



SEP
SECRETARÍA DE
EDUCACIÓN PÚBLICA



SEV
Secretaría de
Educación de Veracruz



TECNM
INSTITUTO TECNOLÓGICO NACIONAL DE
MÉXICO



**INSTITUTO TECNOLÓGICO SUPERIOR
DE ALVARADO - Campus Medellín**

INGENIERÍA EN SISTEMAS COMPUTACIONALES

Materia:

Probabilidad y Estadística

Semestre:

4° Semestre.

Producto Académico:

Investigación U6.

Presenta(n):

Alejandro Zurita Pérez (226Z0200)

Ricardo de Jesús Hernández Ochoa (236Z0267)

Docente:

Xóchitl Romero Hipólito



Medellín de Bravo, Ver. ENE. – JUN. 2024



VERACRUZ
GOBIERNO
DEL ESTADO

INDICE

Introducción	2
Desarrollo	4
6.1 Muestreo	4
6.1.1 Tipos de Muestreo	6
6.1.2 Teorema de Límite Central.....	8
6.1.3 Distribución Muestral de la Media.....	9
6.1.4 Distribución Muestral de una Proporción	11
Conclusión General	13
Referencias Bibliográficas	15

Introducción

El muestreo es una técnica esencial en el campo de la estadística y juega un papel crucial en diversas disciplinas, incluyendo la investigación científica, la economía, la sociología, y la administración de negocios. Este proceso permite a los investigadores y profesionales extraer conclusiones válidas y representativas sobre una población entera sin necesidad de examinar cada uno de sus elementos. En otras palabras, el muestreo hace posible el estudio de grandes poblaciones de manera práctica y económica, al enfocarse solo en una porción seleccionada de la misma.

Imaginemos un escenario donde queremos entender las preferencias de todos los habitantes de una ciudad sobre un nuevo servicio público. Sería impráctico e imposible encuestar a cada individuo debido a las limitaciones de tiempo, recursos y accesibilidad. Aquí es donde el muestreo entra en juego. Al seleccionar y estudiar una muestra representativa de la población, podemos obtener una estimación precisa de las preferencias generales con un alto grado de confianza. Este enfoque no solo ahorra tiempo y dinero, sino que también proporciona resultados que son suficientemente precisos para tomar decisiones informadas.

Hay diferentes métodos de muestreo que se utilizan para garantizar que la muestra sea representativa de la población. Cada método tiene sus ventajas y desventajas, y la elección del método adecuado depende de varios factores, como la naturaleza de la población, el objetivo del estudio, y los recursos disponibles. Entre los métodos de muestreo más comunes se encuentran el muestreo aleatorio simple, el muestreo sistemático, el muestreo estratificado, y el muestreo por conglomerados. Comprender estos métodos y saber cuándo y cómo aplicarlos es fundamental para cualquier análisis estadístico.

Un concepto clave en el muestreo es el Teorema de Límite Central (TLC). Este teorema establece que, independientemente de la distribución original de la población, la distribución de la media muestral se aproximará a una distribución normal a medida que el tamaño de la muestra aumente. Este principio es vital porque permite a los estadísticos aplicar técnicas inferenciales basadas en la distribución normal, facilitando la estimación de parámetros poblacionales y la realización de pruebas de hipótesis.

Además del Teorema de Límite Central, las distribuciones muestrales de la media y de la proporción son conceptos fundamentales en la teoría del muestreo. La distribución muestral de la media se refiere a la distribución de las medias de todas las posibles muestras de un tamaño específico extraídas de una población. De manera similar, la distribución muestral de una proporción describe la distribución

de las proporciones obtenidas de todas las posibles muestras de un tamaño específico. Estas distribuciones permiten a los estadísticos calcular intervalos de confianza y realizar pruebas de hipótesis con un mayor grado de precisión y certeza.

La importancia del muestreo no se limita a la teoría estadística; sus aplicaciones prácticas son vastas y variadas. En el ámbito empresarial, el muestreo se utiliza para realizar estudios de mercado, evaluar la satisfacción del cliente, y mejorar la calidad del producto. En la investigación médica, permite llevar a cabo ensayos clínicos y estudios epidemiológicos que son esenciales para el desarrollo de tratamientos y políticas de salud pública. En las ciencias sociales, el muestreo es fundamental para entender comportamientos, actitudes, y tendencias dentro de diferentes grupos de la sociedad.

En conclusión, el muestreo es una herramienta indispensable en la estadística y en muchas otras áreas del conocimiento. Nos permite estudiar y entender poblaciones de manera eficiente y efectiva, proporcionando información valiosa que de otro modo sería inaccesible. A lo largo de este documento, exploraremos en detalle los diferentes tipos de muestreo, el Teorema de Límite Central, y las distribuciones muestrales de la media y de una proporción, con el objetivo de proporcionar una comprensión clara y práctica de estos conceptos esenciales.

Desarrollo

6.1 Muestreo

El muestreo es un procedimiento fundamental en la estadística que nos permite estudiar una población grande a partir de una pequeña porción de la misma, llamada muestra. Esta técnica es esencial cuando no es viable o práctico examinar a cada miembro de la población debido a limitaciones de tiempo, recursos y accesibilidad. Al elegir una muestra representativa, podemos hacer inferencias y tomar decisiones informadas sin necesidad de analizar cada individuo.

El objetivo principal del muestreo es obtener una muestra que sea un reflejo fiel de la población total. Esto significa que las características de la muestra deben ser proporcionales a las de la población para garantizar que los resultados del análisis sean válidos y aplicables a toda la población. El proceso de muestreo involucra varias etapas, desde la definición de la población hasta la selección de la muestra y el análisis de los datos.

