La Universidad Católica de Loja

Vicerrectorado de Modalidad Abierta y a Distancia

Sistemas de Conocimiento de Estadística Inferencial y su Didáctica

Guía didáctica

UTPL

Facultad Ciencias Sociales, Educación y Humanidades

Sistemas de Conocimiento de Estadística Inferencial y su Didáctica

Guía didáctica

Carrera	PAO Nivel
Pedagogía de las Ciencias Experimentales (Pedagogía de las Matemáticas y la Física)	VI

Autora:

Sonia Patricia Granda Sivisapa

EDUC_3152

Universidad Técnica Particular de Loja

Sistemas de Conocimiento de Estadística Inferencial y su Didáctica

Guía didáctica

Sonia Patricia Granda Sivisapa

Diagramación y diseño digital

Ediloja Cía. Ltda.
Marcelino Champagnat s/n y París
edilojacialtda@ediloja.com.ec
www.ediloja.com.ec

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

Año de edición: marzo, 2021

Edición: primera edición reestructurada en julio 2025 (con un cambio del 1%)

Loja-Ecuador

Índice

1. Datos de información 9
- 1.1 Presentación de la asignatura 9
- 1.2 Competencias genéricas de la UTPL 9
- 1.3 Competencias del perfil profesional 9
- 1.4 Problemática que aborda la asignatura 10

2. Metodología de aprendizaje 12

3. Orientaciones didácticas por resultados de aprendizaje 13

Primer bimestre 13

Resultado de aprendizaje 1: 13

Contenidos, recursos y actividades de aprendizaje recomendadas 13

Semana 1 13
- Unidad 1. Distribuciones de probabilidad 14
  - 1.1. Conceptos y aplicaciones teórico-prácticas de las distribuciones de probabilidad 15
  - 1.2. Variable aleatoria discreta: Distribución Binomial 17
- Actividad de aprendizaje recomendada 18

Contenidos, recursos y actividades de aprendizaje recomendadas 19

Semana 2 19
- Unidad 1. Distribuciones de probabilidad 20
  - 1.3. Distribución probabilística de Poisson 20
- Actividades de aprendizaje recomendadas 24

Contenidos, recursos y actividades de aprendizaje recomendadas 25

Semana 3 25
- Unidad 1. Distribuciones de probabilidad 25
  - 1.4. Variable aleatoria continua: Distribución normal 26
- Actividades de aprendizaje recomendadas 32
- Autoevaluación 1 34

Contenidos, recursos y actividades de aprendizaje recomendadas 37

Semana 4 ..... **37**
Unidad 2. Distribuciones muestrales 38
2.1. Algunos conceptos sobre el muestreo. 38
2.2. Distribución muestral 40
2.3. Distribución de medias muestrales: $\bar{x}$ 41
Actividades de aprendizaje recomendadas 45
Contenidos, recursos y actividades de aprendizaje recomendadas **47**
Semana 5 ..... **47**
Unidad 2. Distribuciones muestrales 48
2.4. Distribución muestral de una proporción: ρ 48
Actividades de aprendizaje recomendadas 54
Contenidos, recursos y actividades de aprendizaje recomendadas **55**
Semana 6 ..... **55**
Unidad 2. Distribuciones muestrales 56
2.5. Distribución de diferencias entre las medias muestrales $\bar{x} - \bar{y}$ 56
2.6. Distribución de diferencias entre dos proporciones muestrales $P_1 - P_2$ 59
Actividades de aprendizaje recomendadas 63
Autoevaluación 2 64
Resultado de aprendizaje 2: **71**
Contenidos, recursos y actividades de aprendizaje recomendadas **71**
Semana 7 ..... **71**
Unidad 3. Pruebas de hipótesis 72
3.1. Conceptos generales 72
3.2. Prueba unilateral y bilateral 73
3.3 Distribución de medias muestrales $(\bar{x})$ 74
Actividades de aprendizaje recomendadas 76
Resultado de aprendizaje 1 y 2: **78**
Contenidos, recursos y actividades de aprendizaje recomendadas **78**

Semana 8 .... **78**
Actividades finales del bimestre 78
Actividades de aprendizaje recomendadas 79
Segundo bimestre **80**
Resultado de aprendizaje 2: **80**
Contenidos, recursos y actividades de aprendizaje recomendadas **80**
Semana 9 .... **80**
Unidad 3. Pruebas de hipótesis 81
3.4. Paquete estadístico SPSS Statistical Package for the Social Sciences/ Paquete estadístico para ciencias sociales. 81
Actividades de aprendizaje recomendadas 84
Contenidos, recursos y actividades de aprendizaje recomendadas **86**
Semana 10 .... **86**
Unidad 3. Pruebas de hipótesis 86
3.5. Distribución de proporciones muestrales ρ 86
3.6. Distribución de diferencias entre dos medias muestrales $\bar{x} - \bar{y}$ 90
Actividades de aprendizaje recomendadas 93
Contenidos, recursos y actividades de aprendizaje recomendadas **94**
Semana 11 .... **94**
Unidad 3. Pruebas de hipótesis 95
3.7. Distribución de diferencias entre dos proporciones muestrales: $P_1 - P_2$ 95
3.8. Distribución t Student 97
3.9. Distribución de medias muestrales. 98
Actividades de aprendizaje recomendadas 101
Contenidos, recursos y actividades de aprendizaje recomendadas **103**
Semana 12 .... **103**
Unidad 3. Pruebas de hipótesis 103
3.10. Distribución de una proporción muestral 103

3.11. Distribución de diferencias entre dos medias muestrales $\bar{x} - \bar{y}$. 107
Actividades de aprendizaje recomendadas 109
Autoevaluación 3 111
Resultado de aprendizaje 3: 120
Contenidos, recursos y actividades de aprendizaje recomendadas........ 120
Semana 13 120
Unidad 4. Otras Pruebas de hipótesis........ 121
4.1. Prueba de hipótesis de una varianza 121
Actividades de aprendizaje recomendadas 128
Contenidos, recursos y actividades de aprendizaje recomendadas........ 129
Semana 14 129
Unidad 4. Otras Pruebas de hipótesis........ 130
4.2. Comparación entre varianzas de dos poblaciones........ 130
4.3. Prueba del coeficiente de correlación de Pearson r = R........ 133
Actividades de aprendizaje recomendadas 136
Resultado de aprendizaje 4: 138
Contenidos, recursos y actividades de aprendizaje recomendadas........ 138
Semana 15 138
Unidad 4. Otras Pruebas de hipótesis........ 138
4.4. Pruebas con observaciones apareadas. 138
4.5. Pruebas no paramétricas 144
Actividades de aprendizaje recomendadas 146
Autoevaluación 4 148
Resultado de aprendizaje 2 a 4:........ 154
Contenidos, recursos y actividades de aprendizaje recomendadas........ 154
Semana 16 154
Actividades finales del bimestre 154
Actividades de aprendizaje recomendadas 155
4. Autoevaluaciones 156

5. Glosario.. 169
6. Referencias bibliográficas 171

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 del perfil profesional

- Diseñar, ejecutar, evaluar y orientar secuencias didácticas con elementos pedagógicos y curriculares orientados a los campos de la matemática y la física mediante la fundamentación teórico-práctico de los sistemas de conocimiento que, faciliten la adaptación a los cambios permanentes de la realidad actual y de un mundo globalizado.
- Identificar, diseñar e integrar los sistemas de conocimiento de la física y la matemática relacionados con el entorno natural y social de los estudiantes, aplicando metodologías y didácticas específicas que faciliten la

contextualización de estas áreas con la realidad de un mundo globalizado y cambiante.

- Seleccionar, adaptar y aplicar herramientas tecnológicas apropiadas para el desarrollo de metodologías activas e innovadoras que faciliten la ejecución del proceso de enseñanza aprendizaje mediante talleres práctico-experimentales permanentes, empleando contenidos contextualizados a la realidad estudiantil, nacional y mundial.
- Seleccionar, adaptar, construir y aplicar criterios, indicadores, técnicas e instrumentos de evaluación idóneos para los procesos de enseñanza aprendizaje de la matemática y la física, considerando diferencias individuales, interculturales e inclusivas; integrando adecuadamente los elementos curriculares, conocimientos, estrategias y metodologías en función de la realidad natural y social del estudiante.
- Diseñar, ejecutar y evaluar modelos pedagógicos y de organización escolar para brindar soluciones a las diferencias individuales, interculturales e inclusivas, mediante la adaptación de los elementos curriculares y contenidos con estrategias y metodologías adaptadas a la realidad de la comunidad.
- Elaborar, ejecutar y evaluar proyectos y/o procesos de investigación que conlleven a la recopilación, organización y análisis de información en el ámbito de las matemáticas y la física enfocados a la generación de nuevos conocimientos, habilidades y actitudes que aporten a la solución de problemas prácticos de su comunidad.
- Desarrollar, ejecutar y difundir proyectos pedagógicos y didácticos con metodologías activas e innovadoras, involucrando la matemática y la física, vinculados a la solución de problemas de la realidad y que apoyen la integración de los docentes con el entorno natural y social de la comunidad y del país en general.

1.4 Problemática que aborda la asignatura

La estadística inferencial, como parte fundamental de las matemáticas, tiene el propósito de analizar el comportamiento de sucesos o eventos mediante el uso de datos obtenidos de una muestra para realizar inferencias sobre una población más amplia. Este proceso permite representar y comprender la realidad, facilitando la toma de decisiones en diferentes contextos investigativos. Sin embargo, no basta con recopilar y calcular datos; lo esencial es aplicar correctamente métodos que permitan formular generalizaciones válidas y confiables, lo cual es el núcleo de esta disciplina. El principal desafío para los investigadores radica en adquirir las habilidades necesarias para elegir y aplicar de manera adecuada pruebas estadísticas paramétricas y no paramétricas, herramientas indispensables en el análisis inferencial. La falta de conocimiento en estos métodos puede limitar la calidad de las conclusiones obtenidas y, por ende, afectar la toma de decisiones basada en datos. Por ello, esta asignatura busca abordar dicha problemática, formando profesionales capaces de realizar análisis estadísticos sólidos y eficientes.

2. Metodología de aprendizaje

El estudio de la asignatura **Sistemas de Conocimientos de Estadística Inferencial y su Didáctica** busca fomentar un aprendizaje permanente, permitiendo a los estudiantes aplicar técnicas de muestreo, métodos de estimación paramétrica y pruebas estadísticas en casos de estudio. Para ello, se implementan actividades basadas en una **metodología activa**, donde los estudiantes acceden inicialmente a la conceptualización teórica a través de microvideos y recursos disponibles en el EVA. Posteriormente, el trabajo con el docente se desarrolla en talleres prácticos, consolidando el aprendizaje mediante la **experimentación.** Además, se realizan videoconferencias semanales, que serán grabadas para quienes no puedan asistir en vivo.

Si bien no existe un procedimiento único, el curso se apoya en la bibliografía recomendada, la guía didáctica y recursos educativos abiertos. Herramientas como cuestionarios interactivos e infografías facilitan la comprensión de los aspectos teóricos antes de ser reforzados con la práctica experimental. La participación del profesor se da de manera síncrona en algunos casos y, principalmente, a través de encuentros asincrónicos, promoviendo el trabajo colaborativo y adaptándose a distintos estilos de aprendizaje para el desarrollo de competencias clave.

3. Orientaciones didácticas por resultados de aprendizaje

Primer bimestre

Resultado de aprendizaje 1:

Aplica las técnicas de muestreo en base al caso de estudio. Identifica métodos de estimación de parámetros y su interpretación en casos de estudio.

Para alcanzar el resultado de aprendizaje propuesto, usted desarrollará la capacidad de aplicar técnicas de muestreo y métodos de estimación paramétrica en el análisis de casos de estudio. Esto le permitirá seleccionar muestras representativas, interpretar datos de manera precisa y realizar inferencias estadísticas fundamentadas. A través de este aprendizaje, adquirirá habilidades esenciales para la toma de decisiones basadas en evidencia, optimizando la precisión y confiabilidad en distintos contextos de aplicación.

Contenidos, recursos y actividades de aprendizaje recomendadas

Recuerde revisar de manera paralela los contenidos con las actividades de aprendizaje recomendadas y actividades de aprendizaje evaluadas.

Semana 1

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. Por esta razón, es fundamental que complete todas las actividades establecidas en el **plan docente**, apoyándose en la bibliografía sugerida, los recursos abiertos en línea y esta guía didáctica. Esto le permitirá fortalecer sus conocimientos y prepararse adecuadamente para sus evaluaciones.

Por otra parte, se lleva a cabo una revisión sistemática de los conocimientos previos en matemática especialmente en estadística descriptiva. Este proceso, 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.

Unidad 1. Distribuciones de probabilidad

La distribución de probabilidad es una herramienta fundamental en la estadística, ya que permite representar de manera matemática y gráfica el comportamiento de un conjunto de datos o sucesos aleatorios. Esta unidad tiene como objetivo principal desarrollar la comprensión y aplicación de las distribuciones más comunes, como la binomial y la normal, para modelar fenómenos en distintos contextos. A lo largo del estudio, se explorará cómo estas distribuciones describen la probabilidad de ocurrencia de diferentes resultados y cómo sus propiedades facilitan el análisis de datos y la toma de decisiones. Además, se utilizarán recursos interactivos y visuales para comprender la relación entre las características teóricas de estas distribuciones y su representación gráfica, consolidando así el aprendizaje a través de ejemplos prácticos.

Por ello, comenzaremos abordando conceptos clave que complementan lo aprendido previamente en estadística descriptiva, enfocados en la recolección de datos, la organización en tablas o cuadros, la representación gráfica y el cálculo de medidas de posición y dispersión. Estos fundamentos serán la base para profundizar en el análisis y la interpretación de distribuciones de probabilidad.

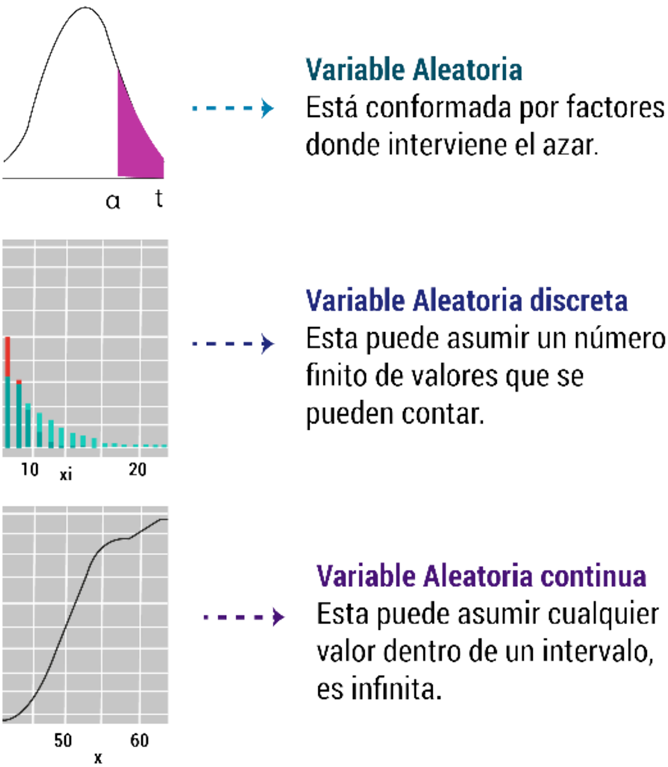
¡Bienvenido a esta fascinante unidad de estudio!

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

Comenzaremos abordando conceptos clave que complementan lo aprendido previamente en estadística descriptiva, enfocados en la recolección de datos, la organización en tablas o cuadros, la representación gráfica y el cálculo de medidas de posición y dispersión. Estos fundamentos serán la base para profundizar en el análisis y la interpretación de distribuciones de probabilidad.

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

Figura 1
Tipos de variables



Nota. Adaptado de *Estadística y muestreo* (14a. ed.) [Ilustración], por Martínez, C., 2019, Ecoe Ediciones, CC BY 4.0.

La figura 1 muestra los tipos de variables necesarias para el aprendizaje de la distribución de probabilidades.

Para diferenciar la distribución de frecuencias y la distribución de probabilidad, se recomienda consultar el capítulo 2 del texto Probabilidad y Estadística (Obando & Arango, 2019), disponible en la biblioteca virtual de la universidad.

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 anterior es.

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 $(nx) = (n\ n - x) = (\frac{n!}{(n-x)!n!})$ 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 en la siguiente presentación interactiva para consolidar su comprensión y facilitar su aplicación práctica.

