d}} = \frac{\sum d_i}{N}$, $a_{\bar{d}} = 0$	Media de la diferencia de la población
$S_d = \sqrt{\frac{\sum(d_i - \bar{d})^2}{n-1}}$	Desviación típica en observaciones apareadas
$t = \frac{\bar{d} - a_{\underline{d}}}{\frac{S_d}{\sqrt{n}}} = \frac{\bar{d} - 0}{S_{\bar{d}}}$	Variante estadística
$S_{\bar{d}} = \frac{S_d}{\sqrt{n}}$	Error estándar en observaciones apareadas.

Esta prueba tiene aplicaciones muy importantes por lo que le recomiendo revisar la explicación de las páginas 419-423 de su texto base, donde consta la fundamentación de esta prueba y la simbología a utilizar la misma que será utilizada a continuación en la resolución de un ejemplo modelo, le invito a revisarlo paso a paso, así como los problemas resueltos en su texto base.

PROBLEMA MODELO #22 Pruebas no paramétricas-Prueba de Chi-Cuadrado χ2

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

PROBLEMA MODELO #22

Solución de un problema de prueba con observaciones apareadas.

Una manzana recién descubierta tiene un sabor delicioso. Se ha decidido docimar (someter a prueba) su rendimiento, plantando este tipo de manzanas junto a otras corrientes, en ocho huertos diseminados en una región apropiada para la producción de ambas variedades. Cuando los árboles empiezan a rendir frutos, se mide su producción en cajas. Los datos obtenidos son los siguientes:

HUERTO	NUEVA MANZANA	MANZANA CORRIENTE
1	13	12
2	14	16
3	19	17
4	10	9
5	15	16
6	14	12
7	12	10
8	11	8

¿Señalan estos resultados una mayor producción para la nueva manzana que para la corriente?

Solución:

En primer lugar, organizamos la información en una tabla para obtener los parámetros necesarios para aplicar la prueba de observaciones pareadas.

HUERTO	x_i	y_i	$d_i = x_i - y_1$	$d_i - (\bar{d}_i)$	$(d_i - \bar{d})^2$
1	13	12	1	0	0
2	14	16	-2	-3	9
3	19	17	2	1	1
4	10	9	1	0	0
5	15	16	-1	-2	4
6	14	12	2	1	1
7	12	10	2	1	1
8	11	8	3	2	4
∑	-	-	8	0	20

Cálculo de la desviación típica en observaciones apareadas.

$$S_d = \sqrt{\frac{\sum\left((d_i - \bar{d})^2\right)}{n-1}}$$

Se requiere la media aritmética de las diferencias $\bar{d} = \frac{\sum d_i}{n} \longrightarrow \bar{d} = \frac{8}{8} = 1$

$$S_d = \sqrt{\frac{20}{7}} = 1.69$$

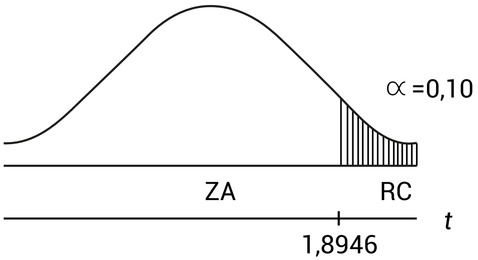
Con este valor procedemos a determinar el error estándar de las observaciones propuestas: $S_{\bar{d}} = \frac{S_d}{\sqrt{n}} \longrightarrow \frac{1.69}{\sqrt{8}} = 0.60$

Solución:

Establecemos las hipótesis de prueba:

$$H_0 = a_d = 0$$
$$H_a = a_d > 0$$

Los grados de libertad $v = n - 2 = 8 - 1 = 7$ y el nivel de significación a = 0.10 nos obtenemos el valor crítico t = 1.8946.



Con todo esto se determina el parámetro estadístico $t = \frac{\bar{d}}{S_{\bar{d}}} = \frac{1}{0.60} = 1.67$

Con todo esto podemos concluir que al nivel del 5% estos resultados no señalan una mayor producción para la nueva manzana.

Nota. Se explica el proceso realizar una prueba de observaciones apareadas en problema de aplicación. Fuente: Martínez. (2019). Elaboración: Granda S. (2020).

4.5. Pruebas no paramétricas

En los estudios de comportamiento o de conducta dentro de las ciencias sociales u otras, se utiliza métodos no paramétricos para el análisis de datos, siempre que sea posible establecer la forma de la distribución poblacional, o cuando dichos datos estén dados a escala ordinal y ordenados por rangos. Las pruebas no paramétricas son contrastes que no plantean hipótesis sobre parámetros y que se limitan a analizar propiedades nominales u ordinales de los datos en una distribución libre. Las pruebas no paramétricas permiten poner a prueba hipótesis no referidas a parámetros poblacionales ya que no necesitan establecer supuestos exigentes sobre las poblaciones de donde se extrae las muestras y no necesitan trabajar con datos obtenidos con una escala de medida o intervalo. Dentro de estas pruebas no paramétricas estudiaremos la Prueba del Chi o Ji cuadrado.