Una de las decisiones más importantes en el muestreo es la elección del método adecuado. Existen varios métodos de muestreo, cada uno con sus ventajas y desventajas. La elección del método depende de factores como la estructura de la población, el objetivo del estudio y los recursos disponibles. Los métodos de muestreo se dividen principalmente en dos categorías: muestreo probabilístico y muestreo no probabilístico.

El muestreo probabilístico se basa en el principio de que cada miembro de la población tiene una probabilidad conocida y no nula de ser seleccionado en la muestra. Este tipo de muestreo incluye métodos como el muestreo aleatorio simple, el muestreo sistemático, el muestreo estratificado y el muestreo por conglomerados. Estos métodos son preferidos en estudios donde es crucial obtener una muestra representativa y se necesita hacer inferencias precisas sobre la población.

Por otro lado, el muestreo no probabilístico no proporciona a todos los miembros de la población una probabilidad conocida de ser seleccionados. Este tipo de muestreo incluye métodos como el muestreo por conveniencia, el muestreo por juicio y el muestreo por cuotas. Aunque estos métodos son más fáciles y rápidos de implementar, tienen el riesgo de introducir sesgos y no siempre garantizan una muestra representativa.

Además de elegir el método de muestreo, es crucial determinar el tamaño de la muestra. Un tamaño de muestra adecuado es necesario para garantizar que los resultados sean precisos y confiables. Un tamaño de muestra muy pequeño puede no captar la variabilidad de la población, mientras que un tamaño de muestra demasiado grande puede ser innecesariamente costoso y llevar más tiempo. La

determinación del tamaño de la muestra generalmente se basa en el nivel de confianza deseado, el margen de error aceptable y la variabilidad de la población.

En resumen, el muestreo es una herramienta esencial en la estadística que nos permite estudiar poblaciones grandes de manera eficiente y efectiva. La elección del método de muestreo y el tamaño de la muestra son decisiones críticas que afectan la validez y la precisión de los resultados. Comprender y aplicar correctamente estas técnicas es fundamental para cualquier análisis estadístico riguroso.



6.1.1 Tipos de Muestreo

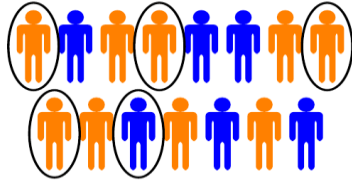
Existen varios tipos de muestreo que se utilizan en la estadística, cada uno con sus propias características y aplicaciones específicas. A continuación, se describen los tipos más comunes:

1. **Muestreo Aleatorio Simple:** En este método, cada miembro de la población tiene la misma probabilidad de ser seleccionado. Se utiliza un proceso aleatorio, como un sorteo o un generador de números aleatorios, para seleccionar los miembros de la muestra. Este método es fácil de entender y aplicar, pero puede ser ineficiente si la población es muy grande y dispersa.
2. **Muestreo Sistemático:** Consiste en seleccionar a cada n ésimo miembro de la población después de elegir un punto de partida aleatorio. Por ejemplo, si se desea una muestra del 10% de una población, se seleccionaría a cada décimo individuo. Este método es más sencillo de implementar que el muestreo aleatorio simple y es útil cuando la población es homogénea y está bien organizada.
3. **Muestreo Estratificado:** La población se divide en subgrupos homogéneos llamados estratos, y se toma una muestra aleatoria de cada estrato. Este método asegura que todos los subgrupos estén representados en la muestra y es útil cuando hay variabilidad significativa entre los estratos. Es comúnmente utilizado en estudios de opinión pública y encuestas de mercado.
4. **Muestreo por Conglomerados:** La población se divide en grupos heterogéneos llamados conglomerados, y se seleccionan uno o más conglomerados al azar para ser estudiados en su totalidad. Este método es eficiente cuando la población está dispersa geográficamente. Un ejemplo típico es un estudio en escuelas donde cada escuela representa un conglomerado.
5. **Muestreo por Conveniencia:** Se seleccionan miembros de la población que son más fáciles de acceder. Este método es rápido y económico, pero tiene un alto riesgo de sesgo y no siempre produce resultados representativos. Es comúnmente utilizado en estudios exploratorios y en fases iniciales de investigación.
6. **Muestreo por Juicio:** Se seleccionan miembros de la población basándose en el juicio del investigador sobre quiénes son los más adecuados para proporcionar la información requerida. Este método es subjetivo y depende de la experiencia del investigador, lo que puede introducir sesgos.
7. **Muestreo por Cuotas:** Se divide la población en subgrupos y se fija una cuota para cada subgrupo. Los miembros se seleccionan hasta que se cumple la cuota. Este método es similar al muestreo estratificado, pero no

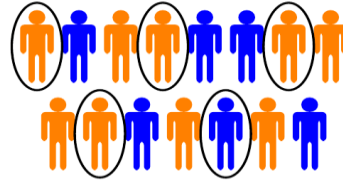
utiliza la aleatoriedad para la selección de los miembros dentro de cada cuota.

Cada uno de estos métodos tiene aplicaciones específicas y se elige según las necesidades y limitaciones del estudio. La clave es seleccionar el método que mejor garantice una muestra representativa y minimice los posibles sesgos.

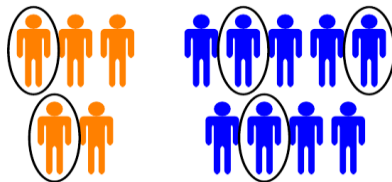
Muestreo aleatorio



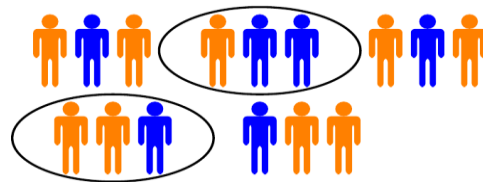
Muestreo sistemático



Muestreo estratificado



Muestreo por conglomerados



6.1.2 Teorema de Límite Central

El Teorema de Límite Central (TLC) es uno de los conceptos más importantes en la estadística. Este teorema establece que, bajo ciertas condiciones, la distribución de la media muestral de una variable aleatoria se aproximará a una distribución normal a medida que el tamaño de la muestra aumente, independientemente de la forma de la distribución original de la población.