Problema Modelo #1: Distribución Binomial

Actividad de aprendizaje recomendada

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

1. Observar el video ¿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, ¿cuál es la probabilidad de que como máximo 6 ganen?

   Nota: por favor complete la actividad en un cuaderno o documento Word.

RETROALIMENTACIÓN

Los videos 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 caso adicional al explicado en el problema modelo #1, donde para resolverlo utilizamos: $n = 8, x = 0, 1, 2, 3, 4, 5, 6,\ p = 0.8, q = 0,2$ y $P(x \leq 6)$ $P(0 \leq x \leq 6)$ y la probabilidad se determina utilizando el

complemento $P(x \leq 6) = 1 — [P(x = 7) + P(x = 8)]$. En caso de haber existido alguna dificultad, es necesario que revise el texto Probabilidad y estadística de Islas, Colín y Morales (2018), disponible en la biblioteca virtual de la universidad.

Con el contenido desarrollado, concluimos la actividad planificada para la primera semana. Agradezco profundamente su compromiso, entusiasmo y participación activa en cada etapa. Lo animo a mantener este mismo nivel de dedicación y esfuerzo en las actividades de la próxima semana.

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

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.

Contenidos, recursos y actividades de aprendizaje recomendadas

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 GeoGebra 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 específico. 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: $P(x) = \frac{\lambda^x e^{-\lambda}}{x!}$ 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 v e 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:

Resolución Problema Modelo 2: Distribución Probabilidad de Poisson

A continuación, 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 (Martínez, 2019, elaborado por Granda, 2020).

Determinamos la media: $\lambda = 10.000(0.001) = 1$ y el número de casos favorables $x = 5$. Recuerde $e = 2.71828$

Tomamos la función de probabilidad

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

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

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

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

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

Expresado como porcentaje:

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

Conclusión:

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

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 del 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:

Resolución Problema Modelo 3: Distribución de Probabilidad de Poisson

A continuación, se muestra el proceso de aplicación de la distribución de probabilidad de Poisson al determinar la probabilidad de concurrencia a una cafetería (Martínez, 2019, elaborado por Granda, 2020).

a. Determinamos la media: λ = 9 en 30 min y el número de casos favorables $x \geq 4$. Recuerde $e = 2.71828$

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

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

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

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

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

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

Expresado como porcentaje:

$$P(X \geq 4)98.3\%$$

Conclusión:

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 $\frac{30 \ minutos}{10 \ minutos} = 3$ y el valor de la media se ve afectada por dicho intervalo $\lambda = \frac{9}{3} = 3$

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

$$P(x = 0) = \frac{3^0 \cdot e^{-3}}{0!} = 0.0497 = 4.97\%$$

Conclusión:

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

Actividades de aprendizaje recomendadas

Al terminar la segunda semana, se ha abordado la definición, aplicación e importancia de la distribución de Poisson. Para afianzar estos conocimientos, se recomienda revisar el capítulo 6 del texto Estadística y muestreo de Martínez (2018), disponible en la biblioteca virtual de la universidad. Finalmente, 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 los procesos expuestos en la herramienta GeoGebra: Distribución Probabilística de Poisson, donde se resuelven problemas de aplicación relacionados con el tema estudiado esta semana.

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

4. Resolver los problemas propuestos en la herramienta GeoGebra: Distribución Probabilística de Poisson

   Nota: Por favor complete la actividad en un cuaderno o documento Word.

5. Revise el sitio web de OpenStax, específicamente la sección 4.4 sobre la Distribución de Poisson. Interactúe con la teoría y los ejemplos propuestos para enriquecer su comprensión del tema.

RETROALIMENTACIÓN

Al resolver las actividades planteadas, usted pondrá en práctica su habilidad de análisis, comparando los diferentes tipos de distribuciones y evaluando cómo se ven afectadas por el tamaño de la muestra y las variables involucradas. Asimismo, podrá realizar operaciones de manera rápida y confiable, apoyándose en el uso de la tecnología.

El video incluido explica detalladamente el proceso para trabajar con la distribución de Poisson, identificando cuándo es conveniente aplicarla y evaluando la pertinencia de los resultados obtenidos. Al concluir esta semana de aprendizaje sobre la distribución de Poisson, le animo a continuar explorando el fascinante mundo de las distribuciones de probabilidad.

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

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

Crea en usted mismo y todo será posible

Contenidos, recursos y actividades de aprendizaje recomendadas

Semana 3

Unidad 1. Distribuciones de probabilidad

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 conoce, más comúnmente, como la "**campana de Gauss**". (Pertegás y Pita ,2001, p.1).

A continuación, se presenta una breve definición de la distribución normal. No obstante, recomiendo explorar el sitio web OpenStax, sección 6, donde se detalla ampliamente la definición y los elementos característicos de esta distribución.

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.

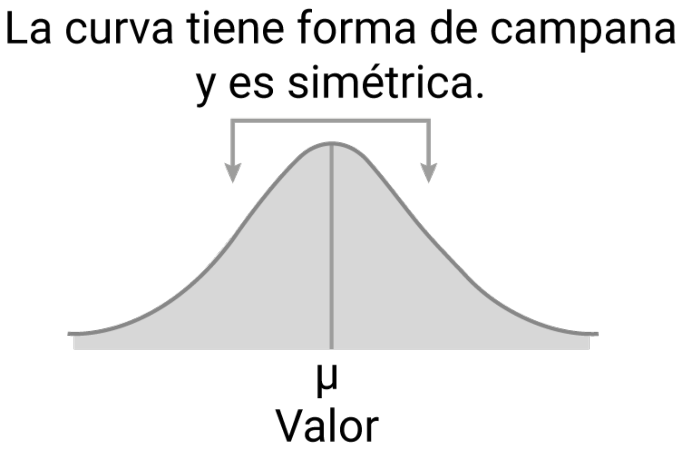
A continuación, se presenta la definición de la distribución normal. No obstante, se recomienda consultar el libro Estadística y muestreo de Martínez (2018), disponible en la biblioteca virtual de la universidad, donde en la página 247 encontrará una explicación detallada sobre su definición y elementos.

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. La siguiente figura explica esta distribución para una variable aleatoria continua.

Figura 2

Definición de Distribución Normal



Nota. Tomado de *Estadística (13ª. ed.)*(p. 246)[Ilustración], por Triola, M., 2013, Pearson Educación, CC BY 4.0.

La distribución normal estándar tiene las siguientes tres propiedades (Triola, 2013, p. 251):

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, $\sigma = 1$)

Estudiaremos la curva normal probabilística, por lo que resulta fundamental revisar detenidamente la unidad 2 en el texto [Estadística inferencial 1 para ingeniería y ciencias]y la herramienta GeoGebra: Distribución Normal, donde se explican las condiciones que esta curva debe cumplir. Una vez que haya revisado dichas condiciones, le invito a poner en práctica sus conocimientos resolviendo los siguientes problemas modelo:

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 $yZ = 1.87; c)Z =—2.35$ y $Z =—0.50$.

Solución:

Resolución Problema Modelo 4: Ejercicios del área bajo la curva normal

Para resolver estos problemas, es necesario saber cómo usar la tabla de distribución normal. Para ello, debe ingresar al video proporcionado y observar detenidamente la explicación sobre el uso de una tabla de distribución normal, como la tabla 6.9 del texto Estadística y muestreo de Martínez (2018).

Es fundamental comprender que la probabilidad es igual al área bajo la curva y que al usar la figura 3 correspondiente a Áreas bajo la curva normal, ésta tomará a valores del centro hacia la derecha y del centro hacia la izquierda. Así, $P_{(x)} = A$

A continuación, se muestra el proceso para determinar el área bajo la curva normal, en diferentes casos del valor de Z (Martínez, 2019, elaborado por Granda, 2020).

a. $Z = -1,20$ y $Z = 2,40$

Revisando en la figura para Z tenemos:

$$P(-1.20 \leq Z \leq 2.40) = 0.3849 + 0.4918 = 0.8767$$

Figura 3

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	0.1063	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	0.2513	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	0.2935	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	0.3264	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3486	0.3503	0.3508	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	0.4756	0.4767
2.0	0.4773	0.4778	0.4778	0.4788	0.4793	0.4798	0.4808	0.4808	0.4808	0.4817
2.1	0.4821	0.4826	0.4862	0.4834	0.4838	0.4842	0.4850	0.4850	0.4850	0.4857

Nota. Tomado de *Estadística y muestreo: 13 ed. (p. 247)* [Ilustración], por Martínez, C., 2018, Ecoe Ediciones, CC BY 4.0.

b. $Z = 1,23$ y $Z = 1.87$

Revisando en la figura para Z tenemos:

$$P(1.23 \leq Z \leq 1.87) = 0.4693 + 0.3907 = 0.078$$

c. $Z = -2.35$ y $Z = -0.50$

Revisando en la figura para Z tenemos:

$$P(—2.35 \leq Z \leq 0.50) = 0.4906—0.1915 = 0.299$$

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 figura 3 que presenta las 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:

Resolución Problema Modelo 5: Aplicación del área bajo la curva normal

A continuación, 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 (Martínez, 2019, elaborado por Granda, 2020).

a. Tomamos los datos e incógnitas del problema

$\mu = 169$

$\sigma = 3$

$P(x < 165) = ?$

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 del texto Estadística y muestreo de Martínez (2018), disponible en la biblioteca virtual de la universidad, 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 $\mu = 0$ 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 tomó de la figura, 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:

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

$$P(x < 165) = 0.0918 \times 100\% = 9.18\%$$

Conclusión:

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

b. ¿Qué porcentaje de alumnos tendrán una estatura entre 1,65 y 1,70?

$$P(165 < x < 170) =?$$

En este caso, determinamos dos puntuaciones z

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

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

Como las áreas encontradas están a la derecha e izquierda de la $\mu = 0$, 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:

$P(165 < x < 170) = 0.5375x100$

Conclusión:

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

Actividades de aprendizaje recomendadas

Hemos concluido el estudio correspondiente a la semana 3, durante el cual se analizaron las condiciones que debe cumplir la curva normal o probabilística para interpretar la distribución normal de datos en una población determinada. Le recordamos revisar detenidamente la bibliografía y los recursos en línea sugeridos, donde encontrará explicaciones detalladas sobre el tema de esta semana, así como ejemplos de problemas resueltos y propuestos. Una vez finalizado este análisis, le invitamos a realizar las siguientes actividades para consolidar su aprendizaje.

1. Analizar el video denominado Distribución Normal de Probabilidades-Campana de Gauss.
2. Resolver la Práctica de la sección 6 del sitio web Openstax.
3. Revisar la definición y características de la Distribución Normal y luego, contestar el siguiente cuestionario, seleccionando el valor de verdad que corresponda en la tabla.

Condiciones de la curva normal o probabilística para completar

Condiciones de la curva normal	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 $\mu_z = 0$ y $\sigma_z = 1$	

Nota. Por favor complete las actividades en un cuaderno o documento Word.

RETROALIMENTACIÓN

En los materiales sugeridos se explican las principales características de la Distribución Normal de Probabilidades, así como los conceptos necesarios para abordar y resolver problemas prácticos relacionados con esta distribución. Estos materiales le servirán de apoyo para enfrentar los ejercicios planteados.

Al completar este cuestionario, tendrá la oportunidad de autoevaluar sus conocimientos sobre las condiciones esenciales de la distribución normal, las cuales son fundamentales para resolver problemas de aplicación.

Con esto, damos por finalizado el estudio de la distribución normal y nos preparamos para avanzar hacia nuevos aprendizajes.

¡Felicitaciones por su esfuerzo activo y compromiso responsable! Continúe avanzando con la misma dedicación, ¡el éxito está a su alcance!

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

Trabajar duro le llevará a la cima, disfrutar el camino le llevará más lejos.

4. 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.

Autoevaluación 1

Instrucción: Marque la alternativa correcta.

1. Dentro de los criterios para satisfacer una experiencia binomial, ¿qué 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 fija.

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 que 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, 35 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 II. Variable Aleatoria discreta III. Variable Aleatoria continua	a. Esta puede asumir cualquier valor dentro de un intervalo, es infinita b. Está conformada por factores donde interviene el azar. 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

Ir al solucionario

Le recuerdo que dispone del horario de tutoría, durante el cual su profesor tutor estará disponible para atender sus dudas e inquietudes. Este espacio está diseñado para ayudarle a resolver las dificultades que puedan surgir durante la semana de trabajo y para validar su desempeño académico. Además, le animo a participar activamente en las actividades síncronas planificadas a lo largo de la asignatura. Estas sesiones representan una excelente oportunidad para superar barreras como el tiempo y la distancia, y para optimizar los resultados de su aprendizaje. ¡Aproveche al máximo estas herramientas para su desarrollo académico!

Contenidos, recursos y actividades de aprendizaje recomendadas

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 7 del texto Estadística y muestreo de Martínez (2018), disponible en la biblioteca virtual de la universidad. 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 GeoGebra 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!

2.1. Algunos conceptos sobre el muestreo.

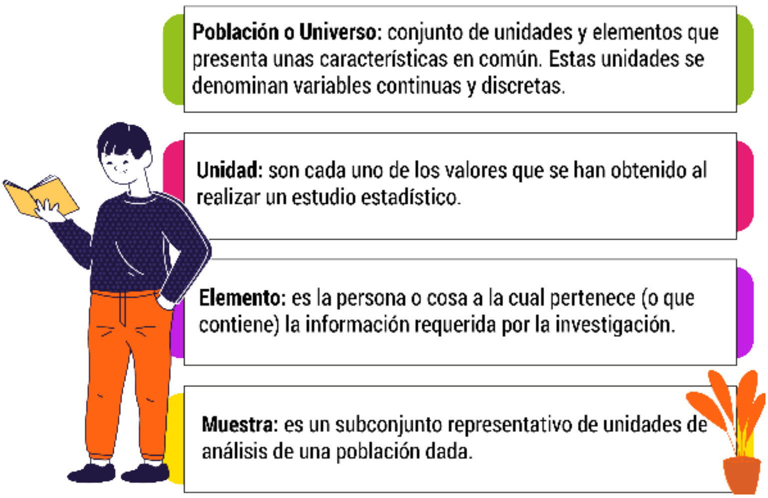
En esta unidad nos ocuparemos de la inferencia estadística o método inductivo, considerando que dicha inferencia es el conjunto de métodos que permite inducir a través de una muestra estadística el comportamiento de una determinada población, lo que nos permitirá extraer conclusiones sobre los parámetros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

Comenzaremos revisando algunos conceptos fundamentales sobre muestreo. No obstante, se recomienda ampliar estas nociones mediante una lectura comprensiva del capítulo 1 del texto Estadística Inferencial de Llinás (2017). En este capítulo se presenta el marco teórico que sustenta esta unidad, el cual he sintetizado en los esquemas conceptuales que se detallan a continuación.

La figura 4 explica los conceptos básicos de población, unidad, elemento y muestra fundamentales en el estudio de las distribuciones muestrales.

Figura 4

Definiciones básicas



Nota. Adaptado de *Estadística y muestreo (14a. ed.)* [Ilustración], por Martínez, C., 2019, Ecoe Ediciones, CC BY 4.0.

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)

Métodos del muestreo aleatorio

El muestreo aleatorio es un método de selección de muestras en el que cada elemento de la población tiene la misma probabilidad de ser elegido. Este tipo de muestreo se basa en la aleatoriedad, lo que garantiza que la muestra sea representativa de la población y minimiza los sesgos en el proceso de selección. Existen diferentes tipos de muestreo aleatorio, como el simple, sistemático, estratificado y por conglomerados, cada uno diseñado para ajustarse a las características específicas del estudio y de la población en cuestión.

La teoría del muestreo se desarrolla en el capítulo 13 del libro Estadística y Muestreo de Martínez (2018), disponible en la biblioteca virtual de la universidad, específicamente en las páginas 663-666. En este capítulo se define el concepto de muestreo aleatorio, cuya esencia se ha sintetizado en la infografía que se presenta a continuación.

Métodos de muestreo aleatorio

La infografía presentada señala la utilidad de estos métodos para generalizar conclusiones a partir de una muestra.

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).

Para profundizar en el tema, es importante revisar los **métodos de selección de unidades.** Estos se explican detalladamente en el capítulo 13 en la página 668 del libro Estadística y Muestreo de Martínez (2018), disponible en la biblioteca virtual de la universidad. Le invito a consultar este libro para estudiar y comprender a fondo los diferentes métodos de selección de unidades al azar, descritos en la siguiente figura.

Figura 5

Métodos de selección de unidades .