4.5.1. Prueba de CHI-CUADRADO χ^2

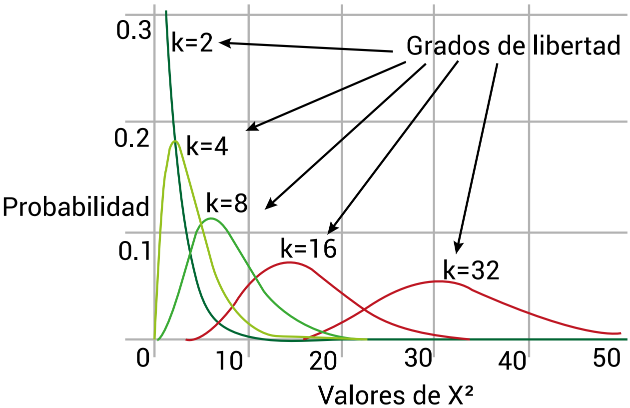
Es una distribución de probabilidad continua con un parámetro "k", que representa los grados de libertad de la variable aleatoria, la distribución de Chi-Cuadrado es denotada por la letra griega χ^2, es frecuentemente usada para probar hipótesis, concernientes a la diferencia entre un conjunto de frecuencias observadas de una muestra y un conjunto correspondientes de frecuencias teóricas esperadas.

Propiedades

- No toma valores negativos.
- La gráfica de la función toma una curva sesgada a la derecha.
- A medida que aumentan los grados de libertad la curva va tomando simetría.

Figura 8.

Función de distribución de probabilidad.



Nota. La figura muestra la distribución de probabilidad continua con un parámetro k que representa los grados de libertad de la variable aleatoria. Fuente: Triola (2013).

Esta prueba no paramétrica es una de las más utilizadas, como explica Martínez(2019) "$J_{i-cuadrado}$ es la suma de las fracciones que tienen por numerador el cuadrado de las diferencias entre las frecuencias reales u observadas y las frecuencias esperadas o teóricas y por denominador la frecuencia esperada"(p.428) así mismo le invito a revisar el objeto de estudio Estadístico ji-cuadrada para prueba de hipótesis esto lo preparará para emprender en la siguiente sección de estudio.

Utilizaremos la expresión como $\chi^2 = \sum \frac{(n_i - n_i^*)^2}{n_i^*}$ donde:

$$F_O = n_i = Frecuencia\ observada\ \ o\ real.$$

$$F_e = n_i^* = Frecuencia\ teórica\ o\ esperada$$

La fundamentación teórica de esta prueba se encuentra en su texto base dentro de las paginas 427-430 por lo que le recomiendo revisar a través de una lectura comprensiva y analítica.

El proceso para aplicar la prueba ***Chi-cuadrado*** se recomienda los siguientes pasos:

1. Formular la hipótesis.
2. Establecer la diferencia entre las frecuencias observadas y las esperadas.
3. Elevar cada diferencia al cuadrado y dividirlas para su frecuencia teórica o esperada. (una a una) para sumarlas.
4. Calcular χ^2

A continuación, se explicará este proceso en la solución de un problema modelo, le invito a seguirlo y analizarlo detenidamente, así como los problemas resueltos de su texto base.

PROBLEMA MODELO # 23

Solución de un problema con la prueba Chi-cuadrado.

Un dado se lanza 100 veces y se obtiene los siguientes resultados:

CARA	1	2	3	4	5	6
n_i	12	17	20	22	13	16

Determine si el lado puede considerarse perfecto, al nivel del 1% de significación.

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Solución:

Se organiza la información en una tabla y se determina los valores necesarios para calcular χ^2.

n_i	n_i^*	$n_i - n_i^*$	$(n_i - n_i^*)^2$	$\frac{(n_i - n_i^*)^2}{n_i^*}$
12	**16.666**	**-4.666**	**21.77**	**1.306**
17	**16.666**	**0.334**	**0.11**	**0.007**
20	**16.666**	**3.334**	**11.12**	**0.667**
22	**16.666**	**5.334**	**28.45**	**1.707**
13	**16.666**	**-3.334**	**13.44**	**0.806**
16	**16.666**	**-0.666**	**0.44**	**0.026**
100	**99.999**	**0.004**	**-**	**4.519**

Para completar esta tabla calculamos $p = \frac{1}{6}$, la frecuencia teórica esperada $n_i^* = pn = \frac{1}{6}(100) = 16.666$, los grados de libertad v = n — 1 = 6 —1 = 5, con esto y el nivel de significación *a* = 0.01 determinar el valor critico $\chi^2_{0.01} = 15.09$ que es el valor del estadístico χ^2 cuando el área del extremo derecho de la cuerva para 5 grados de libertad, como se indica a continuación

Tabla IV, distribución de J, cuadrado (continuación)

v	A=0,30	A=0,20	A=0,10	A=0,05	A=0,02	A=0,01
1	1,07	1,64	2,71	3,84	5,41	6,64
2	2,41	3,22	4,60	5,99	7,82	9,21
3	3,66	4,64	6,25	7,82	9,84	11,34
4	4,88	5,99	7,78	9,49	11,67	13,28
5	6,06	7,29	9,24	11,07	13,39	$\boxed{15,09}$

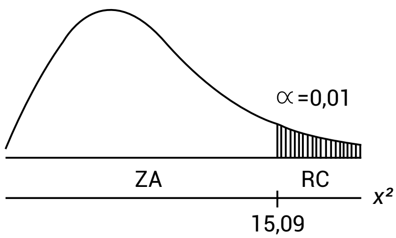
Solución:
Establecemos las hipótesis de prueba:

$$H_0 = n_i = n^*_i$$
$$H_a = n_i = n^*_i$$

Esta hipótesis también se puede escribir en términos de la frecuencia:

$$H_0 = F_o = F_e$$
$$H_a = F_o \neq F_e$$

Representamos el valor crítico en la distribución:



Con todo esto se determina el parámetro estadístico $\chi^2 = \sum \frac{(n_i - n_i^*)^2}{n_i^*}$

$\chi^2 = 4.519$

Con todo esto podemos concluir que al nivel del 5% estos resultados no señalan una mayor producción para la nueva manzana.