Para entender el TLC, consideremos una población con cualquier distribución, no necesariamente normal, con una media μ y una desviación estándar σ . Si tomamos múltiples muestras aleatorias de tamaño n de esta población y calculamos la media de cada muestra, el TLC nos dice que la distribución de estas medias muestrales se aproximará a una distribución normal a medida que n se haga grande. Esta aproximación es válida incluso si la distribución original no es normal.

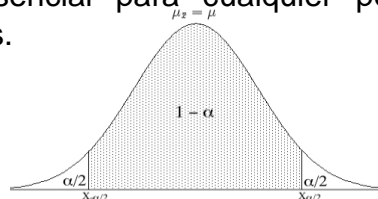
El TLC es crucial porque permite a los estadísticos hacer inferencias sobre la media poblacional utilizando la distribución normal, facilitando el cálculo de intervalos de confianza y la realización de pruebas de hipótesis. Esto simplifica enormemente el análisis estadístico, ya que la distribución normal está bien comprendida y dispone de tablas y fórmulas establecidas para su uso.

Hay algunas condiciones importantes que deben cumplirse para que el TLC sea aplicable:

1. **Independencia:** Las muestras deben ser independientes entre sí. Esto significa que la selección de un individuo no debe influir en la selección de otro.
2. **Tamaño de la Muestra:** Generalmente, se considera que un tamaño de muestra $n \geq 30$ es suficiente para que el TLC sea aplicable. Sin embargo, si la distribución original es muy asimétrica o tiene colas largas, puede requerirse un tamaño de muestra mayor.
3. **Sin sesgo extremo:** La población original no debe tener sesgos extremos, como una varianza infinita. En la práctica, esto rara vez es un problema.

El TLC no solo se aplica a las medias, sino también a otras estadísticas muestrales, como sumas y proporciones. Por ejemplo, la proporción de éxitos en una muestra grande de una distribución binomial también se aproximará a una distribución normal.

El TLC es fundamental para muchas aplicaciones estadísticas, incluyendo el control de calidad, la investigación científica, y las encuestas de opinión. Permite a los estadísticos y a los investigadores trabajar con muestras grandes y hacer inferencias precisas sobre la población de interés. La comprensión y aplicación del TLC es esencial para cualquier persona que trabaje con datos y análisis estadísticos.



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

6.1.3 Distribución Muestral de la Media

La distribución muestral de la media es una distribución teórica que describe las medias de todas las posibles muestras de un tamaño específico extraídas de una población. Esta distribución es fundamental en la estadística inferencial, ya que permite hacer inferencias sobre la media poblacional basándose en la media muestral.

Para ilustrar este concepto, supongamos que tenemos una población con una media μ y una desviación estándar σ . Si tomamos una muestra aleatoria de tamaño n de esta población y calculamos la media de la muestra (\bar{X}), esta media muestral será una estimación de la media poblacional. Si repetimos este proceso muchas veces, obteniendo múltiples muestras de tamaño n y calculando la media de cada una, las medias muestrales formarán una distribución.

El Teorema de Límite Central nos dice que esta distribución muestral de la media se aproximará a una distribución normal, con una media igual a la media poblacional (μ) y una desviación estándar igual a la desviación estándar poblacional dividida por la raíz cuadrada del tamaño de la muestra ($\frac{\sigma}{\sqrt{n}}$). Esta desviación estándar de la distribución muestral se conoce como error estándar de la media.

La fórmula para el error estándar de la media es:

$$\text{Error Estándar} = \frac{\sigma}{\sqrt{n}}$$

El error estándar es una medida de la precisión con la que la media muestral estima la media poblacional. Un error estándar más pequeño indica una estimación más precisa. A medida que el tamaño de la muestra aumenta, el error estándar disminuye, lo que significa que las medias muestrales estarán más estrechamente agrupadas alrededor de la media poblacional.

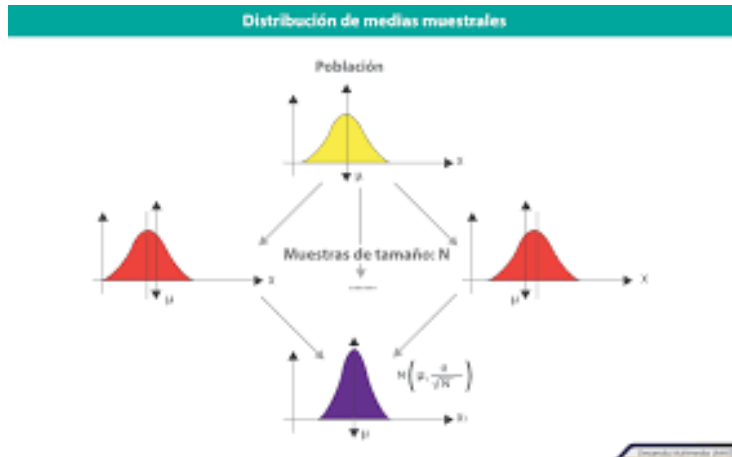
La distribución muestral de la media es crucial para la construcción de intervalos de confianza y la realización de pruebas de hipótesis. Un intervalo de confianza proporciona un rango de valores dentro del cual es probable que se encuentre la media poblacional. Por ejemplo, un intervalo de confianza del 95% se calcula como:

$$\bar{X} \pm Z \times \frac{\sigma}{\sqrt{n}}$$

donde Z es el valor crítico de la distribución normal estándar correspondiente al nivel de confianza deseado.