Nota. Adaptado de *Estadística y muestreo (14a. ed.)* [Ilustración], por Martínez, C., 2019, Ecoe Ediciones, CC BY 4.0.

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 resultará orientador observar el video 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 la resolución de los problemas modelo de la distribución muestral de la media utilizaremos la siguiente simbología, la cual se muestra en la tabla:

Tabla 1

Simbología estadística

Medidas	Población	Muestra	Definiciones y fórmulas	
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. Adaptado de *Estadística y muestreo* (14a. ed.), por Martínez, C., 2019, Ecoe Ediciones.

Completemos 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).

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 cuenta a Martínez (2019), quien señala que en los casos de poblaciones finitas se puede aplicar el factor de corrección, el cual se presenta de muchas formas que se resumen en la expresión dada a continuación. Este 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.

PROBLEMA MODELO #6

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.

Solución:

Solución Problema Modelo 6: Distribución de medias muestrales

A continuación, se explica el proceso para determinar una probabilidad empleando distribución de medias muestrales:

1. Anotamos los datos e incógnitas del problema:

$\bar{x} = 71.7$

$\sigma = 3.1$

$\mu = 72.1$

$n = 90$

$P(\underline{x} < 71.7) = ?$

2. Calculamos la probabilidad solicitada.

$$z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{71.7-72.1}{\frac{3.1}{\sqrt{90}}} = \frac{-0.4}{\frac{3.1}{9.48}} = \frac{-0.4}{0.33} = -1.22$$

Con el valor de $Z =— 1,22$ buscamos el área bajo la curva por medio de la figura 3 Áreas de una distribución normal ordinaria.

$$Z =— 1,22 \rightarrow A(0.3888)$$

Determinamos la probabilidad requerida:

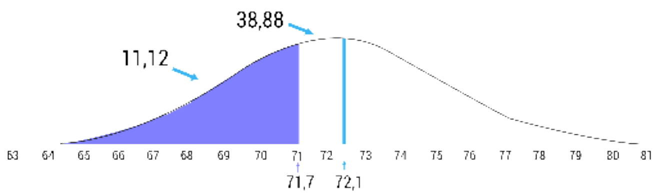
$$P(\bar{x} < 71.7) = A(0.5000) — A(0.3888) = 0.1112$$

$$P(\bar{x} < 71.7) = 0.1112 \times 100\% = 11.12\%$$

3. Finalmente, la solución gráfica se ilustra en la siguiente figura.

Figura 6

Solución gráfica del problema de distribución de medias muestrales



Nota. Adaptado de Estadística y muestreo (14a. ed.) [Ilustración], por Martínez, C., 2019, Ecoe Ediciones, CC BY 4.0.

Al finalizar la cuarta semana de trabajo, se concluye el estudio de las distribuciones muestrales, habiendo abordado la revisión de los conceptos básicos sobre muestreo, la definición de la distribución muestral y la distribución de medias muestrales. Le recomiendo analizar detenidamente la sección 7 del sitio web OpenStax, titulada El teorema del límite central. En este material encontrará explicaciones fundamentales para comprender el tema de esta semana, así como ejemplos de problemas resueltos y propuestos.

Actividades de aprendizaje recomendadas

Luego de esta revisión y análisis, lo invito a desarrollar las siguientes actividades de aprendizaje:

1. Observar atentamente el video titulado Inferencia Estadística/ Distribución muestral de medias y proporciones.
2. Responda el siguiente cuestionario indicando el valor de verdad que corresponda a cada enunciado. Para ello, utilice como referencia la presente guía didáctica, así como la bibliografía y los recursos sugeridos.

Conceptos de distribución muestral para completar

Conceptos de 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.	

Nota: Copie la tabla en un Word o cuaderno para rellenar.

RETROALIMENTACIÓN

En el video se explican las definiciones clave de la inferencia estadística, acompañadas de ejercicios que ilustran la aplicación de los conceptos relacionados con la distribución muestral de la media.

Al responder el cuestionario, tendrá la oportunidad de autoevaluar sus conocimientos sobre la fundamentación teórica presentada en esta guía, así como en la bibliografía y los recursos sugeridos. Estas definiciones son esenciales para resolver problemas relacionados con las distribuciones muestrales, que serán el eje de estudio en las próximas semanas.

Hemos concluido exitosamente la semana de trabajo, avanzando en el análisis de las distribuciones muestrales, incluida la distribución de medias muestrales. Ahora, estamos listos para continuar con nuevos aprendizajes.

¡Felicitaciones por el esfuerzo y el trabajo realizado hasta esta semana! Su dedicación fortalece su proceso de aprendizaje y lo impulsa hacia el éxito.

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

"**Da siempre lo mejor de ti. Lo que plantes ahora, lo cosecharás más tarde**."
— O.G. Mandino

Contenidos, recursos y actividades de aprendizaje recomendadas

Semana 5

En esta semana continuaremos con el estudio de las distribuciones muestrales, centrándonos específicamente en la **aplicación de la distribución muestral de una proporción**: p. Para una mejor comprensión, se recomienda consultar el capítulo 7 del libro Estadística y muestreo de Martínez (2018), donde se ofrece una explicación detallada sobre este tema.

Es fundamental que el estudiante comprenda que, al incrementar el tamaño de la muestra, se obtienen estimaciones más precisas de la probabilidad y de las características poblacionales. Además, es importante destacar que las mediciones repetidas de un mismo fenómeno constituyen un proceso aleatorio, lo que puede generar resultados diferentes, y que las inferencias están influenciadas por la muestra seleccionada (Inzunza, 2019). Con estas consideraciones en mente, le invito a analizar de manera crítica el contenido que se presentará a continuación.

¡Bienvenido!

2.4. Distribución muestral de una proporción: ρ

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 video 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 de la muestra (n)

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

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, presentada en la tabla. Observe detenidamente cada símbolo y su significado, ya que será clave para la comprensión y aplicación de los conceptos abordados.

Tabla 2

Simbología estadística para proporciones muestrales y poblacionales

Fórmula matemática	Descripción y notación
$p = \frac{a}{n} = \frac{\sum a_i}{a_n}$	Proporción muestral $\longleftrightarrow \underline{p}$
$p = \frac{A}{N}$	Proporción poblacional $\longleftrightarrow \mu p = \underline{P}$
$\sigma_p^2 = PQ$	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

Nota. Adaptado de *Estadística y muestreo* (p. 289), por Martínez, C. (2019), Ecoe Ediciones.

Es importante revisar la sección 2.3.3 Distribución de la proporción muestral del texto Estadística inferencial aplicada (2.a ed.) de Díaz (2022), disponible en la biblioteca virtual de la universidad. En esta sección encontrará la simbología y la demostración de las expresiones necesarias para determinar la distribución muestral de una proporción (p). Además, le invito a consultar la resolución de los siguientes problemas modelo, donde se presenta la aplicación práctica del tema estudiado 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 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:

Solución Problema Modelo 7: Distribución muestral de una proporción

A continuación, 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.

1. Anotamos los datos e incógnitas del problema:

$$P = 0,01$$

$$n = 400$$

$$P_{(p>0,02)} = ?$$

2. Calculamos la probabilidad solicitada.

$$Z = \frac{p-P}{\sqrt{\frac{PQ}{n}}} = \frac{0.02-0.01}{\sqrt{\frac{0.01(0.99)}{400}}}$$

Con el valor de $Z = 2,01$ buscamos el área bajo la curva por medio de la figura 3 Áreas de una distribución normal ordinaria.

$Z = 2,01 \rightarrow A(0.4778)$

Determinamos la probabilidad requerida:

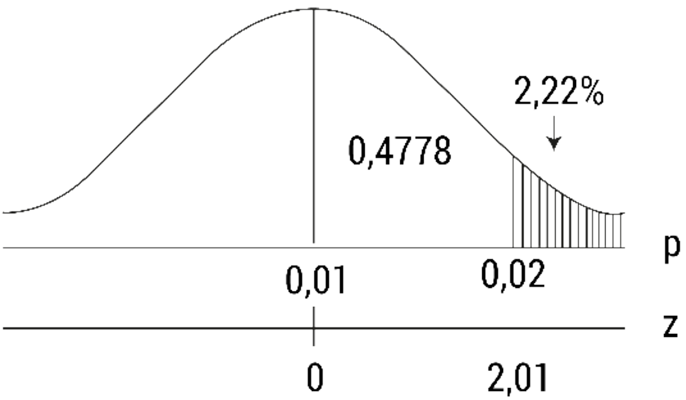
$$P = 0,5000 - 0,4778 = 0,0222$$

$$P_{(p>0,02)} = 0,0222x100 = 2,22$$

3. Finalmente, la región crítica se ilustra en la siguiente figura.

Figura 7

Representación gráfica de la distribución muestral de una proporción



Nota. Adaptado de *Estadística y muestreo (14a. ed.)* (p- 410)[Ilustración], por Martínez, C., 2019, Ecoe Ediciones, CC BY 4.0.

Conclusión:

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

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

Solución Problema Modelo 8: Distribución muestral de una proporción

A continuación, se explica el proceso para determinar la probabilidad de que 42 o más víctimas queden con alguna incapacidad empleando distribución muestral de una proporción.

1. Anotamos los datos e incógnitas del problema:

$$P = 0,25$$

$$n = 150$$

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

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

2. 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 figura 9. Áreas de una distribución normal ordinaria.

$$Z = 0.85 \rightarrow A(0.3023)$$

Figura 8

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	0.1063	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	0.2513	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	0.2935	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	0.3264	0.3315	0.3340	0.3365	0.3389

Nota. Adaptado de *Estadística y muestreo: 13 ed. (p. 411)* [Ilustración], por Martínez, C., 2019, Ecoe Ediciones, CC BY 4.0.

Determinamos la probabilidad requerida:

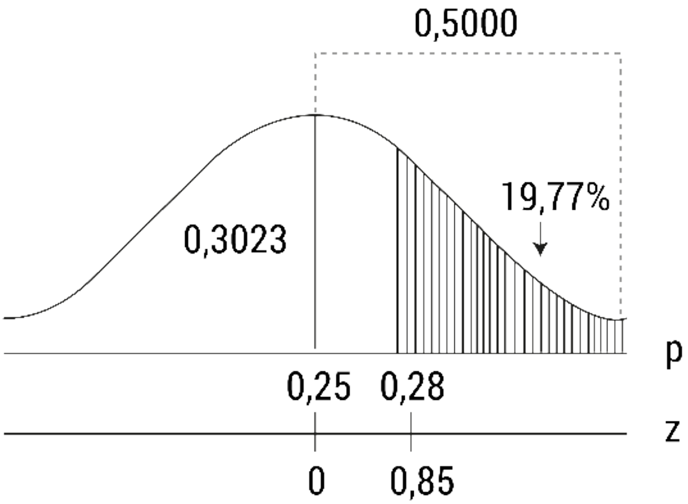
$$P = 0,5000 - 0,3023 = 0,1977$$

$$P(p > 0,28) == 0,1977x100 = 19,77$$

3. Finalmente, la región crítica y la probabilidad requerida se ilustran en la siguiente figura

Figura 9

Representación gráfica de la distribución muestral de una proporción



Nota. Adaptado de *Estadística y muestreo: 13 ed. (p. 410)*[Ilustración], por Martínez, C., 2019, Ecoe Ediciones, CC BY 4.0.

Conclusión:

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

Una vez finalizada la quinta semana de estudio, en la que se analizó la distribución muestral de una proporción, tanto en sus características como en su aplicación, le recuerdo que debe revisar los problemas resueltos de la sección 2.3.3 Distribución de la proporción muestral del texto Estadística inferencial aplicada (2.a ed.) de Díaz (2022).

Actividades de aprendizaje recomendadas

Tras realizar esta revisión y analizado cada uno de los pasos de resolución de los problemas modelos dados, le invito a desarrollar las siguientes actividades:

1. Observar atentamente el video titulado Inferencia Estadística | Distribución muestral de medias y proporciones.
2. Estudiar detenidamente la sección Distribución muestral de proporciones del material educativo Distribuciones muestrales. Enfóquese especialmente en los procesos de resolución de los problemas modelo, ya que esto le permitirá consolidar su comprensión y afrontar con éxito el examen propuesto sobre este tema.

RETROALIMENTACIÓN

En este video se explican las definiciones clave de inferencia estadística, acompañadas de ejercicios que ilustran la aplicación de los conceptos de distribución muestral de proporciones. Además, los materiales sugeridos esta semana le proporcionarán los fundamentos teóricos necesarios para resolver todos los ejercicios propuestos al final de la unidad.

Una vez que haya completado las actividades recomendadas, sus conocimientos se habrán reforzado, preparándolo para continuar avanzando en su proceso de aprendizaje.

Felicitaciones por el arduo trabajo realizado durante esta semana. ¡Ahora, adelante con el siguiente reto!

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

La vida le pondrá obstáculos, pero los límites los pone usted.

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 video [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 μx y μy 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.

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

- Error estándar de las diferencias entre las medias muestrales.

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

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

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

Para ampliar este contenido, se sugiere revisar la sección 1.6 del texto Estadística Inferencial de Llinás (2017) y el capítulo 7, Distribuciones Muestrales, páginas 293-295, del texto Estadística y muestreo de Martínez (2018). En estos materiales educativos se presenta la simbología necesaria para determinar la distribución de diferencias entre dos medias muestrales $(\underline{x} - \underline{y})$. Asimismo, se recomienda revisar la resolución del siguiente problema modelo, donde se expone la aplicación de esta distribución.

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:

Solución Problema Modelo 9: Distribución de diferencias entre dos medias muestrales.

A continuación, se explica el proceso para determinar 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 empleando distribución de diferencias entre dos medias muestrales (Martínez, 2019, elaborado por Granda, 2020).

1. Anotamos los datos e incógnitas del problema:

$$\mu_x = 50 \qquad \mu_y = 60$$

$\sigma_x = 15 \qquad \sigma_y = 18$

$n_1 = 25 \qquad n_2 = 20$

$P_{(x—y<0)} = ?$

2. Calculamos la probabilidad solicitada.

Vamos a determinar la variable estadística Z

$$Z = \frac{(\bar{z}-\bar{y})-(\mu_x-\mu_y)}{\sqrt{\frac{\sigma_z^2}{n_1}+\frac{\sigma_{\bar{y}}^2}{n_2}}}$$

Medias de muestras

$(\bar{x} - \bar{y}) = 0 \longrightarrow$ ya que no nos dan un valor específico, 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 = 1.99 buscamos el área bajo la curva por medio de la figura 3 que presenta las áreas de una distribución normal ordinaria.

$z = 1.99 \rightarrow A(0.4767)$

Determinamos la probabilidad requerida:

$$P = 0,5000—0,4767 = 0,0233$$

$$P(\underline{x} - \underline{y} < 0) = 0.0233 \times 100\% = 2.33\%$$

Conclusión:

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%.**

Para profundizar en el estudio de la distribución muestral, es importante enfocarse en la distribución de las diferencias entre dos proporciones muestrales $P_1 - P_2$. Se recomienda revisar con atención la página 299 del libro Estadística y Muestreo de Martínez (2018), disponible en la biblioteca virtual de la universidad, donde se explican los conceptos clave y su aplicación en el análisis estadístico. A partir de este material, observar el video Distribución de la diferencia de proporciones muestrales, donde se explica en qué ocasiones las distribuciones muestrales se pueden aproximar por normales y cómo, 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 las páginas 299 y 300 del texto Estadística y Muestreo de Martínez (2018), disponible en la biblioteca virtual de la universidad, donde se explica cómo determinar el error estándar de las diferencias entre las dos medias proporcionales. Después de esta revisión, es recomendable resolver un problema modelo que expone la aplicación de esta 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 un 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:

Solución Problema Modelo 10: Distribución de diferencias entre dos proporciones muestrales

A continuación, 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 (Martínez, 2019, elaborado por Granda, 2020).

1. Anotamos los datos e incógnitas del problema:

$$P_1 = 25\% \qquad P_2 = 33\%$$

$$Q_1 = 75\% \qquad Q_2 = 67\%$$

$$n_1 = 150 \qquad n_2 = 100$$

$$\text{hombres} \qquad \text{mujeres}$$

$$P_{(P_1 - P_2 > 0)} = ?$$

2. Calculamos la probabilidad solicitada.

Vamos a determinar la variable estadística Z

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

Proporciones muestrales

$(p_1 — p_2) = 0 \longrightarrow$ 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,

$$Z = \frac{(0) - (-0.08)}{\left(\frac{(0.25)(0.75)}{150} + \frac{(0.33)(0.67)}{100}\right)} = 1.36$$

Con el valor de $Z = 1.36$ buscamos el área bajo la curva por medio de la figura 3 Áreas de una distribución normal ordinaria.