En la gráfica se puede notar que el valor de $\chi^2 = 4.519$ se ubica en la zona de aceptación por lo que podemos concluir que el dado no está cargado al nivel de significación del 1%.

Nota. Se explica el proceso para aplicar la prueba Chi-cuadrado para establecer si el dado del problema está o no cargado. Fuente: Martínez. (2019). Elaboración: Granda S. (2020).

Actividades de aprendizaje recomendadas

Al terminar la semana quince, luego de analizar las pruebas con observaciones apareadas y la prueba no paramétricas Chi-Cuadrado, es necesario afianzar estos conocimientos analizando en su texto base los conceptos y aplicaciones teórico prácticas de este contenido, revise también la explicación para realizar cálculos aplicando la tecnología, por medio del programa EXCEL y luego proceda a realizar las siguientes actividades:

- Observar cuidadosamente el siguiente recurso audiovisual: Prueba Chi-Cuadrada, la prueba no paramétrica más utilizada.
- Leer en su texto básico el Capítulo 9. Otras Pruebas de hipótesis específicamente la prueba con observaciones apareadas y la prueba no paramétrica Chi-cuadrado.
- Revisar su texto base la aplicación del Excel en pruebas de Chi-cuadrado, página 436-440.
- Resolver los problemas # 44 y 45 de su texto base, en la página 441.

RETROALIMENTACIÓN

El video explica el proceso para desarrollar una prueba de Chi-Cuadrada, además el texto muestra los procesos a seguir para resolver problemas de aplicación del contenido estudiado en esta semana, por medio de ejemplos desarrollados paso a paso usted podrá desarrollar las actividades planteadas en esta unidad.

¡Felicitaciones, su trabajo responsable le ha permitido avanzar en su aprendizaje!

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

AUTOEVALUACIÓN CUARTA UNIDAD

Estimado estudiante, luego de culminar con el estudio de la cuarta unidad **Otras Pruebas de Hipótesis** es recomendable que usted realice la autoevaluación de lo aprendido, con la finalidad de retroalimentar aquellas inquietudes más importantes. Recuerde que la unidad dos de la presente guía corresponde al capítulo nueve de su texto base.

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Autoevaluación 4

Instrucción: Marque la alternativa correcta.

1. Las pruebas no paramétricas no requieren que las muestras provengan de poblaciones con distribuciones normales o con cualquier otro tipo particular de distribución. En consecuencia, las pruebas de hipótesis no paramétricas suelen llamarse

   A. Pruebas de distribución aleatoria.
   B. Pruebas de distribución compuesta.
   C. Pruebas de distribución libre.

2. Relacione la simbología con la definición pertinente

   i. $\boldsymbol{H_0: \sigma^2 = \sigma_0^2 \leftrightarrow H_0: \frac{\sigma^2}{\sigma_0^2} = 1}$

   ii. $\boldsymbol{H_a: \frac{\sigma^2}{\sigma_0^2} \neq 1}$

   $\boldsymbol{H_a: \frac{\sigma^2}{\sigma_0^2} > 1}$

   $\boldsymbol{H_a: \frac{\sigma^2}{\sigma_0^2} < 1}$

   iii. $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$

   a. VARIANTE ESTADÍSTICA
   b. HIPÓTESIS ALTERNATIVA
   c. HIPÓTESIS NULA

3. Al docimar la hipótesis σ = 20, dada $\hat{S}$ = 10 para una muestra de tamaño 25. Se establece las hipótesis:

A. $\left\{H_o: \frac{\sigma^2}{\sigma_o^2} = \frac{\sigma^2}{20^2} = 1 \; H_a: \frac{\sigma^2}{20^2} \neq 1\right\}$

B. $\left\{H_o: \frac{\sigma^2}{\sigma_o^2} = \frac{\sigma^2}{20^2} \neq 1 \; H_a: \frac{\sigma^2}{20^2} = 1\right\}$

C. $\left\{H_o: \frac{\sigma^2}{\sigma_o^2} = \frac{\sigma^2}{20^2} \geq 1 \; H_a: \frac{\sigma^2}{20^2} \leq 1\right\}$

4. Para 100 empleados de una Cía. Se observa el salario mínimo mensual expresado en miles de pesos (x), con el fin de determinar si la varianza anterior de todo los salarios fue de $\sigma^2 = 8$, se obtuvieron los siguientes resultados de esa observación: $\sum x_i^2 = 2000 \quad \frac{(\sum x_i)^2}{n} = 1000$

Suponiendo que se corre un riesgo de equivocarse el 5% ¿qué conclusión obtiene la Cía.?

A. Aceptamos que $\frac{\sigma^2}{8} \neq 1$, por lo tanto se puede rechazar que la varianza anterior era 8, al nivel de significación del 5%.

B. Aceptamos que $\frac{\sigma^2}{8} = 1$, por lo tanto se puede admitir que la varianza anterior era 8, al nivel de significación del 5%.

C. Rechazamos que $\frac{\sigma^2}{8} = 1$, por lo tanto se puede descartar que la varianza anterior era 8, al nivel de significación del 5%.

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

5. En una muestra de 35 pares, el Coeficiente de correlación obtenido es 0,80. Si se usa un nivel de significación del 5%, para docimar la hipótesis de que r = 0,90 se debe calcular la desviación estándar con la expresión

A. $\sigma_z = \frac{1}{\sqrt{n-3}}$

B. $\sigma_z = \frac{1}{\sqrt{n+3}}$

C. $\sigma_z = \frac{2}{\sqrt{n-3}}$

6. En una población con Coeficiente de correlación 0,65, se toman muchas muestras de igual tamaño. Si el 15% de estas muestras tienen un coeficiente de correlación de 0,75 o más, ¿cuál es el tamaño de la muestra?