En las pruebas de hipótesis, la distribución muestral de la media se utiliza para determinar la probabilidad de observar una media muestral particular, dado que se conoce o se asume un valor para la media poblacional. Esto permite a los estadísticos evaluar la evidencia contra una hipótesis nula y decidir si deben rechazarla en favor de una hipótesis alternativa.

En resumen, la distribución muestral de la media es una herramienta poderosa en la estadística inferencial. Proporciona una base para hacer estimaciones precisas y realizar inferencias sobre la media poblacional a partir de muestras aleatorias. La comprensión de este concepto es esencial para cualquier análisis estadístico riguroso.



6.1.4 Distribución Muestral de una Proporción

La distribución muestral de una proporción es una distribución teórica que describe las proporciones de éxito en todas las posibles muestras de un tamaño específico extraídas de una población. Este concepto es fundamental en la estadística inferencial, especialmente en estudios donde se analizan datos categóricos, como encuestas de opinión y estudios de prevalencia.

Para entender la distribución muestral de una proporción, consideremos una población en la que una proporción p de individuos tiene una característica particular (éxito). Si tomamos una muestra aleatoria de tamaño n de esta población y calculamos la proporción muestral (\hat{p}) de éxitos en la muestra, esta proporción muestral será una estimación de la proporción poblacional.

Al igual que con la media muestral, si repetimos el proceso de muestreo muchas veces, obteniendo múltiples muestras de tamaño n y calculando la proporción de éxitos en cada una, las proporciones muestrales formarán una distribución. Según el Teorema de Límite Central, cuando n es suficientemente grande, la distribución de la proporción muestral \hat{p} se aproximará a una distribución normal con una media igual a la proporción poblacional (p) y una desviación estándar igual a $\sqrt{\frac{p(1-p)}{n}}$.

La fórmula para el error estándar de la proporción es:

$$\text{Error estándar} = \sqrt{\frac{p(1-p)}{n}}$$

El error estándar de la proporción es una medida de la precisión con la que la proporción muestral estima la proporción poblacional. Un error estándar más pequeño indica una estimación más precisa. A medida que el tamaño de la muestra aumenta, el error estándar disminuye, lo que significa que las proporciones muestrales estarán más estrechamente agrupadas alrededor de la proporción poblacional.

La distribución muestral de una proporción es fundamental para la construcción de intervalos de confianza y la realización de pruebas de hipótesis en estudios de datos categóricos. Un intervalo de confianza para la proporción poblacional se calcula como:

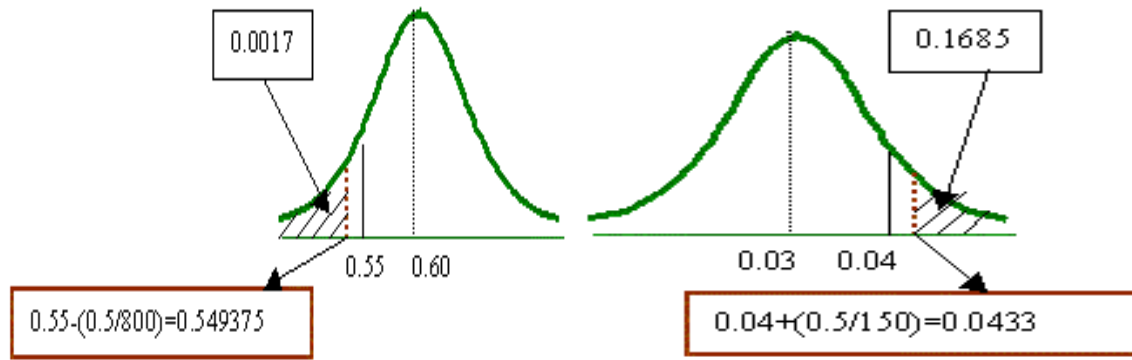
$$\hat{p} \pm Z \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

donde \hat{p} es la proporción muestral y Z es el valor crítico de la distribución normal estándar correspondiente al nivel de confianza deseado.

En las pruebas de hipótesis, la distribución muestral de una proporción se utiliza para determinar la probabilidad de observar una proporción muestral particular, dado que se conoce o se asume un valor para la proporción poblacional. Esto

permite a los estadísticos evaluar la evidencia contra una hipótesis nula y decidir si deben rechazarla en favor de una hipótesis alternativa.

En conclusión, la distribución muestral de una proporción es una herramienta esencial en la estadística inferencial para el análisis de datos categóricos. Proporciona una base para hacer estimaciones precisas y realizar inferencias sobre la proporción poblacional a partir de muestras aleatorias. La comprensión de este concepto es fundamental para cualquier análisis estadístico riguroso.



Conclusión General

El muestreo y sus técnicas son fundamentales en la estadística, permitiéndonos estudiar grandes poblaciones de manera eficiente y efectiva. A través de los diversos métodos de muestreo, podemos obtener muestras representativas que reflejan fielmente las características de la población, lo que nos permite hacer inferencias y tomar decisiones informadas.

En este trabajo, hemos explorado varios aspectos cruciales del muestreo, comenzando con una introducción al concepto y su importancia. Posteriormente, profundizamos en los diferentes tipos de muestreo, cada uno con sus ventajas y desventajas, y su aplicabilidad en distintos contextos. Entendimos que la elección del método de muestreo depende de la estructura de la población, el objetivo del estudio y los recursos disponibles. Desde el muestreo aleatorio simple, que asegura una representación imparcial de la población, hasta el muestreo por conveniencia, que, aunque rápido y económico, puede introducir sesgos significativos.