$Z = 1.36 \longrightarrow A(0.4131)$

Determinamos la probabilidad requerida:

$$P = 0.5000 - 0.4131 = 0.0869$$

$$P_{(P_1 - P_2)} > 0 = 0.0869 \times 100\% = 8.69\%$$

Conclusión:

La proporción de mujeres que escuchan el programa es menor o igual a la proporción de hombres de **8.69%.**

Al finalizar la semana de estudio número seis, dedicada al análisis de la distribución de diferencias entre dos medias y dos proporciones muestrales en su aplicación práctica, se recomienda revisar la bibliografía sugerida. Allí encontrará explicaciones detalladas del tema, ejemplos de problemas resueltos y propuestos, así como ejercicios misceláneos acompañados de una síntesis correspondiente al capítulo.

Actividades de aprendizaje recomendadas

Tras revisar el fundamento teórico de esta semana y analizar detalladamente los problemas modelo presentados, es hora de proceder con el desarrollo de las siguientes actividades para afianzar sus conocimientos:

1. Consulte el texto Estadística inferencial aplicada (2.a ed.) de Díaz (2022), específicamente las secciones 2.3.2 y 2.3.4, donde se explican en detalle la distribución de diferencias entre las medias muestrales y las diferencias entre dos proporciones muestrales.
2. Observar atentamente el video Distribución muestral de diferencias de medias.
3. Observar atentamente el video Distribución de diferencia entre dos proporciones.
4. Revise la herramienta GeoGebra: Distribuciones Muestrales, donde se aborda de manera detallada la teoría de las distribuciones muestrales, complementada con ejercicios y problemas resueltos que facilitan su comprensión y aplicación.

RETROALIMENTACIÓN

Los videos de esta semana presentan las definiciones clave de inferencia estadística, acompañadas de ejercicios que ilustran los conceptos de distribución de diferencias entre las medias muestrales y entre dos proporciones muestrales. Además, la bibliografía sugerida ofrece los fundamentos teóricos necesarios para aplicar los conceptos de esta unidad.

Tras completar las actividades propuestas, estará listo para avanzar en su proceso de aprendizaje. Antes de continuar, es fundamental realizar la autoevaluación correspondiente, recordando que esta unidad se centra en las distribuciones muestrales.

¡Enhorabuena! Todo su trabajo responsable le ha permitido culminar la sexta semana. ¡Continúe con el mismo ahínco!

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

5. Estimado estudiante, al finalizar el estudio de la segunda unidad, Distribuciones Muestrales, le recomendamos realizar una autoevaluación para identificar posibles inquietudes y fortalecer su comprensión. No olvide consultar la bibliografía sugerida en las secciones correspondientes para ampliar y consolidar su conocimiento.

Autoevaluación 2

Instrucción: Marque la alternativa correcta.

1. Relacionar las definiciones dadas

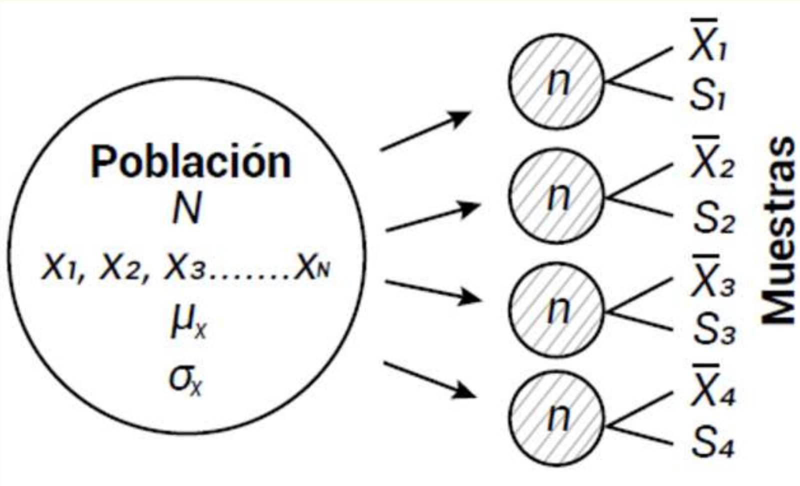
I. Población o Universo II. Unidad III. Muestra	a. Es un subconjunto representativo de unidades de análisis de una población dada b. Conjunto de unidades y elementos que presentan unas características comunes. Estas unidades se denominan variables continuas y discretas. c. Son cada uno de los valores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

2. Lea, analice y responda. Identifique los factores determinantes para las dificultades de aprendizaje y el desarrollo.

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 aleatorio estratificado
c. Muestreo aleatorio estratificado

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:



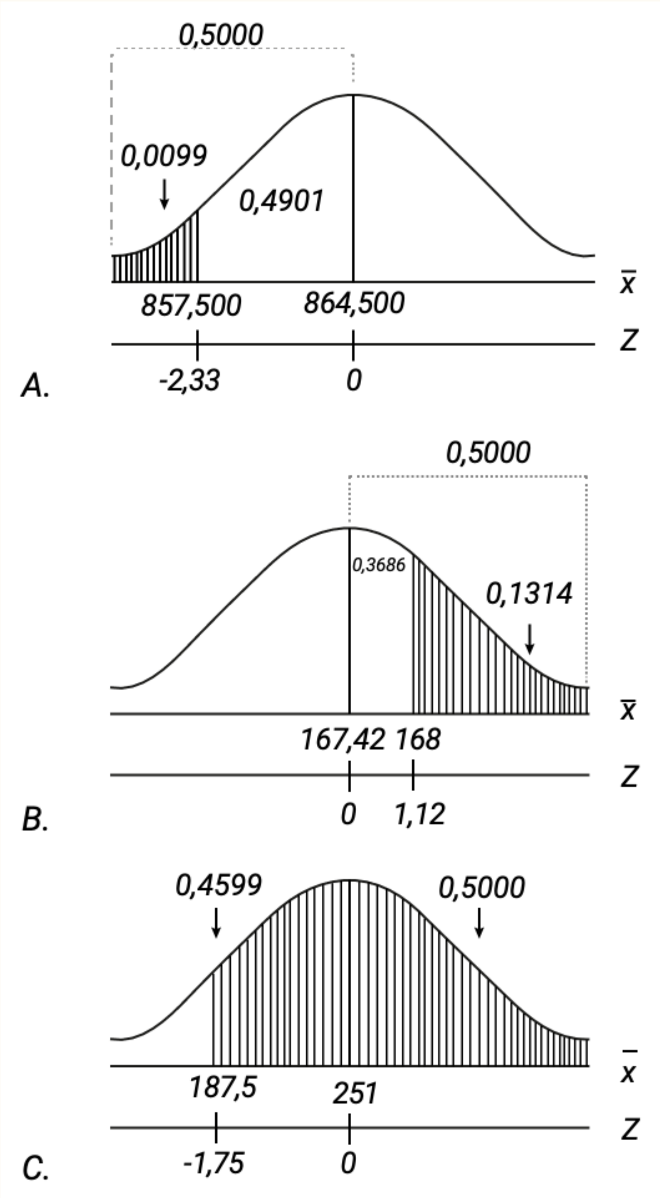
A. Media poblacional.
B. Media muestral.
C. Media proporcional.

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.

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.



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

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

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

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.
$$n = (\frac{Z\sigma}{E})^2$$

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

C.
$$n = (\frac{Z\sigma}{E})$$

Ir al solucionario

Estimado estudiante, le recuerdo que cuenta con el horario de tutoría, durante el cual su profesor tutor está disponible para atender sus dudas e inquietudes. Este espacio le permitirá resolver cualquier dificultad que surja durante la semana de trabajo y validar su desempeño. Asimismo, le animo a participar activamente en las actividades síncronas planificadas a lo largo de la asignatura, ya que estas representan una excelente oportunidad para superar limitaciones de tiempo y espacio, optimizando así sus resultados de aprendizaje

Resultado de aprendizaje 2:

Valida las hipótesis que se generan en cada uno de los casos analizados.

Para alcanzar el resultado planteado, usted desarrollará la capacidad de validar hipótesis en distintos casos analizados, aplicando métodos estadísticos y criterios de rigor científico. Esto le permitirá evaluar la veracidad de los planteamientos iniciales, fundamentar sus conclusiones con base en evidencias y fortalecer la toma de decisiones en diversos contextos. A través de este proceso, adquirirá habilidades analíticas esenciales para la interpretación y el uso adecuado de la información en la resolución de problemas.

Contenidos, recursos y actividades de aprendizaje recomendadas

Recuerde revisar de manera paralela los contenidos con las actividades de aprendizaje recomendadas y actividades de aprendizaje evaluadas.

Semana 7

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 video Conceptos, que le permitirá comprender argumentos estadísticos de moderada complejidad en resultados de diversos tipos de estudios. Así mismo, le presento una explicación sobre un Test de hipótesis, que le sugiero escuchar atentamente antes de seguir estudiando con ahínco las secciones siguientes. *¡Bienvenido!*

Como bien lo explican Inzunza y Vidal (2013), las **pruebas de hipótesis** son uno de los métodos fundamentales de la inferencia estadística. Este procedimiento permite verificar una afirmación sobre el valor de un parámetro poblacional. Dado que los datos provienen de una muestra, los resultados pueden estar sujetos a variaciones aleatorias.

Por ello, una prueba de hipótesis evalúa si las desviaciones observadas con respecto al resultado que idealmente debería haber ocurrido según la hipótesis planteada, son atribuibles al azar o, por el contrario, reflejan que los resultados no coinciden con la hipótesis formulada sobre el valor del parámetro (p. 181).

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.

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).

Le recomiendo consultar las páginas 324 a 326 del libro Estadística y muestreo de Martínez (2018), disponible en la biblioteca virtual de la universidad. En este libro encontrará una explicación clara y detallada, acompañada de ejemplos, sobre la definición de las pruebas de hipótesis y su clasificación, las cuales se explican a continuación en base a lo expuesto por Triola (2013).

- La **hipótesis nula** (denotada con H0) 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 Hα) 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 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 cuando se debió rechazar.
- Si rechazamos una hipótesis falsa, la decisión es correcta.

3.2. Prueba unilateral y bilateral

En el texto Estadística y muestreo Estadística y muestreo de Martínez (2018), usted encontrará la explicación sobre hipótesis unilateral, bilateral, nivel de significación y puntos críticos. Al respecto, Martínez (2018) 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 un área bajo la curva de probabilidad o normal, denominada región crítica o zona de rechazo.

Finalmente, revise la página 329 del texto Estadística y muestreo, donde se detalla el procedimiento para realizar las pruebas de hipótesis. En esta sección, se incluye una síntesis de los pasos a seguir, acompañada de explicaciones ampliadas y ejemplos ilustrativos para facilitar la comprensión de dicho proceso.

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

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:

Solución Problema Modelo 11: Prueba de hipótesis.

A continuación, se explica el proceso para probar que los estudiantes de la ciudad son superiores en inglés.

1. Anotamos los datos e incógnitas del problema:

$\mu = 64$

$\sigma = 8$

$n = 64$

$\bar{x} = 68$

2. Calculamos la probabilidad solicitada.

Hipótesis nula $H_0 : \mu = 64$

Hipótesis alternativa a $H_0 : \mu > 64$

Dócima unilateral a la derecha

Nivel de significación $\alpha = 0.05$

Variante estadística

Distribución de medias muestrales

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

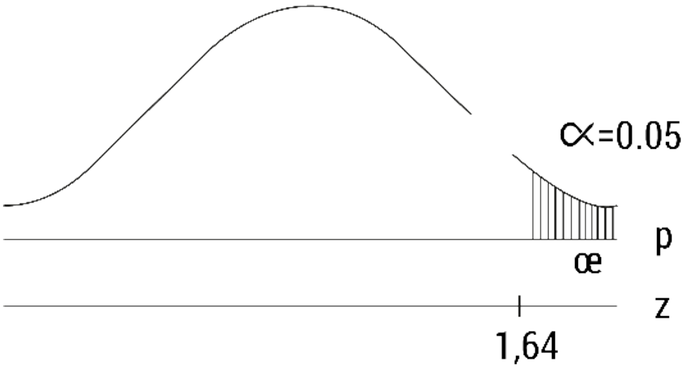
$$z = \frac{68-64}{\frac{8}{\sqrt{64}}}$$

La región crítica de la dócima unilateral a la derecha $Z_s < 1,64$

3. La descripción de la región crítica se ilustra en la siguiente figura.

Figura 10

Representación típica de una prueba de hipótesis unilateral derecha en estadística



Nota. Tomado de *Estadística y muestreo: 13 ed. (p. 531)* [Ilustración], por Martínez, C., 2019, Ecoe Ediciones, CC BY 4.0.

Conclusión:

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.

Una vez terminada la semana de estudio número siete, en la que se abordó el análisis de las pruebas de hipótesis, se recomienda estudiar la bibliografía sugerida, esta guía y los demás materiales propuestos. En vista de la necesidad para que usted, como estudiante y futuro profesional, comprenda la integración y relación que guardan entre sí los conceptos estadísticos en el proceso de prueba.

Actividades de aprendizaje recomendadas

Para reforzar este aprendizaje, se sugiere desarrollar las siguientes actividades:

1. Observar cuidadosamente el video Prueba de Hipótesis para la media.
2. Interactuar con la aplicación GeoGebra Contraste de Hipótesis para la media.

3. Revisar el capítulo 4 (páginas 178-187) del libro Estadística inferencial aplicada (2.a ed.) de Díaz (2022), disponible en la biblioteca virtual de la universidad. En este capítulo encontrará el fundamento teórico y el desarrollo de problemas resueltos que le serán de utilidad para su aprendizaje y práctica en esta unidad.
4. Acceder al sitio web OpenStax y diríjase a la sección 9, Pruebas de hipótesis con una muestra. Allí podrá analizar el contenido correspondiente y profundizar en el material de estudio asignado para esta semana.

RETROALIMENTACIÓN

El video recomendado explica la prueba de hipótesis para la media, brindando una síntesis clara de los parámetros clave para aplicar los contenidos de esta unidad. Asimismo, mediante la aplicación GeoGebra, podrá interactuar ajustando los parámetros de un contraste de hipótesis, lo que enriquecerá su comprensión de manera práctica. En la bibliografía sugerida y los recursos en línea, encontrará los fundamentos teóricos necesarios para resolver los problemas planteados al final de la unidad.

Al finalizar las actividades recomendadas, estará preparado para cerrar con éxito las tareas correspondientes al bimestre.

¡Enhorabuena por su esfuerzo y dedicación en cada actividad! Este logro refleja un paso significativo en su trayectoria de aprendizaje.

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

Resultado de aprendizaje 1 y 2:

- Aplica las técnicas de muestreo en base al caso de estudio. Identifica métodos de estimación de parámetros y su interpretación en casos de estudio.
- Valida las hipótesis que se generan en cada uno de los casos analizados.

Contenidos, recursos y actividades de aprendizaje recomendadas

Semana 8

Actividades finales del bimestre

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.
- 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 las siguientes actividades:

1. Revise cada uno de los conceptos estudiados en las dos unidades planificadas y desarrolladas en este primer bimestre.
2. Realice diversos ejercicios y problemas aplicados a los conceptos, propiedades y leyes abordados en cada una de las unidades estudiadas. Utilice para ello las actividades disponibles en el sitio web OpenStax, donde podrá reforzar su comprensión y dominio de los temas.
3. Para cada unidad, es recomendable sistematizar el conocimiento adquirido mediante la elaboración de un mapa conceptual u otro organizador gráfico que considere adecuado. Esta práctica facilitará la comprensión y estructuración de los contenidos aprendidos.

Nota: por favor complete las actividades en un cuaderno o documento Word.

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:

Valida las hipótesis que se generan en cada uno de los casos analizados.

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 recomendadas

Recuerde revisar de manera paralela los contenidos con las actividades de aprendizaje recomendadas y actividades de aprendizaje evaluadas.

Semana 9

Haciendo un paréntesis en el estudio de la prueba de hipótesis, en la presente 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.

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 PowerPoint. Finalmente, SPSS permite manejar bancos de datos de gran magnitud y también efectuar análisis estadísticos muy complejos.

- 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

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. Comparamos 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?
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.

Tabla 3
Comparación de niveles de autoestima entre hombres y mujeres

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

Nota. Adaptado de *Estadística y muestreo* (p. 345), por Martínez, C. (2019), Ecoe Ediciones.

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.