En este problema el valor de la transformación de *r* = 0,75 a *z* según la TABLA V es:

Tabla V. transformación de *r A z (es decir,* $z=0,5\ Log_e \frac{1+r}{1-r}$*)*

r	z	r	z	r	z
.00	.000	.36	.377	.71	.887
.01	.010	.37	.388	.72	.908
.02	.020	.38	.400	.73	.929
.03	.030	.39	.412	.74	.950
.04	.040	.40	.424	.75	.973
.05	.050				
.06	.060	.41	.436	.76	.996
.07	.070	.42	.448	.77	1.020
.08	.080	.43	.460	.78	1.045
.09	.090	.44	.472	.79	1.071
.10	.100	.45	.485	.80	1.099

A. z = 0,377
B. z = 0,973
C. z = 0,377

7. La distribución de Chi-Cuadrado es denotada por la letra griega χ^2, es frecuentemente usada para probar hipótesis, concernientes a la diferencia entre un conjunto de frecuencias observadas de una muestra y un conjunto correspondientes de frecuencias teóricas esperadas, cuyas propiedades son:

   A. Toma valores negativos, la gráfica de la función toma una curva sesgada a la derecha y a medida que aumentan los grados de libertad la curva va tomando simetría.
   B. No toma valores negativos, la gráfica de la función toma una curva sesgada a la izquierda y a medida que aumentan los grados de libertad la curva va tomando simetría.
   C. No toma valores negativos, la gráfica de la función toma una curva sesgada a la derecha y a medida que aumentan los grados de libertad la curva va tomando simetría.

8. El proceso que se recomienda para aplicar la prueba Chi-cuadrado es:

   A. Formular la hipótesis, establecer la diferencia entre las frecuencias observadas y las esperadas, elevar cada diferencia al cuadrado y dividirlas para su frecuencia teórica o esperada y calcular χ^2.
   B. Calcular χ^2, elevar cada diferencia al cuadrado y dividirlas para su frecuencia teórica o esperada, Formular la hipótesis, establecer la diferencia entre las frecuencias observadas y las esperadas,
   C. Establecer la diferencia entre las frecuencias observadas y las esperadas, Formular la hipótesis, elevar cada diferencia al cuadrado y dividirlas para su frecuencia teórica o esperada y calcular χ^2.

9. Una estadística de accidentes leves, ocurridos en dos fábricas A y B muestran que, de 102 accidentes, 59 han tenido lugar en la fábrica A y 43 en la fábrica B. La hipótesis de que no existe relación entre el número de accidentes y el hecho de que ocurran en la fábrica A o en la fábrica B es:

   A. $H_0 : n_i \neq n^*_i$
   B. $H_0 : n_i = n^*_i$
   C. $H_0 : n_i > n^*_i$

10. El valor de verdad de los siguientes enunciados es:

   - χ^2 puede tener valores negativos
   - Si se tiene 4 filas y 3 columnas, los grados de libertas son 5
   - La suma de los valores esperados es igual a la suma de los valores observados

   A. FFV
   B. FVF
   C. VVF

Una vez terminada la autoevaluación de la cuarta unidad, le recomiendo verificar sus respuestas con el solucionario que se encuentra al final de esta guía didáctica, esto le permitirá validar su desempeño. Si surgen dudas en una o más preguntas vuelva al leer el contenido científico para que identifique la validez de su respuesta.

Ir al solucionario

Le recuerdo estimado estudiante que usted dispone del horario de tutoría, donde su profesor tutor se encuentra presto a atender sus dudas e inquietudes, él ayudará a resolver las dificultades que se le presenten en el transcurso de la semana de trabajo y a validar su desempeño. Finalmente le motivo a participar de las actividades síncronas planificadas a lo largo de la asignatura, ya que estas son la oportunidad de reducir obstáculos como el tiempo y el espacio y obtener así mejores resultados en su aprendizaje.

> ***Recuerde que:***
>
> *A través del chat de tutoría y consultas puede aclarar sus dudas e inquietudes con su docente tutor.*
>
> *La motivación es lo que te pone en marcha, el hábito es lo que hace que sigas.*
>
> *Jim Rohn*

Actividades finales del bimestre

Semana 16

Le recomiendo llevar a cabo las siguientes actividades, como trabajo previo a rendir la evaluación presencial del segundo bimestre.

Actividad 1. Recupere las actividades de aprendizaje calificadas desarrolladas durante el segundo bimestre.

Actividad 2. Organice un documento que contenga dicha información cuidando la presentación; si lo hizo a mano, puede fotografiar y organizar un documento en PDF.

Actividad 3. Revise los contenidos del segundo bimestre y participe de la evaluación presencial.

Para ello considere:

- Su diario de notas.
- Actividades de aprendizaje recomendadas.
- Actividades de aprendizaje calificadas.
- Actividades de aprendizaje interactivas; y,
- Evaluaciones parciales.

Actividades de aprendizaje recomendadas

Además de las actividades anteriores, para consolidar sus aprendizajes del segundo bimestre, y prepararse a la prueba presencial en físico o virtual, sugiero:

- Revise cada uno de los conceptos estudiados en las dos unidades planificadas y desarrolladas en este segundo bimestre.

- Realice suficientes ejercicios y problemas de aplicación de los diferentes conceptos, propiedades y leyes, de cada una de las unidades estudiadas, desarrollando los problemas propuestos al final de cada unidad del texto básico.