El Teorema de Límite Central (TLC) fue otro tema fundamental abordado en este trabajo. Este teorema nos proporciona una herramienta poderosa para hacer inferencias sobre la media poblacional a partir de muestras, independientemente de la forma de la distribución original. La comprensión del TLC nos permite utilizar la distribución normal para realizar cálculos precisos de intervalos de confianza y pruebas de hipótesis, simplificando el análisis estadístico y mejorando la precisión de nuestras conclusiones.

La distribución muestral de la media y la distribución muestral de una proporción son conceptos esenciales que permiten a los estadísticos y a los investigadores hacer estimaciones precisas sobre las poblaciones. La distribución muestral de la media nos ayuda a comprender cómo varían las medias muestrales alrededor de la media poblacional, y el error estándar juega un papel crucial en determinar la precisión de estas estimaciones. De manera similar, la distribución muestral de una proporción nos permite analizar datos categóricos y hacer inferencias sobre la proporción poblacional con un alto grado de precisión.

En la práctica, estos conceptos y técnicas tienen una amplia gama de aplicaciones. En el campo de la investigación científica, el muestreo es vital para realizar estudios clínicos y ensayos controlados, donde no es práctico examinar a todos los pacientes. En la investigación de mercado, las empresas utilizan el muestreo para comprender las preferencias y comportamientos de los consumidores, lo que les permite tomar decisiones estratégicas basadas en datos representativos. En las encuestas de opinión pública, el muestreo proporciona una visión precisa de las actitudes y opiniones de la población, lo que es crucial para la toma de decisiones políticas y sociales.

A pesar de su importancia, es fundamental reconocer las limitaciones y desafíos asociados con el muestreo. La selección de una muestra adecuada y la minimización de sesgos son aspectos críticos que deben gestionarse cuidadosamente. Los errores en el diseño del muestreo o en la ejecución del muestreo pueden conducir a resultados incorrectos y conclusiones equivocadas. Por lo tanto, es esencial que los estadísticos y los investigadores sigan buenas prácticas y principios éticos en la realización de estudios basados en el muestreo.

En conclusión, el muestreo es una herramienta indispensable en la estadística que nos permite estudiar poblaciones grandes de manera eficiente y precisa. La comprensión de los diferentes tipos de muestreo, el Teorema de Límite Central, y las distribuciones muestrales de la media y de una proporción son fundamentales para cualquier análisis estadístico riguroso. Estos conceptos y técnicas no solo facilitan la investigación y la toma de decisiones en una variedad de campos, sino que también contribuyen significativamente al avance del conocimiento y la comprensión de fenómenos complejos. La aplicación cuidadosa y ética de estas herramientas asegura que las conclusiones derivadas sean válidas y aplicables, beneficiando así a la sociedad en general.

Referencias Bibliográficas

1. Muestreo y Tipos de Muestreo

- Cochran, W. G. (1977). *Técnicas de muestreo* (3rd ed.). Wiley.
- Levy, P. S., & Lemeshow, S. (1999). *Muestreo de Poblaciones: Métodos y Aplicaciones*. (3rd ed.). Wiley.

2. Teorema de Límite Central

- Hogg, R. V., McKean, J., & Craig, A. T. (2005). *Introducción a la estadística matemática* (6th ed.). Pearson.
- Larsen, R. J., & Marx, M. L. (2012). *Una introducción a la estadística matemática y sus aplicaciones* (5th ed.). Pearson.

3. Distribución Muestral de la Media

- Devore, J. L. (2012). *Probabilidad y Estadística para la Ingeniería y las Ciencias* (8th ed.). Cengage Learning.
- Freund, J. E., & Perles, B. M. (2006). *Estadística elemental moderna* (12th ed.). Pearson.

4. Distribución Muestral de una Proporción

- Agresti, A., & Franklin, C. (2013). *Estadísticas: el arte y la ciencia de aprender de los datos* (3rd ed.). Pearson.
- Moore, D. S., McCabe, G. P., & Craig, B. A. (2012). *Introducción a la práctica de la estadística*. (8th ed.). W. H. Freeman.

Índice	1
Introducción	2
6.1 Muestreo	3
Principales técnicas de muestreo	3
¿Qué es el muestreo no probabilístico?	4
6.1.1 Tipos de muestreo.....	5
Muestreo no probabilístico.....	9
6.1.2 El teorema del límite central	10
6.1.3 Distribución Muestral de Medias	12
6.1.4 Distribución muestral de una proporción.	14
Conclusión	15
Referencias Bibliográficas	16

Introducción

En el ámbito de la estadística y la investigación, el muestreo y los conceptos relacionados es la recolección y análisis de datos. Estos conceptos permiten a los investigadores hacer inferencias sobre una población más amplia a partir de una muestra más pequeña y manejable: muestreo, tipos de muestreo, teorema de límite central, distribución muestral de la media y distribución muestral de una proporción.

Muestreo

El muestreo es el proceso mediante el cual se selecciona un subconjunto representativo de una población para realizar un estudio.

Tipos de Muestreo

Existen diversos métodos de muestreo que se pueden clasificar en dos grandes categorías: muestreo probabilístico y no probabilístico.

Teorema de Límite Central

El Teorema de Límite Central (TLC) es un pilar de la teoría estadística que establece que, bajo ciertas condiciones, la distribución de la suma (o promedio) de un gran número de variables aleatorias independientes e idénticamente distribuidas tiende a ser normal, independientemente de la distribución original de las variables.

Distribución Muestral de la Media

La distribución muestral de la media se refiere a la distribución de las medias de todas las posibles muestras de un tamaño específico extraídas de una población.

Distribución Muestral de una Proporción

Similar a la distribución muestral de la media, la distribución muestral de una proporción describe la distribución de las proporciones observadas en todas las posibles muestras de un tamaño específico extraídas de una población.