RESULTADOS SPSS

Tabla 4. *Estadísticos de grupo*

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

Nota. Granda, S., 2025.

Tabla 5
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			
Se han asumido varianzas iguales.			2,534	18	0,021
No se han asumido varianzas iguales			2,534	17,098	0,021

Nota. Granda, S., 2025.

Actividades de aprendizaje recomendadas

Es hora de reforzar los conocimientos adquiridos resolviendo las siguientes actividades:

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

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

Tabla 6

Eficacia de terapia en capacidad lectoescritora de niños disléxicos

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

Nota. Adaptado de *Manual básico SPSS* (p.17), por J.A. González G.(2009), Universidad de Talca.

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.

Nota: por favor complete la actividad en un cuaderno o documento Word.

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!

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 ha 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 artículo de aprendizaje Distribución muestral de la proporción que lo preparará para iniciar el estudio de esta sección ¡Bienvenido!

Unidad 3. Pruebas de hipótesis

3.5. Distribución de proporciones muestrales ρ

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 determina 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 fracaso

n = Tamaño de la muestra

A continuación, se presentará este contenido mediante un problema modelo. Le invito a seguir y reflexionar sobre el proceso de resolución, así como a revisar la sección 4.3.5 del libro Estadística inferencial aplicada (2.a ed.) de Díaz (2022), disponible en la biblioteca virtual de la universidad.

PROBLEMA MODELO #12

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.

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

Solución:

Solución Problema Modelo 12: Prueba de hipótesis

A continuación, se explica el proceso para aceptar o rechazar la hipótesis aplicando la distribución de proporciones muestrales.

1. Anotamos los datos del problema:

   $n = 360$

   $\alpha = 0.05$

   $P = 0.14$

   Cálculo de la probabilidad o proporción:

   $p = \frac{48}{360} = 0.13$

   $q = 1 - 0.13 = 0.87$

2. Calculamos la probabilidad solicitada.

   Hipótesis nula $H_0 : P = 0.14$

   Hipótesis alternativa $H_a : P < 0.14$

   Dócima unilateral a la izquierda.

   Nivel de significación $\alpha = 0.05$

   Variante estadística

   Distribución de medias muestrales

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

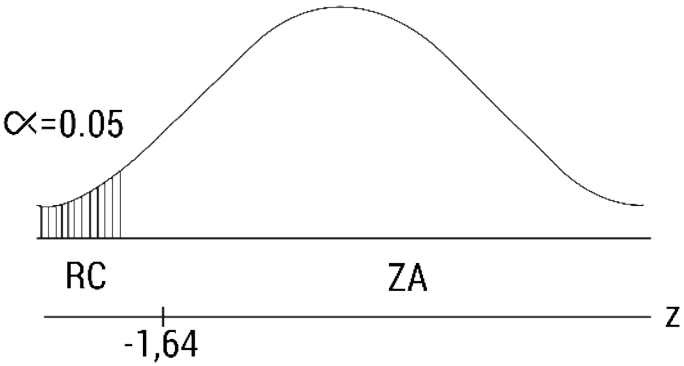
$$z = \frac{0.13-0.14}{\sqrt{\frac{(0.13)(0.87)}{360}}} = -0.5642$$

La región crítica de la dócima unilateral a la izquierda $Z_s < 1,64$

3. La descripción de la región crítica se ilustra en la siguiente figura.

Figura 11

Representación típica de una prueba de hipótesis unilateral derecha en estadística



Nota. Tomado de *Estadística y muestreo: 13 ed. (p. 531)* [Ilustración], por Martínez, C., 2019, Ecoe Ediciones, CC BY 4.0.

El valor de $z =— 0.5642$ caen en zona de aceptación a la izquierda por lo tanto se acepta H_0.

Conclusión:

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

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 que 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 video de estudio PRUEBA DE HIPÓTESIS PARA DIFERENCIAS DE MEDIAS. MUESTRAS GRANDES Y VARIANZAS 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 $\bar{x} - \bar{y}$

Ahora analizaremos el caso de dos poblaciones independientes, en el que se busca determinar si la diferencia entre sus medias muestrales es significativa o si una media es mayor o menor que la otra. Para ello, le recomiendo, estimado estudiante, revisar y analizar detenidamente la sección 4.3 del libro Estadística inferencial 1 para ingeniería y ciencias 1 de Vladimirovna & Gutiérrez (2016), disponible en la biblioteca virtual de la universidad.

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:

Solución Problema Modelo 13: Prueba de hipótesis

A continuación, se explica el proceso para aceptar o rechazar la hipótesis nula del problema, aplicando la distribución de diferencias entre dos modelos.

1. Anotamos los datos del problema

$$n_1 = 100$$

$$n_2 = 90$$

$$a = 0.01$$

$$\bar{x} = 107$$

$$\bar{y} = 103$$

$$S_x = 17$$

$$S_y = 16$$

2. Calculamos la probabilidad solicitada.

Hipótesis nula $H_0 : \mu_x = \mu_y$

Hipótesis alternativa $H_a : \mu_x \neq \mu_y$
Dócima bilateral

Nivel de significación $\alpha = 0.05$

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

$$Z = \frac{\bar{x}-\bar{y}}{\sqrt{\frac{s_x^2}{n_1} + \frac{s_y^2}{n_2}}}$$

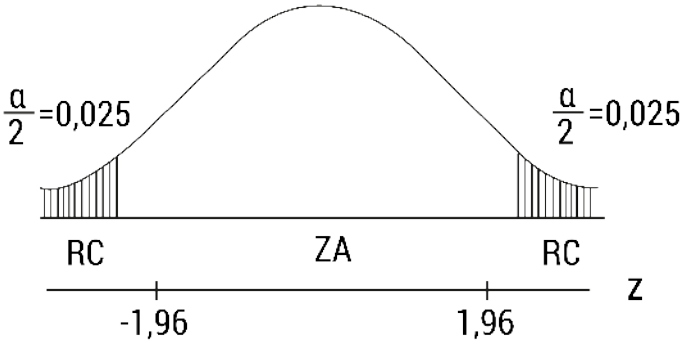
$$Z = \frac{107-103}{2.3947} = 1.67$$

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

La descripción de la región crítica se ilustra en la siguiente figura.

Figura 12

Representación típica de una prueba de hipótesis bilateral en estadística.



Nota. Tomado de *Estadística y muestreo: 13 ed. (p. 531)* [Ilustración], por Martínez, C., 2019, Ecoe Ediciones, CC BY 4.0.

Conclusión:

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

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:

1. Observar cuidadosamente el video Prueba de hipótesis.
2. Revisar minuciosamente el video de estudio PRUEBA DE HIPÓTESIS PARA DIFERENCIAS DE MEDIAS CON MUESTRAS GRANDES.
3. Leer el capítulo 8, *Pruebas de hipótesis*, del texto Estadística y Muestreo de Martínez (2018), con énfasis en los temas sobre la distribución de proporciones muestrales y las diferencias entre medias muestrales.
4. Revisar el sitio web OpenStax, enfocándose en las secciones 10.3 y 10.4 del tema Prueba de hipótesis con dos muestras. Preste especial atención a los procesos de solución de problemas de aplicación, ya que estos serán clave para su comprensión y práctica del contenido.

RETROALIMENTACIÓN

Los videos proporcionados explican la teoría básica de esta semana y presentan un ejemplo resuelto paso a paso sobre la validación de una hipótesis de medias. Además, en la bibliografía sugerida encontrará los fundamentos teóricos necesarios para resolver 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!

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

¡No se rinda! Cada logro por más pequeño que sea le aproxima cada día a su objetivo.

Contenidos, recursos y actividades de aprendizaje recomendadas

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 estas 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 video 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!

Unidad 3. Pruebas de hipótesis

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 H0 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

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

$$S_{p_2} = \sqrt{p_2 q_2}$$

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

PROBLEMA MODELO #14

Solución de un problema de prueba de hipótesis.

Un estudio del consumo de café en el trabajo, por sex o, 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:

Solución Problema Modelo 14: Prueba de hipótesis

A continuación, se explica el proceso para aceptar o rechazar la hipótesis nula del problema, aplicando la distribución de diferencias entre dos proporciones.

1. Anotamos los datos del problema.

   $$n_1 = 200$$

   $$n_2 = 150$$

   $$\alpha = 0.05$$

   Determinamos:

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

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

2. Calculamos la probabilidad solicitada.

   Hipótesis nula $H_0 : P_1 = P_2$

   Hipótesis alternativa $H_\alpha : P_1 \neq P_2$

   Dócima bilateral

   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

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

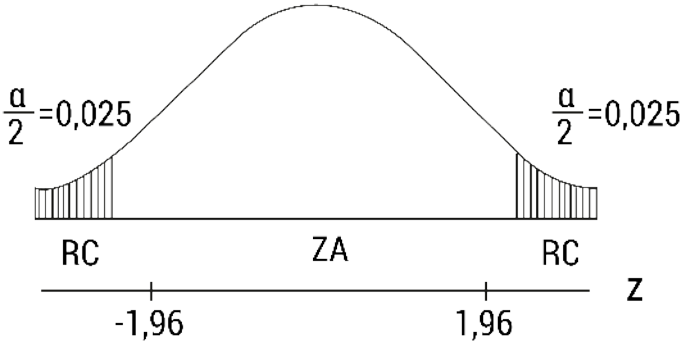
$$Z = \frac{0.64 - 0.71}{\sqrt{\frac{0.64(0.36)}{200} + \frac{0.71(0.29)}{150}}} = -1.39$$

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

La descripción de la región crítica se ilustra en la siguiente figura.

Figura 13

Representación típica de una prueba de hipótesis bilateral en estadística



Nota. Tomado de Estadística y muestreo: 13 ed. (p. 531) [Ilustración], por Martínez, C., 2019, Ecoe Ediciones, CC BY 4.0.

Conclusión:

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

3.8. Distribución t Student

La **distribución t de 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 de Student, es necesario que revise detalladamente las páginas 351-353 del texto Estadística y muestreo de Martínez (2018), disponible en la biblioteca virtual de la universidad.

Recuerde la función de la distribución “t” es:

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

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, estimado 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 denominada Tabla de distribución “t” de Student de la página 354 del texto Estadística y muestreo de Martínez (2018), 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. Revisar sobre la Comparación de la distribución T con la normal estándar, donde podrá observar cómo cambia la distribución t al modificar los grados de libertad.

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 considerará que la desviación está sin corregir y teniendo que corregirla para la aplicación en la variante estadística “t”.

Se utilizará los siguientes símbolos:

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

$$S = \text{desviación típica corregida}$$

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

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

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

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:

Solución Problema Modelo 15: Prueba de hipótesis.

A continuación, se explica el proceso para aceptar o rechazar la hipótesis nula del problema, aplicando la distribución de medias muestrales.

1. 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.

$$v = 16 — 1 = — 15$$

$$v = 15$$

2. Calculamos la probabilidad solicitada.

Hipótesis nula $H_0 : \mu = 10$

Hipótesis alternativa $H_\alpha : \mu > 10$

Dócima unilateral

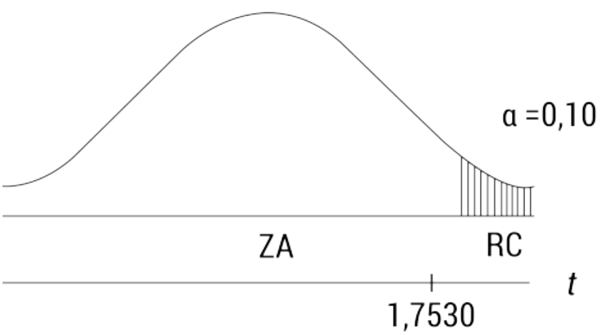
Por ser unilateral se toma el doble del nivel de significación en la RC (zona de rechazo)

Nivel de significación $\alpha = 0.10$

El valor del punto crítico obtenido de la tabla 8.3 es Z = 1.7530 como se indica en la figura

Figura 14

Representación típica de una prueba de hipótesis unilateral derecha en estadística



Nota. Tomado de *Estadística y muestreo: 13 ed. (p. 574)* [Ilustración], por Martínez, C., 2019, Ecoe Ediciones, CC BY 4.0.

Cálculo de la variante estadística "t".

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

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

Conclusión:

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

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 de Student y distribución de medias muestrales. Le recordamos revisar detenidamente la bibliografía, donde encontrará explicaciones detalladas sobre el tema de esta semana, así como ejemplos de problemas resueltos y propuestos.

Actividades de aprendizaje recomendadas

Luego de profundizar en estos temas, le recomiendo desarrollar las siguientes actividades:

1. Observar cuidadosamente el video: Uso de la Tabla T-Student.
2. Interactuar con la aplicación de GeoGebra Student t Distribution cambiando los grados de libertad para que compare el área bajo la curva de una distribución normal con una distribución t.
3. Leer el texto Estadística y muestreo de Martínez (2018), específicamente los temas sobre la distribución t de Student y la distribución de medias muestrales para muestras pequeñas. En este material encontrará, además del fundamento teórico, el desarrollo de problemas resueltos, un resumen sobre el manejo de la varianza para muestras grandes y pequeñas, así como la tabla de distribución t de

Student. Todo ello, será de gran utilidad para su aprendizaje y práctica en esta unidad.

4. Revisar la sección 4 del texto Estadística descriptiva e inferencial de Proaño (2020).

RETROALIMENTACIÓN

El video presenta el uso de la tabla t de Student, incluyendo su origen, definición y aplicación, así como la comparación de áreas entre distintas distribuciones. Además, en el libro Estadística y muestreo de Martínez (2018), se encuentran los fundamentos teóricos necesarios para resolver los problemas propuestos al final de la unidad.

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!

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

Contenidos, recursos y actividades de aprendizaje recomendadas

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 video 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!

Unidad 3. Pruebas de hipótesis

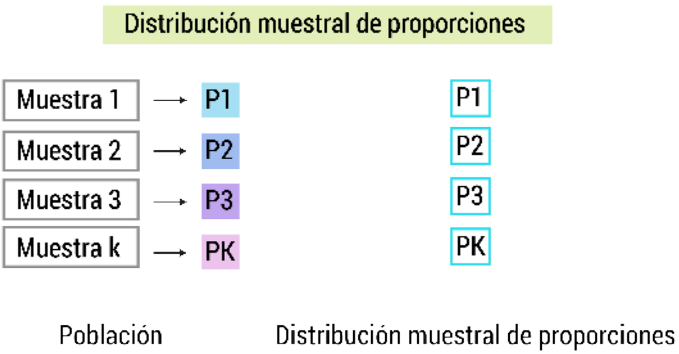
3.10. Distribución de una proporción muestral

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.

La figura siguiente 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.

Figura 15

Distribución muestral de proporciones



Nota. Adaptado de *Distribuciones muestrales y estimación por intervalos* [Ilustración], por Facultad de Estudios a Distancia, 2018, UniversidadMilitarNuevaGranada, CC BY 4.0.

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.

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:

Tabla 7
Formulación de hipótesis estadísticas para pruebas de proporciones

Hipótesis nula	Hipótesis alternativa
$H_0 : P = p_o$ Donde p_o es la constante determinada.	Se tienen las siguientes tres posibilidades: 1. $H_1 : P > p_o$ 2. $H_1 : P < p_o$ 3. $H_1 : P = p_o$

Nota. Granda, S., 2025.

Así mismo, para determinar el estadístico de prueba, emplearemos $t = \frac{p-p}{\sqrt{\frac{pq}{n-1}}}$ expresión, la que aplicaremos en la resolución del siguiente problema modelo. Le invito a revisar detenidamente el proceso desarrollado y a consultar el libro Estadística y muestreo de Martínez (2018), donde encontrará el análisis de los problemas resueltos relacionados.

PROBLEMA MODELO #16

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?

Solución:

Solución Problema Modelo 16: Prueba de hipótesis.

A continuación, se explica el proceso para aceptar o rechazar la hipótesis nula del problema, aplicando la distribución de medias muestrales.