- Para cada una de las unidades, es valedero que sistematice el conocimiento aprendido, a través de la construcción de un mapa conceptual o algún otro organizador gráfico que estime conveniente.

Recuerde que:

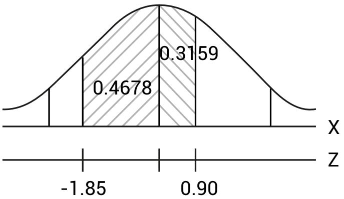
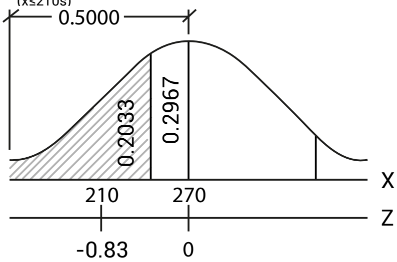
A través del chat de tutoría y consultas puede aclarar sus dudas e inquietudes con su docente tutor.

No dejes que lo que no puedes hacer interfiera con lo que puedes hacer.

John R. Wooden

4. Solucionario

Autoevaluación 1	
Pregunta	**Respuesta y retroalimentación**
1	Respuesta: B En el caso del lanzamiento de una moneda, esta pueda ser: cara o sello, por lo tanto, son sucesos mutuamente excluyentes.
2	Respuesta: B. Aplicando $P_{(x=2)} = \left(\frac{4!}{2!\,2!}\right)\left(\frac{1}{2}\right)^{2}\left(\frac{1}{2}\right)^{4-2} = \frac{6}{16} = 0.375 = 37.5\%$
3	Respuesta: A. Aplicando n=8, p=0.8 (ganar), q=1-0.8=0.2, x=0,1,2,3,4,5 y 6, $P_{(x\leq6)}$ = ? $P_{(x\leq6)}=P_{(0)}+P_{(1)}+P_{(3)}+P_{(4)}+P_{(5)}+P_{(6)}=1-[P_{(7)}+P_{(8)}]$ (aplicando complemento) $P_{(x\leq6)}=1-[(8\ 7)\ (0.8)^{7}\ (0.2)^{1}+(8\ 8)\ (0.8)^{8}\ (0.2)^{0}]=1-[0.3355+0.1678]$ $P_{(x\leq6)}$ = 49.67%
4	Respuesta: B. Calculando la media $\lambda = np$ para p = 0.03 y n = 100 tenemos que λ = 0.03 (100) = 3. Aplicando Poisson: $P_{(x=2)} = \frac{3^2 e^{-3}}{2!} = \frac{3^2(0.04979)}{2!} = 0.2240 = 22.4\%$
5	Respuesta: B. Por ejemplo, para x=0,1,2,3 y 4. Cuando se escribe VERDADERO, la función devolverá $P_{(x\leq4)}=P_{(0)}+P_{(1)}+P_{(3)}+P_{(4)}$, en cambio si se escribe FALSO, la función devolverá solo de $P_{(x=4)}=P_{(4)}$
6	Respuesta: C. $p = \frac{3}{100000} = 0.00003$ y n = 200000 tenemos que λ = 0.00003 (200000) = 6. Aplicando Poisson para $P_{(x\leq2)}$: $P_{(x\leq2)} = \frac{6^0e^{-6}}{0!} + \frac{6^1e^{-6}}{1!} + \frac{6^2e^{-6}}{2!} = 0.002479 + 0.014874 + 0.044622 = 6.20\%$

Autoevaluación 1	
Pregunta	**Respuesta y retroalimentación**
7	Respuesta: B. Utilizando la tabla, para *Z*=0.90 => *A*(0.3159) y para *Z*=-1.85 => *A*(0.4678) por simetría. Sumando las áreas 0.3159 + 0.4678 = 0.7837 
8	Respuesta: F Las características de la distribución normal son: Simetría, asintótica, el área bajo la curva es aproximadamente del 100%, la media, se localiza en el centro de la campana y se puede aplicar en los ejercicios de distribución Binomial, siendo su resultado un valor aproximado.
9	Respuesta: A. Con μ = 4.5 min = 270 s, σ = 72s y $P_{(x\leq 3.5min)} = P_{(x\leq 210s)}$ = ? $Z = \frac{X-\mu}{\sigma} = \frac{210-270}{72} = -0.83 \rightarrow A(0.2967)$ $P_{(x\leq 210s)}$ = 0.5000 — 0.2967 = 0.2033 = 20.33 
10	Respuesta A. i-b, ii-c, iii-a **Variable Aleatoria:** Está conformada por factores donde interviene el azar. **Variable Aleatoria discreta:** Esta puede asumir un número finito de valores que se pueden contar. **Variable Aleatoria continua:** Esta puede asumir cualquier valor dentro de un intervalo, es infinita.