6.1 Muestreo

El muestreo estadístico es el proceso por el que se selecciona una muestra estadística que resulta representativa de la población y con la que se trabajarán los valores estimados de los parámetros a analizar.

Como la población o universo de estudio suele ser muy grande, los investigadores acotan su análisis a una muestra que resulte representativas de la población. Para ello, seleccionan un método de muestreo y realizan una investigación a través de la muestra que se ha seleccionado. A través de la inferencia estadística sobre la muestra, los investigadores obtienen información de la población que están estudiando.

El investigador, además de seleccionar la muestra, debe verificar los siguientes aspectos:

- El tamaño adecuado que debe tener la muestra y el nivel de confianza que desea establecer
- Los marcos muestrales que serán utilizados en la investigación
- Los estimadores que se van a utilizar
- La técnica de selección de los elementos de la investigación

Principales técnicas de muestreo

Hay diferentes técnicas de muestreo que puede usar el investigador para crear una estadística. Las más habituales:

- Muestreo Aleatorio Simple (MAS).
- Muestreo Aleatorio Estratificado (MAE).
- Muestreo Sistemático (MS).
- Muestreo por Conglomerados (MPC).

Todos estos que hemos mencionado son técnicas probabilísticas, aunque existen no probabilística, como el muestreo por conveniencia o por bola de nieve.

¿Qué es el muestreo no probabilístico?

El muestreo no probabilístico es una técnica que selecciona las personas que harán parte de la muestra de una manera subjetiva, esto quiere decir, según la decisión del investigador, evitando hacerlo al azar.

A diferencia en el muestreo probabilístico, donde cada individuo de la población tiene las mismas oportunidades de ser seleccionado, en el muestreo no probabilístico no todas las personas pueden participar.

El muestreo no probabilístico es más útil para estudios exploratorios como la encuesta piloto (una encuesta que se implementa en una muestra más pequeña, en comparación con el tamaño de muestra predeterminado). Se utiliza donde no es posible extraer un muestreo de probabilidad aleatorio debido a consideraciones de tiempo o costo.

También es un método menos estricto, este método de muestreo depende en gran medida de la experiencia de los investigadores. El muestreo no probabilístico comúnmente se lleva a cabo mediante métodos de observación, y se utiliza ampliamente en la investigación cualitativa.

6.1.1 Tipos de muestreo.

Los tipos de muestreo se dividen en dos categorías: probabilísticos y no probabilísticos.

- Muestreo aleatorio simple: este tipo de muestreo elige al azar cada individuo que hará parte de la muestra y todos tienen las mismas oportunidades de ser seleccionados.

Por ejemplo: si hay 400 universidades, eligen 20 aleatoriamente que cumplan ciertas características, tales como edificios con árboles, banderas, entrada de vidrio y grandes ventanas.



- Muestreo sistemático: este muestreo selecciona de forma aleatoria al primer individuo de la población que hará parte del estudio y luego, define un intervalo para completar la muestra.

Por ejemplo: en un grupo de 1000 personas desean obtener una muestra de 100 personas. El primer seleccionado es el número 2 (en azul). A partir de este y, mediante un intervalo de 4 personas decidido por el investigador, se seleccionarán los siguientes individuos. De esta manera, serán los números 6, 10, 14, 18, 22, etc., hasta completar la muestra de 100 individuos.



Este muestreo asegura una selección equitativa de la población, lo que resulta muy útil y rápido, además permite que la población esté correctamente representada.

- Muestreo por conglomerados: se utiliza cuando no se pueden estudiar todos los individuos de una población porque es muy grande o se encuentra dispersa en un área geográfica muy extensa, lo que aumenta el costo de la investigación.

Es útil cuando la población se encuentra agrupada previamente y de estos grupos se extraen los individuos para conformar la muestra.

Por ejemplo: imagina que necesitas realizar una investigación sobre la población de tu país. Como es un poco difícil llegar a cada habitante, aprovechas la división que ya existe por ciudades o por estados. Entonces, eliges 100 personas o 50 por cada zona de manera aleatoria para conformar la muestra.



- **Muestreo estratificado:** el muestreo estratificado divide la población en subgrupos o estratos que comparten ciertas características. Luego, utilizando el método del muestreo aleatorio simple, se eligen individuos de cada estrato para conformar la muestra.

Por ejemplo: una empresa necesita averiguar cuántas personas consumen jugo de naranja en Puerto Rico. La persona encargada de realizar el estudio decide utilizar el muestreo estratificado.

Entonces, divide la población en diferentes estratos, en este caso, por edades. De cada estrato toma un porcentaje de personas y con ellas conforma una muestra de 100.000 personas para el estudio del jugo de naranja.

¿Cuántas personas consumen jugo de naranja? 				
	Rango de edades de personas	Individuo en mm	%	Muestra
Estrato				
1	Menores de 18	2,5	41,67%	41.667
2	Entre 18 a 35	2	33,33%	33.333
3	Mayores de 35	1,5	25,00%	25.000
Total		6	100%	100.000

Total de muestra: 100.000

Muestreo no probabilístico

- Muestreo de bola de nieve: esta técnica ayuda a los investigadores a encontrar muestras cuando son difíciles de localizar. Es útil si el tamaño de la muestra es pequeño y no está disponible fácilmente.

Por ejemplo: si eres un investigador y estás buscando personas con cierta enfermedad, lo que haces es contactar a uno o dos individuos que presenten dicha condición médica y que ellos mismos sean quienes te recomienden a otros individuos con una historia médica parecida.