1. Anotamos los datos del problema.

$P = 0.86$

$n = 18$

$p = \frac{16}{18} = 0.89$

$a = 0.01$

Determinamos los grados de libertad.

$$v = 18 — 1 = 17$$

$$v = 17$$

Cálculo de la desviación típica.

$$S_p = \sqrt{pq}$$

$$S_p = 0.89(0.11)$$

2. Calculamos la probabilidad solicitada.

Hipótesis nula $H_0 : P = 0.86$

Hipótesis alternativa $H_\alpha : P > 0.86$

Dócima unilateral derecha

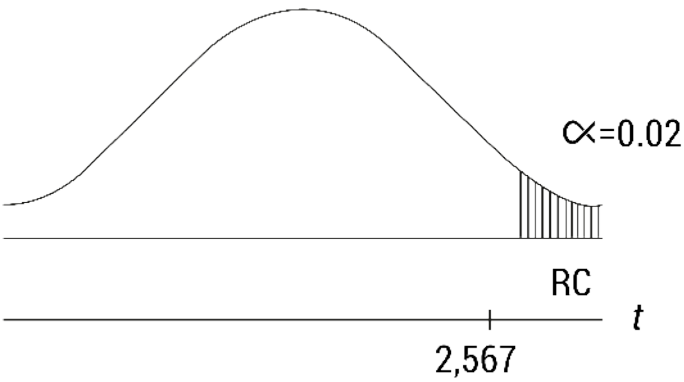
Por ser unilateral se toma el doble del nivel de significación en la RC (zona de rechazo)

Nivel de significación $\alpha = 0.02$

El valor del punto crítico obtenido de la tabla 8.3 es Z= 2.567 como se indica en la figura dada a continuación.

Figura 16

Representación típica de una prueba de hipótesis unilateral derecha en estadística



Nota. Adaptado de *Estadística y muestreo* (p. 352), por Martínez, C. (2019), Ecoe Ediciones.

Cálculo de la variante estadística "t".

$$t = \frac{p-p}{\sqrt{\frac{pq}{n-1}}} = \frac{0.89-0.86}{\sqrt{\frac{0.89(0.11)}{18-1}}} = 0.40$$

Conclusión:

El equipo no es mucho más efectivo que el señalado por el fabricante, al nivel del 1%.

Ahora continuamos con la segunda parte de esta semana, la distribución de diferencias entre dos medias muestrales, recuerde que cuando n1 o n2 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, aplicando diferentes fórmulas según el tipo de información proporcionada. Le recomiendo revisar el texto Estadística y muestreo de

Martínez (2018), específicamente en las páginas 360-361, donde se explican los casos en los que se consideran muestras de dos poblaciones con varianzas poblacionales iguales. Esta condición, que se cumple en la mayoría de los casos, 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.

$$S^2=\frac{\sum(x_i-\bar{x})^2+\sum(y_i-\bar{y})^2}{n_1+n_2-2} \text{ o } S^2=\frac{|\sum x_i^2-n_1\bar{x}^2|+|\sum y_i^2-n_2\bar{y}^2|}{n_1+n_2-2}$$

   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 y a calculadas (supuestamente corregidas) se aplica las siguientes fórmulas:

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

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

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

   Para finalmente, obtener la variante estadística

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

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.

$$\nu = \frac{\left(\frac{S_X^2}{n_1}+\frac{S_Y^2}{n_2}\right)^2}{\frac{\left(\frac{S_X^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{S_Y^2}{n_2}\right)^2}{n_2-1}}$$

A continuación, se desarrollará el contenido previamente explicado mediante la solución del problema modelo. Le invito a revisar el siguiente módulo didáctico para seguir y analizar este proceso, así como a revisar los problemas resueltos en la bibliografía y los recursos abiertos sugeridos.

Problema Modelo #17: Prueba de hipótesis

Al finalizar la semana doce, en la que se analizó la distribución de una proporción muestral y las diferencias entre dos medias muestrales para muestras pequeñas, le recomiendo revisar la bibliografía sugerida, así como los conceptos y aplicaciones teórico-prácticas de estas distribuciones.

Actividades de aprendizaje recomendadas

Para reforzar sus conocimientos en este tema, le invito a realizar las siguientes actividades:

1. Indagar en el sitio web LibreTexts en la sección donde se aborda la comparación de dos medias poblacionales para muestras pequeñas independientes. Este sitio web proporciona una explicación detallada

sobre los fundamentos teóricos, las condiciones necesarias para realizar esta comparación, y los procedimientos estadísticos aplicables, incluyendo el uso de la distribución t de Student.

2. Analizar el video de estudio Prueba de hipótesis para diferencia de medias con muestras pequeñas, aquí se desarrolla paso a paso un ejemplo.
3. Observar el video de estudio Tutorial pasos contraste de hipótesis.

RETROALIMENTACIÓN

La bibliografía sugerida explica cómo resolver problemas de distribución de una proporción muestral y diferencias entre dos medias muestrales mediante ejemplos paso a paso. Esto le ayudará a comprender los temas y simplificar los procesos usando tecnología. Además, los videos de estudio ofrecen una síntesis clara de los tipos de pruebas de hipótesis estudiadas y su aplicación práctica.

¡Felicidades por su esfuerzo hasta aquí!

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

4. Estimado estudiante, al concluir el estudio de la tercera unidad, Pruebas de Hipótesis, le recomendamos realizar una autoevaluación. Esto le permitirá identificar y abordar las inquietudes más relevantes para fortalecer su comprensión del tema.

Autoevaluación 3

Seleccione la opción correcta.

1. El término 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 sabemos 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. Rekacionar 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.

II. Es considerada como la hipótesis nula, y a que hace referencia al valor del parámetro que se quiere probar como verdadero.

III. Aceptar la hipótesis nula (H_a) cuando se ha debido rechazar. En el ejercicio que estamos desarrollando, sería: "Aceptar la moneda como correcta, cuando en verdad no lo es".

IV. Rechazar la hipótesis nula (H_a) cuando se ha debido aceptar. "Rechazar la moneda como incorrecta, cuando en verdad está equilibrada".

a. Error tipo II

b. Error tipo I

c. H_o

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

		VERDADERA	FALSA
D	RECHAZAR	Decisión correcta	Error tipo II

A. A-B
B. C-D
C. A-C

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. La hipótesis nula y alternativa serían:

$$H_O : p = 0,5$$

$$H_1 : p > 0,5$$

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 establece la Federal Aviation Administration). A partir de esto identifique la hipótesis nula y la hipótesis alternativa correcta:

A. $H_O \mu \leq 195$ libras y $H_1 \mu > 195$ libras

B. $H_O \mu > 195$ libras y $H_1 \mu > 195$ libras

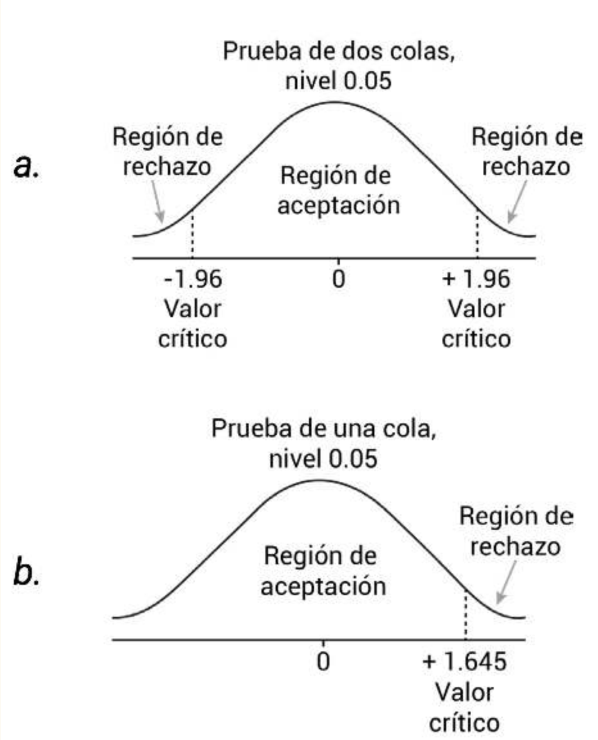
C. $H_O \mu = 195$ libras y $H_1 \mu > 195$ libras

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 aplicará 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, o sea que la hipótesis alternativa es c. diferente.



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: $\mu = 86$ teniendo como datos

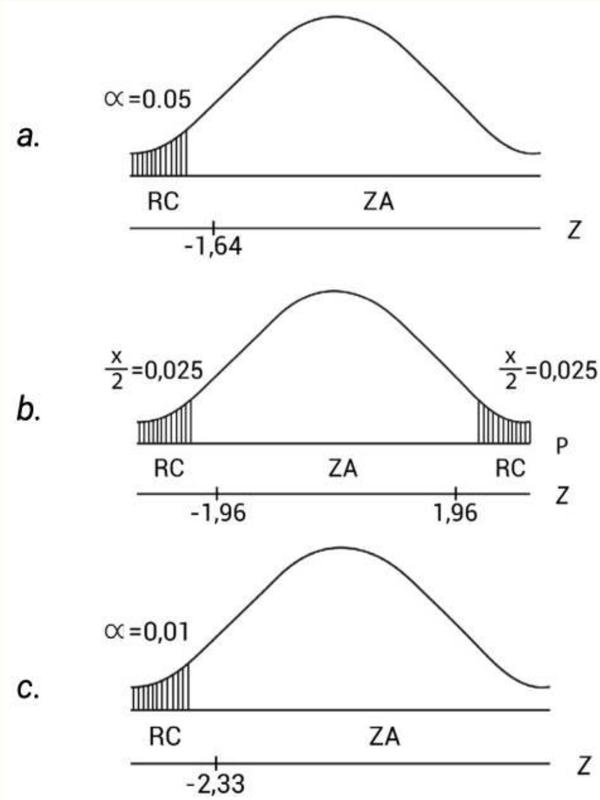
$\bar{x} = 82, \sigma = 15$ y $n = 25$ utilizamos la variante estadística:

A. $Z = \frac{x-\mu}{\sigma}$

B. $Z = \frac{\bar{x}-\mu}{\sigma}$

C. $Z = \frac{\bar{x}-\mu}{\frac{S}{\sqrt{n}}}$

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:



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	n1 = 130	80
B	n2 = 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. $S_{\bar{x}—\bar{y}} = 0.01$

Ir al solucionario

Si surgen dudas en una o más preguntas vuelva a leer el contenido científico para que identifique la validez de su respuesta.

Resultado de aprendizaje 3:

Emplea información no paramétrica para análisis de datos.

Para alcanzar el resultado de aprendizaje planteado, usted aplicará pruebas de hipótesis no paramétricas en el análisis de datos. A través del estudio de la varianza y la comparación entre varianzas de dos poblaciones, comprenderá cómo evaluar la dispersión de los datos y su significancia estadística. Además, mediante la prueba del coeficiente de correlación de Pearson, analizará la relación entre variables. Las actividades de aprendizaje recomendadas permitirán reforzar estos conceptos, promoviendo el desarrollo de habilidades para interpretar y utilizar técnicas estadísticas en diversos contextos.

Contenidos, recursos y actividades de aprendizaje recomendadas

Recuerde revisar de manera paralela los contenidos con las actividades de aprendizaje recomendadas y actividades de aprendizaje evaluadas.

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.

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, y a que estas son la oportunidad de reducir obstáculos como el tiempo y el espacio, y obtener así mejores resultados en su aprendizaje.

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 video 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 presentada en la tabla:

Tabla 8
Simbología en pruebas de hipótesis para varianzas poblacionales

Notación matemática	Descripción
$H_0\colon \sigma^2 = \sigma_0^2 \leftrightarrow H_0\colon \frac{\sigma^2}{\sigma_0^2} = 1$	**HIPÓTESIS NULA**
$H_\alpha : \frac{\sigma^2}{\sigma_0^2} \neq 1$ $H_\alpha : \frac{\sigma^2}{\sigma_0^2} > 1$ $H_\alpha : \frac{\sigma^2}{\sigma_0^2} < 1$	**HIPÓTESIS ALTERNATIVA**
$x^2 = \frac{(n-1)S^2}{\sigma_0^2}$	**VARIANTE ESTADÍSTICA** *Chi-cuadrado* con *n* — 1 grados de libertad

Nota. Adaptado de *Estadística y muestreo* (p. 404), por Martínez, C. (2019), Ecoe Ediciones.

Le recomiendo revisar el texto Estadística y muestreo de Martínez (2018), en las páginas 404-406, donde se detalla el proceso a seguir en este tipo de prueba de hipótesis. Además, encontrará una explicación clara sobre los límites de confianza, acompañada de ejemplos ilustrativos. En esta sección también se incluye la Tabla 9.1, que presenta los percentiles de la distribución χ^2. Es fundamental que usted analice detenidamente esta información para comprender el uso de la tabla en la determinación de los valores críticos.

A continuación, se presenta una explicación detallada del contenido abordado previamente en la solución del problema modelo. Le invito a seguir y analizar este proceso, así como a consultar los problemas resueltos en la bibliografía sugerida.

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 $\sigma = 5$.

DATOS: $n = 15$ y $\hat{S} = 7$, $\sigma = 5$

Solución Problema Modelo 18: Prueba de hipótesis de una varianza

A continuación, se explica el proceso para aceptar o rechazar la hipótesis nula del problema, aplicando la prueba de hipótesis de una varianza.

Vamos a establecer la hipótesis:

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

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

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

$$\frac{x^2}{v} = \frac{\hat{S}^2}{25}$$

Determinamos los valores críticos:

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

$$\left(\frac{x^2}{\overline{\nu}}\right)_{\frac{\alpha}{2};\nu} \longrightarrow \left(\frac{x^2}{\overline{\nu}}\right)_{0.025;14}$$

$$\left(\frac{x^2}{v}\right)_{1-\frac{\alpha}{2},\nu} \longrightarrow \left(\frac{x^2}{v}\right)_{0.975;14}$$

Ahora debemos utilizar la Tabla 9.1 de la página 406, de Percentiles de la distribución $\frac{X^2}{v}$, que se encuentra en el texto Estadística y muestreo de Martínez (2018), para determinar los valores críticos.

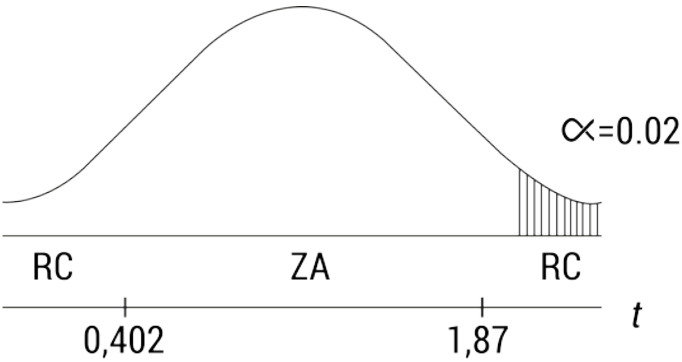
$$\left(\frac{x^2}{v}\right)_{0.025,14} = 0.402$$

$$\left(\frac{x^2}{v}\right)_{0.975,14} = 1.87$$

Los valores críticos obtenidos son 0.402 y 1.87 como se indica en la figura.

Figura 17

Representación típica de una prueba de hipótesis bilateral de una varianza



Nota. Adaptado de Estadística y muestreo: 13 ed. [Ilustración], por Martínez, C., 2019, Ecoe Ediciones, CC BY 4.0.

Finalmente, calculamos el valor estadístico para analizarlo y adoptar una decisión. La variante estadística debe estar entre los valores críticos encontrados.

$$0.402\frac{X^2}{v} > 1.87$$

Entonces,

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

Conclusión:

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

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.

Tabla 9

Simbología de prueba de hipótesis para varianza poblacional

Descripción fórmula	Elementos estadísticos
$H_0 : \sigma^2 = \sigma_0^2$ $H_\alpha : \sigma^2 \neq \sigma_0^2$	**Formulación de hipótesis**
$\alpha = 0.05$	**Grado de significación**
$x^2 = \frac{(n-1)S^2}{\sigma_0^2}$	**Variante estadística**

Nota. Granda, S., 2025.

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

A continuación, se detallará este proceso a través de la solución del problema modelo. Le invito a seguir y analizar este ejemplo, así como los problemas resueltos en el sitio web OpenStax, específicamente en la sección Prueba de una sola varianza.

PROBLEMA MODELO #19

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,$\sigma^2 = 64$, $\alpha = 0.1$

Solución:

Solución Problema Modelo 19: Prueba de hipótesis de una varianza

A continuación, se explica el proceso para aceptar o rechazar la hipótesis nula del problema, aplicando la prueba de hipótesis de una varianza (Martínez, 2019, elaborado por Granda, 2020).

Solución:

Vamos a establecer la hipótesis:

$$H_0 : \sigma^2 = 64$$

$$H_0 : \sigma^2 \neq 64$$

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:

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

En la tabla de distribución de chi cuadrada (X_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 $X_s^2 = 27.204$.

$$\frac{(20-1)64}{27.204} < \sigma^2 < \frac{(20-1)64}{11.651}$$

$$\frac{1216}{27.204} < \sigma^2 < \frac{1216}{11.651}$$

Conclusión:

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.