Ir a la autoevaluación

Autoevaluación 2	
Pregunta	**Respuestas y retroalimentación**
1	Respuesta: I-b, II-c, III-a ▪ **Población o Universo**: conjunto de unidades y elementos que presentan una característica común. Estas unidades se denominan variables continuas y discretas. ▪ **Unidad:** son cada uno de los valores que se han obtenido al realizar un estudio estadístico. ▪ **Muestra:** Es un subconjunto representativo de unidades de análisis de una población dada.
2	Respuesta: I-c, II-a, III-b ▪ **Muestreo aleatorio simple o muestreo aleatorio irrestricto**, en el cual se da igual oportunidad de selección a cada elemento o unidad dentro de la población. ▪ **Muestreo sistemático**. La selección de las unidades se hace a intervalos regulares, en un orden sistemático. ▪ **Muestreo aleatorio estratificado** (Asignación igual, proporcional y óptimo), garantiza la representatividad, reduciendo el error de la muestra al formar grupos o subpoblaciones más o menos homogéneas, en cuanto a su composición interna y heterogénea cuando se comparan entre sí.
3	Respuesta: A. Teorema. Dada una población, si extraemos todas las muestras posibles de un mismo tamaño, entonces la media de la distribución de todas las medias muestrales posibles, será igual a la media poblacional.
4	Respuesta: **Teoría del límite central**. Se cumple, cuando independientemente de la población origen, la distribución de las medias aleatorias se aproximan a una distribución normal a medida que el tamaño de la muestra crece. Se podrá decir también, que, si las muestras provienen de una población que no es normal, es de importancia tener en cuenta el tamaño de la muestra, si el tamaño muestral es pequeño, la distribución obtenida con sus medias muestrales tendrán un comportamiento similar al de la población de donde se extrajeron. Por el contrario, si el tamaño muestral es grande, el comportamiento de estas medias muestrales será igual al de una distribución normal, independientemente de la población de donde fueron extraídas.

Autoevaluación 2	
Pregunta	**Respuestas y retroalimentación**
5	Respuesta: A. Con μ = 864500 σ = 15000, n = 25 $Z = \frac{857500-864500}{\frac{15000}{\sqrt{25}}} = -2{,}33$ Entonces el área vara este valor de Z es *A*=0.4901 y la probabilidad es igual a *P* = 0,5000 — 0,4901 = 0,0099, finalmente $P_{\bar{x}} < 857500$ = 0,99%
6	Respuesta: A Con los datos dados en el problema *P* = 0,25, *Q* = 0,75, $p = \frac{8}{36} = 0{,}22$, $P_{(p<0.22)}$ = ? Procedemos a determinar el estadístico Z, $z = \frac{0{,}22-0{,}25}{\sqrt{\frac{0{,}22(0{,}78)}{36}}} = -0.43$ con este valor identificamos el área bajo la curva en la tabla 6.9. Áreas de una distribución normal ordinaria *A*=0.1664 para este valor la probabilidad es igual a P = 0,5000 — 0,1664 = 0,3336, finalmente $P_{(p<0{,}22)}$ = 33,36%
7	Respuesta: A La desviación típica de las diferencias entre los pares de medias muestrales. $\sigma_{\bar{x}-\bar{y}} = \sqrt{\frac{\sum[(\bar{x}_i - \bar{y}_i) - (\mu_x - \mu_y)]^2}{M}}$
8	Respuesta: B En la distribución de deferencias entre dos proporciones muestrales para determinar el estadístico Z se emplea la expresión $Z = \frac{(p_1-p_2)-(\mu_{p_1}-\mu_{p_2})}{\sqrt{\frac{P_1Q_1}{n_1}+\frac{P_2Q_2}{n_2}}}$
9	Respuesta: A Nivel de confianza. Tiene relación directa con el tamaño de la muestra, por lo tanto, se dirá que a mayor nivel de confianza más grande debe ser el tamaño de la muestra. Los valores de Z se obtienen mediante el uso de tablas.
10	Respuesta: A La fórmula que se utiliza para el cálculo de *n* en poblaciones infinitas en la variable es $n = \left(\frac{Z\sigma}{E}\right)^2$

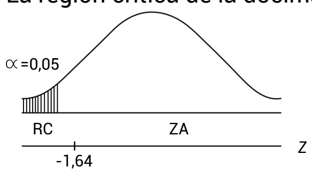
Ir a la autoevaluación

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Autoevaluación 3	
Pregunta	**Respuestas y retroalimentación**
1	Respuesta: A La inferencia estadística comprende dos partes principales, a saber: la estimación de parámetros y la prueba o docimasia de hipótesis.
2	Respuesta: C Este método se basa en la aplicación de técnicas de muestreo, para lo cual se requiere de un buen diseño, además de la aplicación de métodos aleatorios de selección, cuando las probabilidades son iguales para cada elemento de una población. En algunos casos no requieren ser iguales, siempre que se conozcan y sean diferentes a cero.
3	Respuesta: B Una hipótesis estadística, también puede considerarse como la afirmación acerca de una característica ideal de una población sobre la cual hay inseguridad en el momento de formularla y que, a la vez, es expresada de tal forma que puede ser rechazada.
4	Respuesta: i-d, iii-b, ii-c, iv-a H_1: Corresponde a la hipótesis alternativa o falsa, estableciendo que el parámetro puede ser mayor, menor o igual, de acuerdo con la propuesta hecha en la hipótesis nula. Error tipo I: Aceptar la hipótesis nula (Ha) cuando se ha debido rechazar. En el ejercicio que estamos desarrollando, sería: “Aceptar la moneda como correcta, cuando en verdad no lo es”. H_0: Es considerada como la hipótesis nula, ya que hace referencia al valor del parámetro que se quiere probar como verdadero. Error tipo II: Rechazar la hipótesis nula (Ha) cuando se ha debido aceptar. “Rechazar la moneda como incorrecta, cuando en verdad está equilibrada”.
5	Respuesta: A-C
6	Respuesta: B Un error tipo II se comete al no rechazar la hipótesis nula cuando es falsa. El siguiente es un error de tipo II: no rechazar $H_0 : p = 0{,}5$ (y, por lo tanto, no sustentar $p > 0{,}5$) cuando en realidad $p > 0{,}5$. Es decir, comeríamos un error tipo II si concluimos que el método de selección del género no tiene efecto cuando en realidad es eficaz para incrementar la probabilidad de que nazca una niña.