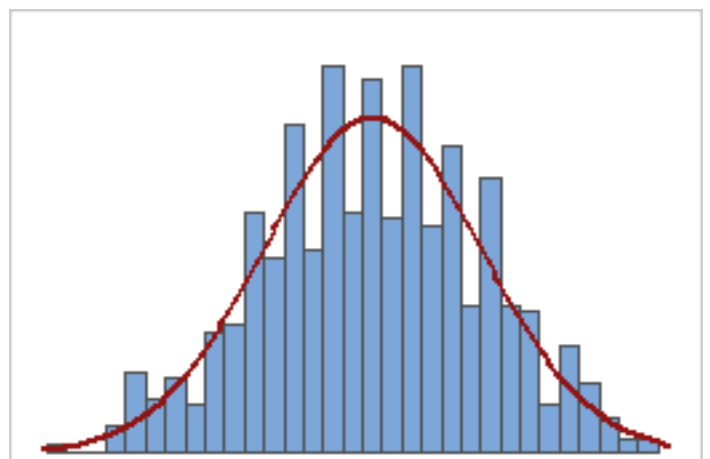
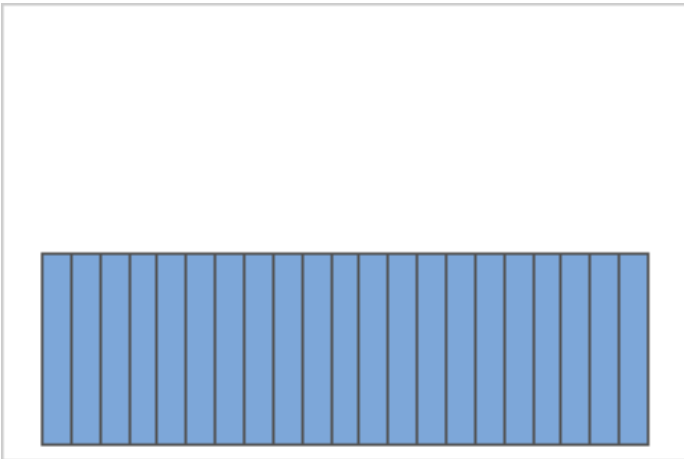
- Muestreo por cuotas: este método permite dividir la población en grupos o estratos que comparten características como el sexo, la edad, los estudios, etc., y de ellos se selecciona una muestra proporcional y representativa.

Por ejemplo: en una empresa grande quieren saber cuáles son los hobbies preferidos por sus empleados. Como son tantas personas, deciden dividirlos en dos estratos, por género: hombres y mujeres.

- Muestreo intencional o por conveniencia: con este método las muestras se seleccionan basándose únicamente en el conocimiento y la credibilidad del investigador. En otras palabras, los investigadores eligen solo aquellas personas que ellos creen que son los adecuados para participar en un estudio de investigación, ya sea porque son fáciles de reclutar o porque los consideran buenos representantes de la población.

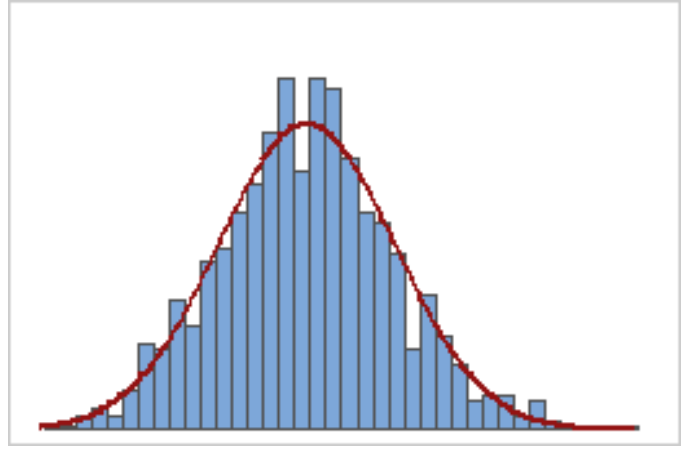
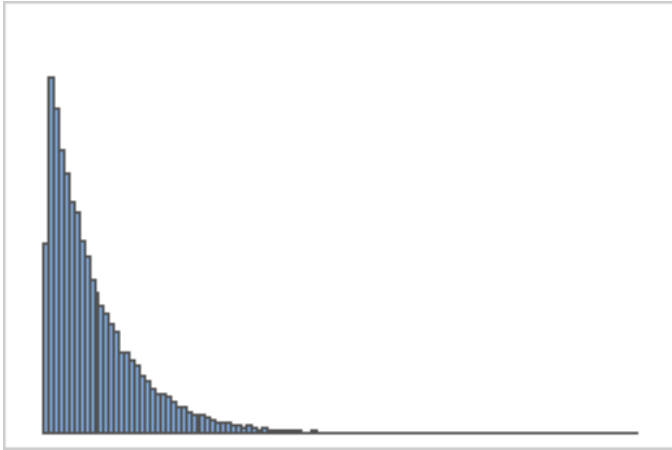
6.1.2 El teorema del límite central

El teorema del límite central es un teorema fundamental de probabilidad y estadística. El teorema describe la distribución de la media de una muestra aleatoria proveniente de una población con varianza finita. Cuando el tamaño de la muestra es lo suficientemente grande, la distribución de las medias sigue aproximadamente una distribución normal. El teorema se aplica independientemente de la forma de la distribución de la población. Muchos procedimientos estadísticos comunes requieren que los datos sean aproximadamente normales. El teorema de límite central le permite aplicar estos procedimientos útiles a poblaciones que son considerablemente no normales. El tamaño que debe tener la muestra depende de la forma de la distribución original. Si la distribución de la población es simétrica, un tamaño de muestra de 5 podría producir una aproximación adecuada. Si la distribución de la población es considerablemente asimétrica, es necesario un tamaño de muestra más grande. Por ejemplo, la distribución de la media puede ser aproximadamente normal si el tamaño de la muestra es mayor que 50. Las siguientes gráficas muestran ejemplos de cómo la distribución afecta el tamaño de la muestra que se necesita.