Al finalizar el trabajo correspondiente a la semana trece, en la cual se abordaron otras pruebas de hipótesis y pruebas no paramétricas, le recuerdo la importancia de revisar los conceptos y las aplicaciones teórico-prácticas de estas pruebas en la bibliografía y recursos sugeridos.

Actividades de aprendizaje recomendadas

A continuación, le invito a realizar las actividades asignadas para reforzar lo estudiado:

1. Observar cuidadosamente los siguientes videos:
   - "Cómo usar la Tabla Chi-cuadrado"
   - "Prueba de hipótesis para la varianza".
2. Resolver la práctica titulada "Datos sobre la distribución Chi-cuadrado", disponible en el sitio web de OpenStax.
3. Revisar y resolver las actividades de la herramienta GeoGebra titulada "Distribución Chi-cuadrado". En esta herramienta encontrará el material necesario para reforzar la teoría de esta distribución.

Nota: por favor complete las actividades en un cuaderno o documento Word.

RETROALIMENTACIÓN

El video explica detalladamente el proceso para llevar a cabo una prueba de hipótesis para la varianza, mientras que el texto y los materiales recomendados complementan los pasos necesarios para resolver problemas de aplicación, utilizando ejemplos desarrollados de manera clara y paso a paso. Estas herramientas le permitirán comprender a profundidad los temas abordados en esta unidad.

¡Felicitaciones por el esfuerzo y dedicación que ha demostrado en su proceso de aprendizaje!

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

Contenidos, recursos y actividades de aprendizaje recomendadas

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 video 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.

Estimado estudiante, para analizar las propiedades de la Distribución F, es fundamental revisar las páginas 410-414 del libro Estadística y muestreo de Martínez (2018). En este apartado encontrará el fundamento teórico, junto con una variedad de ejemplos que le permitirán estudiar y comprender esta prueba de manera efectiva.

Los pasos a seguir en esta prueba son:

1. Plantear las hipótesis H_0 y H_a

$$H_0 : \sigma_1^2 = \sigma_2^2$$

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

2. Establecer el nivel de significación $\alpha = \ldots$
3. Determinar el valor de F en la variante estadística.

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

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

Finalmente, recuerde consultar la tabla F al final del texto Estadística y muestreo de Martínez (2018), ubicada en la página 411. El procedimiento para utilizar dicha tabla se explica a continuación mediante el desarrollo de un problema modelo, el cual le sugiero revisar y analizar detenidamente antes de realizar las actividades programadas para esta semana.

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: $\bar{x} = 725;\ S_x^2 = 61;\ \bar{y} = 661;\ S_y^2 = 86$; 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: $n_x = 7,\ n_y = 13,\ \bar{x} = 725;\ S_x^2 = 61,\ \bar{y} = 661;$

$S_x^2 = 86,\ v_x = n - 1 = 7 - 1 = 6,\ v_y = n - 1 = 13 - 1 = 12$

Solución:

Solución Problema Modelo 20: Distribución F

Posteriormente, se explica el proceso para aceptar o rechazar la hipótesis nula del problema, aplicando la distribución F (Martínez, 2019, elaborado por Granda, 2020).

Vamos a establecer la hipótesis:

$H_0 : \sigma_1^2 < \sigma_2^2$

$H_a : \sigma_2^2 < \sigma_1^2$

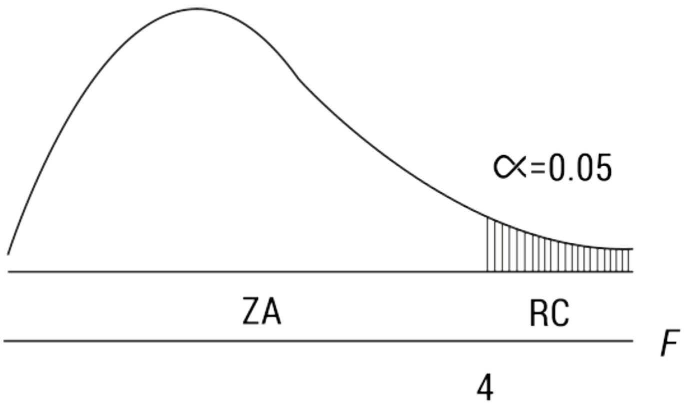
Tomando en cuenta el nivel de significación a = 0.05 Formulamos la variante estadística.

$$F = \frac{S_y^2}{S_x^2} = \frac{86}{61} = 1.41$$

El valor crítico de la distribución F, correspondiente a un nivel de significancia de 0.05, se obtiene de la tabla de valores críticos ubicada al final del libro Estadística y muestreo de Martínez (2018). Este valor es 4, considerando 12 grados de libertad en el numerador y 6 grados de libertad en el denominador, como corresponde en la figura dada a continuación.

Figura 18

Representación típica de una prueba de hipótesis unilateral con distribución F.



Nota. Adaptado de *Estadística y muestreo: 13 ed.* [Ilustración], por Martínez, C., 2019, Ecoe Ediciones, CC BY 4.0.

Conclusión:

Estos resultados permiten aceptar H_O, es decir, la primera prueba no presenta mayor variabilidad que la segunda al nivel del 5% de significancia.

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 v alores 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 realizar la prueba de coeficiente de correlación Pearson, por lo que lo invito a observar el video 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 : coeficiente de correlación $\quad n$: # de pares observados

- También: $t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{n\Sigma\, x_i y_j - (\Sigma\, x_i)(\Sigma\, y_j)}{\sqrt{[n\Sigma\, x_i^2 - (\Sigma\, x_i)^2][n\Sigma\, y_i^2 - (\Sigma\, y_i)^2]}}$
- Para estos casos se considera que la Hipótesis nula H0 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 la bibliografía sugerida.

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 $n = 18, a = 5\%$

Solución:

Solución Problema Modelo 21: Prueba del coeficiente de correlación de Pearson

Posteriormente, 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.

Vamos a establecer la hipótesis:

$$H_0 : R = 0 \longrightarrow \text{no existe correlación}$$

$$H_{\alpha}: R \neq 0 \rightarrow \text{el coeficiente de correlación es significativo}$$

O también para poblaciones tenemos:

$$H_0: p = 0$$

$$H_1: p > 0$$

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

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

Calculamos el parámetro estadístico:

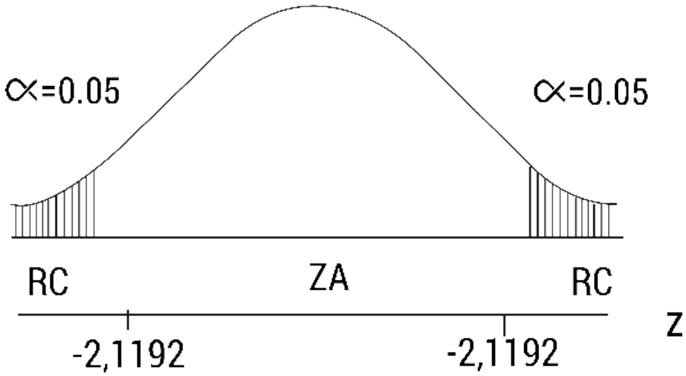
$$t = r\sqrt{\frac{n-2}{1-r^2}}$$

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

Para determinar el valor crítico, se encuentra con la distribución t de Student y debe ser una prueba bilateral, recuerde que, para esta prueba $\alpha = \frac{0.05}{2} = 0.025$, como se ilustra en la siguiente figura.

Figura 19

Representación típica de una prueba de hipótesis bilateral con distribución normal



Nota. Tomado de *Estadística y muestreo: 13 ed. (p. 578)* [Ilustración], por Martínez, C., 2019, Ecoe Ediciones, CC BY 4.0.

Conclusión:

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.

Al finalizar el trabajo de la semana catorce, donde se estudió la comparación entre varianzas de dos poblaciones y la prueba del coeficiente de correlación de Pearson, se recomienda revisar en la bibliografía sugerida y en los materiales proporcionados los conceptos y aplicaciones teórico-prácticas de estas pruebas de hipótesis.

Actividades de aprendizaje recomendadas

Posteriormente, le invito a realizar las siguientes actividades para afianzar sus conocimientos:

1. Observar cuidadosamente el siguiente video: Prueba de hipótesis para la varianza de dos poblaciones.
2. Leer el capítulo 9 del texto Estadística y muestreo de Martínez (2018), disponible en la biblioteca virtual de la universidad. Enfóquese en las secciones dedicadas a la distribución F y a la prueba del coeficiente de correlación de Pearson.
3. Revisar el proceso del video para realizar la Prueba de hipótesis de la correlación.
4. Resolver los problemas 18 y 19 del libro Estadística y muestreo de Martínez (2018), disponibles en la página 418, utilizando métodos analíticos y herramientas tecnológicas.

Nota: por favor complete la actividad en un cuaderno o documento Word.

RETROALIMENTACIÓN

Los videos explican 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!

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

Resultado de aprendizaje 4:

Utiliza la prueba de hipótesis con la distribución ji cuadrada a casos aplicados.

Para alcanzar el resultado de aprendizaje planteado, usted aplicará la prueba de hipótesis con la distribución ji-cuadrada a diversos casos prácticos. A través del estudio de pruebas con observaciones apareadas y pruebas no paramétricas, comprenderá cómo analizar datos cuando no se cumplen los supuestos de normalidad. Mediante actividades prácticas y ejercicios aplicados, desarrollará la habilidad de interpretar resultados y tomar decisiones basadas en evidencia estadística.

Contenidos, recursos y actividades de aprendizaje recomendadas

Recuerde revisar de manera paralela los contenidos con las actividades de aprendizaje recomendadas y actividades de aprendizaje evaluadas.

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. ¡Bienvenido!

Unidad 4. Otras Pruebas de hipótesis

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 video 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 que se presenta en la tabla siguiente:

Tabla 10

Fórmulas para pruebas con observaciones apareadas

Fórmula matemática	Descripción e interpretación
$d_i = x_i - y_i$	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	Número 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
$t = \frac{\bar{d} - a_d}{\frac{S_d}{\sqrt{n}}} = \frac{\bar{d} - 0}{S_{\bar{d}}}$	Variante estadística
$S_{\bar{a}} = \frac{S_d}{\sqrt{n}}$	Error estándar en observaciones apareadas.

Nota. Adaptado de *Estadística y muestreo* (14a. ed.), por Martínez, C., 2019, Ecoe Ediciones.

Esta prueba tiene aplicaciones muy relevantes; por ello, le recomiendo consultar la explicación en las páginas 419-423 del texto Estadística y muestreo de Martínez (2018). En estas páginas encontrará la fundamentación teórica de la prueba y la simbología correspondiente, la cual será empleada en la resolución de un ejemplo modelo. Le invito a seguirlo paso a paso y a analizar los problemas resueltos que se presentan en la bibliografía sugerida.

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 que se presentan en la siguiente tabla:

Tabla 11

Datos comparativos de producción entre variedad nueva de manzana y variedad corriente

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

Nota. Adaptado de *Estadística y muestreo* (14a. ed.), por Martínez, C., 2019, Ecoe Ediciones.

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

Solución:

Solución Problema Modelo 22: Prueba con observaciones apareadas

Seguidamente, se explica el proceso realizar una prueba de observaciones apareadas en problema de aplicación.

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

Tabla 12

Parámetros para aplicar la prueba de observaciones pareadas

HUERTO	x_i	y_i	$d_i = x_i — y_i$	$d_i — (\bar{d}_i)$	$(d - i — \bar{d})^2$
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
Σ	-	-	8	0	20

Nota. Adaptado de *Estadística y muestreo* (14a. ed.), por Martínez, C., 2019, Ecoe Ediciones.

En segundo lugar, se realiza el cálculo de la desviación típica en observaciones apareadas

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

Para este cálculo, se requiere la media aritmética de las diferencias

$$\bar{d} = \frac{\sum d_i}{n} \rightarrow \bar{d} = \frac{8}{8} = 1$$

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

Con este valor procedemos a determinar el error estándar de las observaciones propuestas: $S_{\bar{d}} = \frac{S_d}{\sqrt{n}} = \frac{1.69}{\sqrt{8}} = 0.6$

Luego, establecemos las hipótesis de prueba

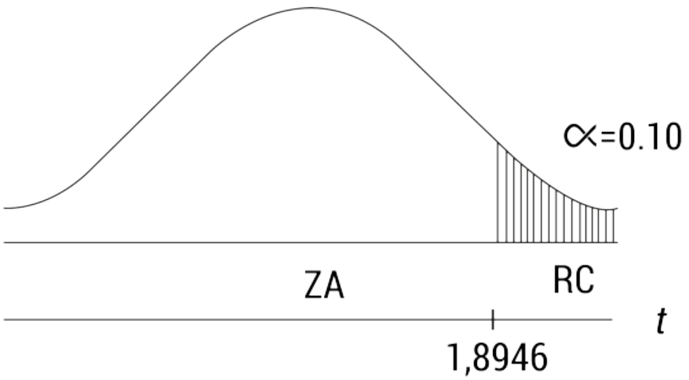
$$H_0 = a_d = 0$$

$$H - \alpha = a_d > 0$$

Los grados de libertad $v = n—2 = 8—1 = 7$, y con un nivel de significación α= 0.10, obtenemos el valor crítico t = 1.8946, como se muestra en la figura a continuación.

Figura 20

Representación típica de una prueba de hipótesis unilateral con distribución t



Nota. Adaptado de *Estadística y muestreo* (p. 415), por Martínez, C. (2019), Ecoe Ediciones.

Con todo esto, se determina el parámetro o estadístico "t":

$$t = \frac{d}{S_d} = \frac{1}{0.60} = 1.67$$

Conclusión:

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

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 X^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 X^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.

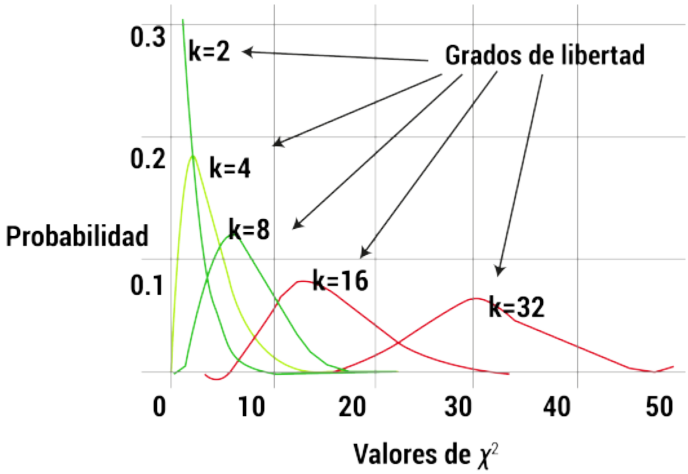
Propiedades

- 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.

La figura 21 muestra la distribución de probabilidad continua con un parámetro k que representa los grados de libertad de la variable aleatoria.

Figura 21

Función de distribución de probabilidad



Nota. Adaptado de *Estadística (13ª. ed.)* [Ilustración], por Triola, M., 2013, Pearson Educación, CC BY 4.0.

Esta prueba no paramétrica es una de las más utilizadas, como explica Martínez (2019) "$Ji—cuadrado$ 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). Asimismo, le invito a revisar el video de estudio Estadístico ji-cuadrada para prueba de hipótesis, esto lo preparará para emprender en la siguiente sección de estudio.

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

$$F_o = n_i = \text{Frecuencia observada o real.}$$

$$F_e = n_i^* = Frecuencia\ te\acute{o}rica\ o\ esperada$$

La fundamentación teórica de esta prueba se encuentra en el libro Estadística y muestreo de Martínez (2018), específicamente en las páginas 427-430. Se recomienda realizar una lectura comprensiva y analítica de este apartado para lograr una mejor comprensión del tema. La fundamentación teórica de esta prueba se encuentra en el libro Estadística y muestreo de Martínez (2018), específicamente en las páginas 427-430. Se recomienda realizar una lectura comprensiva y analítica de este apartado para lograr una mejor comprensión del tema.

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 X^2

A continuación, se explicará este proceso mediante la resolución de un problema modelo. Le invito a seguirlo y analizarlo con atención, así como a revisar los problemas resueltos en la bibliografía y los materiales educativos sugeridos. Para complementar y reforzar la comprensión de estos conceptos, le propongo revisar el siguiente módulo didáctico, donde se desarrolla un ejercicio práctico.