Autoevaluación 3	
Pregunta	**Respuestas y retroalimentación**
7	Respuesta A. Es una hipótesis causal ya que tenemos la presencia de dos variables x: desintegración familiar y: baja autoestima C. Es una hipótesis tiene como finalidad comparar grupos, en este caso x: alumnos de la zona rural y: alumnos de la zona urbana.
8	Respuesta A. Se expresa la afirmación dada en forma simbólica. La afirmación de que la media es a lo sumo de 195 libras que se expresan en forma simbólica como $\mu \leq 195$ *libras*. Si es falso que $\mu \leq 195$ *libras* entonces $\mu \leq 195$ *libras* debe ser verdad. De las dos expresiones simbólicas $\mu \leq 195$ *libras* y $\mu \leq 195$ *libras*, vemos que $\mu \leq 195$ *libras* no contiene igualdad, por lo que permitiremos que la hipótesis alternativa H_1 sea $\mu > 195$ *libras*. A sismo, la hipótesis nula debe ser la afirmación de que la media es igual a 195 libra, por lo que dejamos que H_0 sea $\mu \leq 195$ *libras*
9	Respuesta: I-c, II-b, III-a En los siguientes diagramas se muestran las áreas para una prueba de dos colas y una prueba de una cola para las que se aplicara la prueba con z. Si se aplica una prueba de dos colas, la regla de decisión plantea que si el valor calculado de z queda entre mas y menos 1.96, se acepta la hipótesis nula. De otra manera se rechaza. También se muestra la tabla que explica los tipos de error al aceptar o rechazar una hipótesis.
10	Respuesta: B Cuando se trabaja con un nivel del 5%, el resultado es significativo; si se emplea el 1 %, el resultado es altamente significativo, y si es del 10%, se considera poco significativo.
11	Respuesta En el caso planteado utilizamos la distribución de medias muestrales $z = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}$

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

<table>
<tr><th colspan="2">Autoevaluación 3</th></tr>
<tr><th>Pregunta</th><th>Respuestas y retroalimentación</th></tr>
<tr><td>12</td><td>Respuesta A
Hipótesis nula $H_0 : P < 0.14$
Hipótesis alternativa $H_a : P < 0.14$
Dócima unilateral a la izquierda.
Nivel de significación $a = 0.05$
Variante estadística
Distribución de medias muestrales
$z = \frac{p-P}{\sqrt{\frac{pq}{n}}}, z = \frac{0.13-0.14}{\sqrt{\frac{(0.13)(0.87)}{360}}} = -0.5642$
La región crítica de la dócima unilateral a la izquierda $Z_s < 1,64$


El valor de z = -0.5642 caen en zona de aceptación a la izquierda por lo tanto se acepta H_0</td></tr>
<tr><td>13</td><td>Respuesta. B
En el caso de una distribución de diferencias entre dos medias muestrales, puede plantearse:
$H_0 : \mu_x = \mu_y$
$H_0 : \mu_x > \mu_y$
Dócima unilateral a la derecha.</td></tr>
<tr><td>14</td><td>Respuesta: B
$Z = \frac{p_1 - p_2}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} = \frac{0,62 - 0,96}{\sqrt{\frac{0,62(0,38)}{130} + \frac{0,96(0,04)}{100}}} = -7,25$</td></tr>
<tr><td>15</td><td>Respuesta A
$S_{\bar{x}-\bar{y}} = \sqrt{\frac{(n_1 - 1)S_x^2 + (n_2 - 1)S_y^2}{n_1 + n_2 - 2}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$
$S_{\bar{x}-\bar{y}} = \sqrt{\frac{(15-1)39,06 + (15-1)21,62}{15+15-2}} \sqrt{\frac{1}{15} + \frac{1}{15}} = 5,51(0,365) = 2,01$</td></tr>
</table>

Ir a la autoevaluación

Autoevaluación 4	
Pregunta	**Respuestas y retroalimentación**
1	Respuesta: C La definición de prueba no paramétrica nos la explica Triola (2013) “no requieren que las muestras provengan de poblaciones con distribuciones normales o con cualquier otro tipo particular de distribución. En consecuencia, las pruebas de hipótesis no paramétricas suelen llamarse **pruebas de distribución libre**” (p.662).
2	Respuesta: 1-c, 2-b, 3-a
	$H_0: \sigma^2 = \sigma_0^2 \leftrightarrow H_0: \frac{\sigma^2}{\sigma_0^2} = 1$ HIPÓTESIS NULA
	$H_a: \frac{\sigma^2}{\sigma_0^2} \neq 1$ $H_a: \frac{\sigma^2}{\sigma_0^2} > 1$ $H_a: \frac{\sigma^2}{\sigma_0^2} < 1$ **HIPÓTESIS ALTERNATIVA**
	$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$ **VARIANTE ESTADÍSTICA** *Chi-cuadrado* con *n* — 1 grados de libertad
3	Respuesta: A La hipótesis que se establece para este caso es $\left\{H_o: \frac{\sigma^2}{\sigma_o^2} = \frac{\sigma^2}{20^2} = 1 \; H_a: \frac{\sigma^2}{20^2} \neq 1\right\}$
4	Respuesta: B Aceptamos que $\frac{\sigma^2}{8} = 1$, por lo tanto se puede admitir que la varianza anterior era 8, al nivel de significación del 5%.
5	Respuesta: A La desviación estándar se simbolizará por σ_z y se calcula así: $\sigma_z = \frac{1}{\sqrt{n-3}}$

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

Autoevaluación 4

Pregunta	Respuestas y retroalimentación
6	Respuesta: B

Tabla V. transformación de *r A z (es decir,* $z=0,5\ Log_e \frac{1+r}{1-r}$ *)*

r	z	r	z	r	z
.00	.000	.36	.377	.71	.887
.01	.010	.37	.388	.72	.908
.02	.020	.38	.400	.73	.929
.03	.030	.39	.412	.74	.950
.04	.040	.40	.424	.75	.973
.05	.050				
.06	.060	.41	.436	.76	.996
.07	.070	.42	.448	.77	1.020
.08	.080	.43	.460	.78	1.045
.09	.090	.44	.472	.79	1.071
.10	.100	.45	.485	.80	1.099

7 Respuesta: C

Es una distribución de probabilidad continua con un parámetro “k”, que representa los grados de libertad de la variable aleatoria, la distribución de Chi-Cuadrado es denotada por la letra griega χ^2, es frecuentemente usada para probar hipótesis, concernientes a la diferencia entre un conjunto de frecuencias observadas de una muestra y un conjunto correspondientes de frecuencias teóricas esperadas.

Propiedades

- No toma valores negativos.
- La gráfica de la función toma una curva sesgada a la derecha-
- A medida que aumentan los grados de libertad la curva va tomando simetría.

8 Respuesta: A

El proceso para aplicar la prueba ***Chi-cuadrado*** se recomienda los siguientes pasos:

Formular la hipótesis

Establecer la diferencia entre las frecuencias observadas y las esperadas.

Elevar cada diferencia al cuadrado y dividirlas para su frecuencia teórica o esperada. (una a una) para sumarlas.

Calcular χ^2

9 Respuesta: B

Hipótesis nula

$H_o : n_i = n^*_i, a = 0,05$

Autoevaluación 4	
Pregunta	**Respuestas y retroalimentación**
10	Respuesta: A χ^2 puede tener valores negativos F Si se tiene 4 filas y 3 columnas, los grados de libertas son 5 F La suma de los valores esperados es igual a la suma de los valores observados V

Ir a la autoevaluación

5. Glosario

Distribución normal estándar

Es una distribución normal de probabilidad con $\mu = 0$ y $\sigma = 1$, y el área total debajo de su función de densidad es igual a 1.

Hipótesis

En estadística, una hipótesis es una afirmación o aseveración acerca de una propiedad de una población.

Prueba de hipótesis

Es un procedimiento para someter a prueba una afirmación acerca de una propiedad de una población.

Distribución de probabilidad

Es una distribución que indica la probabilidad de cada valor de la variable aleatoria. A menudo se expresa como gráfica, tabla o fórmula.

Estadístico

Es una medida cuantitativa, derivada de un conjunto de datos de una muestra, con el objetivo de estimar o inferir características de una población o modelo estadístico.

Parámetros

Se llama parámetros poblacionales a cantidades que se obtienen a partir de las observaciones de la variable y sus probabilidades y que determinan perfectamente la distribución de esta, así como las características de la población, por ejemplo: La media, μ, la varianza σ^2, la proporción de determinados sucesos, P.

Muestra aleatoria de tamaño n

Es un conjunto de n individuos tomado de tal manera que cada subconjunto de tamaño n de la población tenga la misma probabilidad de ser elegido como muestra; es decir, si la población tiene tamaño N, cada una de las combinaciones posibles de n elementos debe ser equiprobable.

Índice
Primer bimestre
Segundo bimestre
Solucionario
Glosario
Referencias bibliográficas

6. Referencias bibliográficas

Martínez, C., (2019). *Estadística y muestreo* (14a. ed.). Ecoe Ediciones. https://elibro.net/es/ereader/bibliotecautplecoe/125946.

Triola Mario, F. (2013). Estadística (13ª. ed.). Pearson.

Alvarado, Hugo; Galindo, Maritza; Retamal, Lidia. «Comprensión de la distribución muestral mediante configuraciones didácticas y su implicación en la inferencia estadística». Enseñanza de las ciencias, Vol. 31, Núm. 2 (2013), p. 75-91. DOI 10.5565/rev/ec/v31n2.803 <https://ddd.uab.cat/record/107306> [Consulta: 13 desembre 2020].

Mercado, M., (2013). Exploración de conceptos de probabilidad con Geogebra. Revista de didáctica de la Estadística. N°. 2, 309-327. https://dialnet.unirioja.es/servlet/articulo?codigo=4770324

Rodríguez, M., (2010). Importancia de la Distribución Binomial y de Poisson. Cap&Cua. Vol.3, N°.1,1-5. https://dialnet.unirioja.es/servlet/articulo?codigo=4020400

Inzunza, S. Islas, E. Análisis de una trayectoria de aprendizaje para desarrollar razonamiento sobre muestras, variabilidad y distribuciones muestrales. Educación matemática. Vol.31. N°3, 203-230 https://dialnet.unirioja.es/servlet/articulo?codigo=7504989

Inzunza, S. Jiménez, J. Caracterización del razonamiento estadístico de estudiantes universitarios acerca de las pruebas de hipótesis. Revista Latinoamericana de Investigación en Matemática Educativa. Vol.16. N°2, 179-211 http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-24362013000200003

Universidad Militar Nueva Granada-Facultad de Estudios a Distancia. (2018). Pruebas de hipótesis. Recuperado el 26 de enero de 2021. Enlace web

Camacho García, A. (2017). Distribución muestral de la proporción. Recuperado el 25 de enero de 2021 http://hdl.handle.net/10251/82125