Problema Modelo #23: Pruebas no paramétricas-Prueba de Chi- Cuadrado χ

Actividades de aprendizaje recomendadas

Al finalizar la semana quince, y después de haber analizado las pruebas con observaciones apareadas y la prueba no paramétrica Chi-Cuadrado, es fundamental consolidar estos conocimientos. Para ello, le recomiendo revisar el sitio web de OpenStax, específicamente la sección 11.1 titulada

Datos sobre la distribución Chi-cuadrado. Una vez completado este repaso, proceda a desarrollar las actividades que se presentan a continuación:

1. Observar cuidadosamente el siguiente video: Prueba Chi-Cuadrada, la prueba no paramétrica más utilizada.
2. Leer el capítulo 9, titulado Otras pruebas de hipótesis, del texto Estadística y muestreo de Martínez (2018). Preste especial atención a las secciones dedicadas a la prueba con observaciones apareadas y a la prueba no paramétrica Chi-Cuadrado.
3. Desarrollar la práctica sobre la distribución Chi-Cuadrado, disponible en el sitio web de OpenStax.

   Nota: Por favor complete la actividad en un cuaderno o documento Word.

   RETROALIMENTACIÓN

   El video ofrece una explicación clara del proceso para realizar una prueba de Chi-Cuadrado. Además, la bibliografía y los recursos en línea recomendados presentan los pasos necesarios para resolver problemas de aplicación relacionados con el contenido estudiado esta semana. A través de ejemplos desarrollados de forma detallada y paso a paso, tendrá las herramientas necesarias para completar las actividades planteadas en esta unidad.

   ¡Felicitaciones! Su dedicación y trabajo responsable le han permitido avanzar significativamente en su aprendizaje.

4. Estimado estudiante, al finalizar el estudio de la cuarta unidad, Otras *Pruebas de Hipótesis*, le sugiero realizar una autoevaluación de los conocimientos adquiridos. Este ejercicio le permitirá identificar y reforzar los aspectos que requieran mayor atención, promoviendo una comprensión más sólida de los temas abordados.

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.

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

II.

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

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

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

III.

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

a. VARIANTE ESTADÍSTICA
b. HIPÓTESIS ALTERNATIVA
c. HIPÓTESIS NUL A

3. Al docimar la hipótesis σ = 20, dada Ŝ = 10 para una muestra de tamaño 25. Se establece las hipótesis:

A.
$$\{H_0 \frac{\sigma^2}{\sigma_o^2} = \frac{\sigma^2}{20^2} = 1 \; H_\alpha : \frac{\sigma^2}{20^2} \neq 1\}$$

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

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

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 el salario fue de σ² = 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%.

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
.00	.000
.01	.010
.02	.020
.03	.030
.04	.040
.05	.050
.06	.060
.07	.070
.08	.080
.09	.090
.10	.100

r	z
.36	.377
.37	.388
.38	.400
.39	.412
.40	.424
.41	.436
.42	.448
.43	.460
.44	.472
.45	.485

r	z
.71	.887
.72	.908
.73	.929
.74	.950
.75	.973
.76	.996
.77	1.020
.78	1.045
.79	1.071
.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, cuy as 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 X^2.
   B. Calcular χ², 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 X^2.

9. Una estadística de accidentes leves, ocurridos en dos fábricas A y B muestra 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. $H_0 : n_i \neq n \cdot i$
   B. $H_0 : n_i = n \cdot i$
   C. $H_0 : n_i > n \cdot i$

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

    - X^2 puede tener valores negativos
    - Si se tiene 4 filas y 3 columnas, los grados de libertad son 5

- La suma de los valores esperados es igual a la suma de los valores observados

A. FFV
B. FVF
C. VVF

Ir al solucionario

Si surgen dudas en una o más preguntas vuelva a leer el contenido científico para que identifique la validez de su respuesta.

Estimado estudiante, le recuerdo que cuenta con el horario de tutoría, durante el cual su profesor tutor está disponible para atender sus dudas e inquietudes. Este espacio es ideal para resolver cualquier dificultad que surja durante la semana de trabajo y para validar su desempeño.

Finalmente, le animo a participar activamente en las actividades síncronas planificadas a lo largo de la asignatura, ya que estas representan una excelente oportunidad para superar barreras de tiempo y espacio, contribuyendo a optimizar sus resultados de aprendizaje.

“La motivación es lo que te pone en marcha, el hábito es lo que hace que sigas.”

— Jim Rohn

Resultado de aprendizaje 2 a 4:

- Valida las hipótesis que se generan en cada uno de los casos analizados.
- Emplea información no paramétrica para análisis de datos.
- Utiliza la prueba de hipótesis con la distribución ji cuadrada a casos aplicados.

Contenidos, recursos y actividades de aprendizaje recomendadas

Semana 16

Actividades finales del bimestre

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 anterior es, para consolidar sus aprendizajes del segundo bimestre, y prepararse a la prueba presencial en físico o virtual, sugiero realizar las siguientes actividades:

1. Revise cada uno de los conceptos estudiados en las dos unidades planificadas y desarrolladas en este segundo bimestre.
2. 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 de la bibliografía y recursos sugeridos.
3. 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.

Nota: por favor complete las actividades en un cuaderno o documento Word.

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

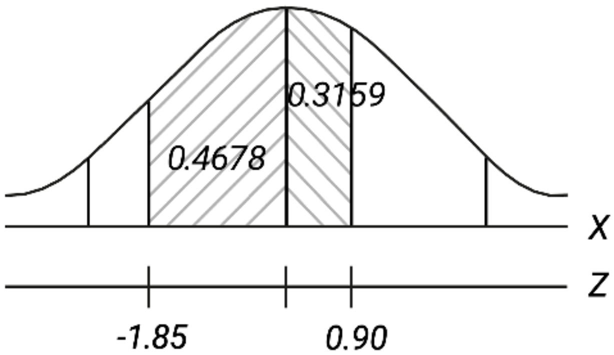
"No dejes que lo que no puedes hacer interfiera con lo que puedes hacer."

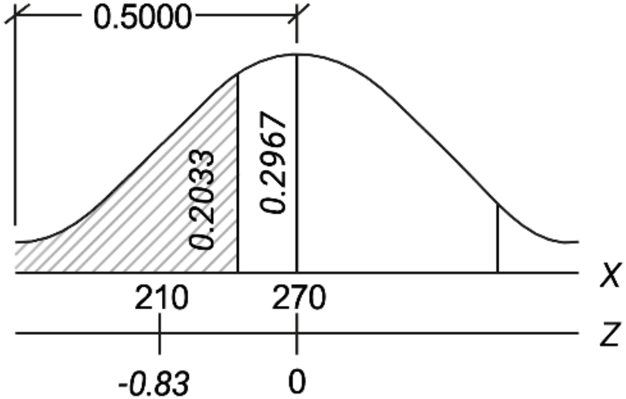
— John R. Wooden

4. Autoevaluaciones

Autoevaluación 1

Pregunta	Respuesta	Retroalimentación
1	B	En el caso del lanzamiento de una moneda, esta pueda ser: cara o sello, por lo tanto, son sucesos mutuamente excluyentes.
2	A	Aplicando $P_{(x=2)} = (\frac{4!}{2!2!})(\frac{1}{2})^2(\frac{1}{2})^{4-2} = \frac{6}{16} = 0.375 = 37.5\%$
3	B	Aplicando $= 8, p = 0.8(ganar), q = 1 - 0.8 = 0.2, x = 0,1,2,3,4,5y6, P(x \leq 6) = ?$ $P(x \leq 6) = P(0) + P(1) + P(3) + P(4) + P(5) + P(6) = 1$ $— [P(7) + P(8)]\textbf{(aplicando complemento)}$ $P(x \leq 6) = 1 — [(87)(0.8)7(0.2)1 + (88)(0.8)8(0.2)0] = 1 — [0.3355 + 0.1678]$ $P(x \leq 6) = 49.67$
4	B	Calculando la media $\lambda = np$ para p = 0.03 y n = 100 tenemos que $\lambda = 0.03(100) = 3$. Aplicando Poisson: $P_{(x=2)} = \frac{3^2e^{-3}}{2!} = \frac{3^2(0.04979)}{2!} = 0.2240 = 22.4\%$
5	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 sí se escribe FALSO, la función devolverá solo de $P_{(x=4)} = P_{(4)}$
6	C	Respuesta: C. y n = 200000 tenemos que $\lambda = 0.00003(200000) = 6$. Aplicando Poisson para $P_{(x\leq2)}$: $P_{(x\leq2)} = \frac{6^0e^{-6}}{0!} + \frac{6^1e^{-6}}{1!} + \frac{6^2e^{-6}}{2!} = 0.002479 + 0.014874 + 0.044622 = 6.20\%$

Pregunta	Respuesta	Retroalimentación
7	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	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.

<table>
<tr><th>Pregunta</th><th>Respuesta</th><th>Retroalimentación</th></tr>
<tr><td>9</td><td>A</td><td>Con $\mu = 4.5 \text{ min} = 270 \text{ s}, \quad \sigma = 72 \text{ s}, \quad P(x \leq 3.5 \text{ min}) = P(x \leq 210 \text{ s}) = ?$
$Z = \frac{X-\mu}{\sigma} = \frac{210-270}{72} = -0.83 \rightarrow A(0.2967)$
$P(x \leq 210 \text{ s}) = 0.5000 - 0.2967 = 0.2033 = 20.33\%$

</td></tr>
<tr><td>10</td><td>A. i-b, ii-c, iii-a</td><td>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.</td></tr>
</table>

Ir a la autoevaluación

Autoevaluación 2

Pregunta	Respuesta	Retroalimentación
1	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	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	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	A	**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 tendrá 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.

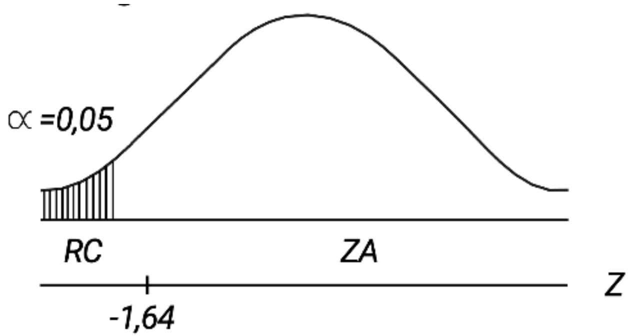
Pregunta	Respuesta	Retroalimentación
5	A	Con $\mu = 864500, \quad \sigma = 15000, \quad n = 25$ $Z = \frac{857500-864500}{\frac{15000}{\sqrt{25}}} = -2,33$ Entonces el área vara este valor de Z es A=0.4901 y la probabilidad es igual a P = 0,5000 — 0,4901 = 0,0099, finalmente$P_x < 857500 = 0,99$
6	A	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	A	La desviación típica de las diferencias entre los pares de medias muestrales. $\sigma_{\overline{x}-\overline{y}} = \sqrt{\frac{\sum[(\overline{x}_i-\overline{y}_i)-(\mu_x-\mu_y)]^2}{M}}$
8	B	En la distribución de diferencias 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	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	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

Autoevaluación 3

Pregunta	Respuesta	Retroalimentación
1	A	La inferencia estadística comprende dos partes principales, a saber: la estimación de parámetros y la prueba o docimasia de hipótesis.
2	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	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	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 (H_a) cuando se ha debido rechazar. En el ejercicio que estamos desarrollando, sería: “Aceptar la moneda como correcta, cuando en verdad no lo es”. H0: 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 (H_a) cuando se ha debido aceptar. “Rechazar la moneda como incorrecta, cuando en verdad está equilibrada”.
5	C	En las pruebas de hipótesis estadísticas, la tabla le muestra las consecuencias de aceptar o rechazar hipótesis que pueden ser verdaderas o falsas. Para responder correctamente, usted debe identificar qué combinación de celdas representa situaciones complementarias considerando que los errores Tipo I (rechazar hipótesis verdadera) y Tipo II (aceptar hipótesis falsa) están relacionados con el nivel de significación de la prueba.
6	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, cometeríamos 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.

Pregunta	Respuesta	Retroalimentación
7	A	Es una hipótesis causal y a 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	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 μ ≤ 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 $\mu \leq 195$ libras. Asimismo, 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	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 aplicará 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 más 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	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	C	En el caso planteado utilizamos la distribución de medias muestrales $Z = \frac{\bar{z}-\mu}{\frac{s}{\sqrt{n}}}$

<table>
<tr><th>Pregunta</th><th>Respuesta</th><th>Retroalimentación</th></tr>
<tr><td>12</td><td>A</td><td>
Hipótesis nula $H_0 : P < 0.14$
Hipótesis alternativa $H_\alpha : P < 0.14$
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 Zs < 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>B</td><td>
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>
</table>

Pregunta	Respuesta	Retroalimentación
14	B	$Z=\frac{p_1-p_2}{\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}}=\frac{0{,}62-0{,}96}{\sqrt{\frac{0{,}62(0{,}38)}{130}+\frac{0{,}96(0{,}04)}{100}}}=-7.25$
15	A	$S_{\overline{x}-\overline{y}}=\sqrt{\frac{(n_1-1)S_x^2+(n_2-1)S_y^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$ $S_{\overline{x}-\overline{y}}=\sqrt{\frac{(15-1)\cdot 39{,}06+(15-1)\cdot 21.62}{15+15-2}}\sqrt{\frac{1}{15}+\frac{1}{15}}=5,51\cdot(0.365)=2.01$
		Ir a la autoevaluación

Autoevaluación 4

Pregunta	Respuesta	Retroalimentación
1	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	1-c, 2-b, 3-a	$H_0\colon \sigma^2 = \sigma_0^2 \leftrightarrow H_0\colon \frac{\sigma^2}{\sigma_0^2} = 1$ HIPÓTESIS NUL A $H_a \frac{\sigma^2}{\sigma_0^2} \neq 1$ $H_a \frac{\sigma^2}{\sigma_0^2} > 1$ HIPÓTESIS ALTERNATIVA $H_a \frac{\sigma^2}{\sigma_0^2} < 1$ $x^2 = \frac{(n-1)S^2}{\sigma_0^2}$ VARIANTE ESTADÍSTICA *Chi-cuadrado* con *n* − *1* grados de libertad
3	A	La hipótesis que se establece para este caso es $\{H_0 : \frac{\sigma^2}{\sigma_0^2} = \frac{\sigma^2}{20^2} = 1 \; H_a : \frac{\sigma^2}{20^2} \neq 1\}$
4	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	A	La desviación estándar se simboliza por σ z y se calcula así: $\sigma_z = \frac{1}{\sqrt{n-3}}$

<table>
<tr><th>Pregunta</th><th>Respuesta</th><th>Retroalimentación</th></tr>
<tr><td>6</td><td>B</td><td>

Tabla V transformación de r A z (es decir, z=0.5 $Log_e \frac{1+r}{1-r}$)

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

</td></tr>
<tr><td>7</td><td>C</td><td>

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 χ2, 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.

</td></tr>
<tr><td>8</td><td>A</td><td>

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

</td></tr>
</table>

Pregunta	Respuesta	Retroalimentación
9	B	Hipótesis nula $H_0 : n_i = n \cdot i, \quad \alpha = 0.05$
10	A	X^2 puede tener valores negativos **F** Si se tiene 4 filas y 3 columnas, los grados de libertad 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.

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

Obando López, J. & Arango Londoño, N. (2019). Probabilidad y estadística: (ed.). Fondo Editorial EIA. https://elibro.net/es/ereader/bibliotecautpl/125705

Martínez Bencardino, C. (2018). *Estadística y muestreo:* (13 ed.). Ecoe Ediciones. https://elibro.net/es/ereader/bibliotecautpl/131880

Islas Salomón, C. A. Colín Uribe, M. P. & Morales Téllez, F. (2018). *Proba bilidad y estadística:* (ed.). Grupo Editorial Éxodo. https://elibro.net/es/ereader/bibliotecautpl/128557

Vladimirovna Panteleeva, O. & Gutiérrez González, E. (2016). *Estadística inferencial 1 para ingeniería y ciencias:* (ed.). Grupo Editorial Patria. h ttps://elibro.net/es/ereader/bibliotecautpl/40474

Llinás Solano, H. (2017). *Estadística Inferencial:* (ed.). Universidad del Norte. https://elibro.net/es/ereader/bibliotecautpl/70060

Díaz Rodríguez, M. (2022). *Estadística inferencial aplicada:* (2 ed.). Universidad del Norte. https://elibro.net/es/ereader/bibliotecautpl/221614

Proaño Rivera, W. B. (2020). *Estadística descriptiva e inferencial:* (1 ed.). Universidad del Azuay. https://elibro.net/es/ereader/bibliotecautpl/233574