Muestras de una población uniforme

Una población que sigue una distribución uniforme es simétrica, pero marcadamente no normal, como lo demuestra el primer histograma. Sin embargo, la distribución de las medias de 1000 muestras de tamaño 5 de esta población es aproximadamente normal debido al teorema del límite central, como lo demuestra el segundo histograma. Este histograma de las medias de las muestras incluye una curva normal superpuesta para ilustrar esta normalidad.



Muestras de una población exponencial

Una población que sigue una distribución exponencial es asimétrica y no normal, como lo demuestra el primer histograma. Sin embargo, la distribución de las medias de 1000 muestras de tamaño 50 de esta población es aproximadamente normal debido al teorema del límite central, como lo demuestra el segundo histograma. Este histograma de las medias de las muestras incluye una curva normal superpuesta para ilustrar esta normalidad.

6.1.3 Distribución Muestral de Medias

Si recordamos a la distribución normal, esta es una distribución continua, en forma de campana en donde la media, la mediana y la moda tienen un mismo valor y es simétrica. Con esta distribución podíamos calcular la probabilidad de algún evento relacionado con la variable aleatoria, mediante la siguiente fórmula:

$$z = \frac{x - \mu}{\sigma}$$

En donde z es una variable estandarizada con media igual a cero y varianza igual a uno. Con esta fórmula se pueden hacer los cálculos de probabilidad para cualquier ejercicio, utilizando la tabla de la distribución z .

Sabemos que cuando se extraen muestras de tamaño mayor a 30 o bien de cualquier tamaño de una población normal, la distribución muestral de medias tiene un comportamiento aproximadamente normal, por lo que se puede utilizar la fórmula de la distribución normal con $\mu = \mu_{\bar{x}}$ y $\sigma = \bar{\sigma}$, entonces la fórmula para calcular la probabilidad del comportamiento del estadístico, en este caso la media de la muestra, quedaría de la siguiente manera:

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

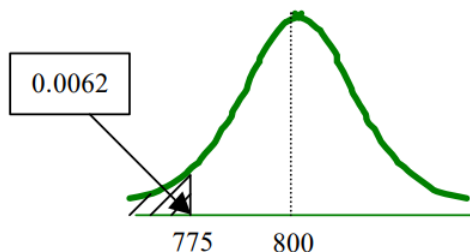
y para poblaciones finitas y muestro con reemplazo:

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n} \sqrt{\frac{N-n}{N-1}}}$$

Ejemplo:

Una empresa eléctrica fabrica focos que tienen una duración que se distribuye aproximadamente en forma normal, con media de 800 horas y desviación estándar de 40 horas. Encuentre la probabilidad de que una muestra aleatoria de 16 focos tenga una vida promedio de menos de 775 horas.

Solución:



$$z = \frac{775 - 800}{40 / \sqrt{16}} = -2.5$$

Este valor se busca en la tabla de z
 $P(\bar{x} \leq 775) = P(z \leq -2.5) = 0.0062$

La interpretación sería que la probabilidad de que la media de la muestra de 16 focos sea menor a 775 horas es de 0.0062.

Ejemplo:

Las estaturas de 1000 estudiantes están distribuidas aproximadamente en forma normal con una media de 174.5 centímetros y una desviación estándar de 6.9 centímetros. Si se extraen 200 muestras aleatorias de tamaño 25 sin reemplazo de esta población, determine:

- El número de las medias muestrales que caen entre 172.5 y 175.8 centímetros.
- El número de medias muestrales que caen por debajo de 172 centímetros.

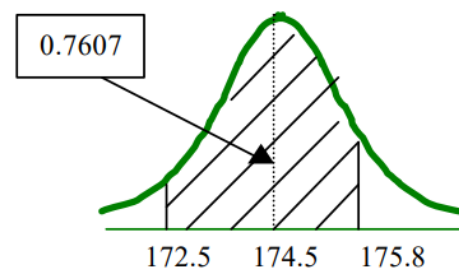
Solución:

Como se puede observar en este ejercicio se cuenta con una población finita y un muestreo sin reemplazo, por lo que se tendrá que agregar el factor de corrección. Se procederá a calcular el denominador de Z para sólo sustituir o en cada inciso.

$$\sigma / \sqrt{n} \sqrt{\frac{N-n}{N-1}} = 6.9 / \sqrt{25} \sqrt{\frac{1000-25}{1000-1}} = 1.36$$

a)

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n} \sqrt{\frac{N-n}{N-1}}} = \frac{172.5 - 174.5}{1.36} = -1.47$$



$$z = \frac{175.8 - 174.5}{1.36} = 0.96$$

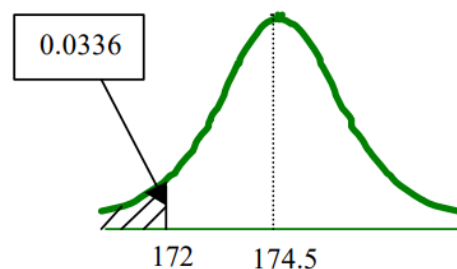
$$p(172.5 \leq \bar{x} \leq 175.8) = 0.7607$$

$$(0.7607)(200) = 152 \text{ medias muestrales}$$

b) $z = \frac{172 - 174.5}{1.36} = -1.83$

$$p(\bar{x} \leq 172) = 0.0336$$

$$(0.0336)(200) = 7 \text{ medias muestrales}$$



6.1.4 Distribución muestral de una proporción.

Cuando en una población procedemos a estudiar una característica con sólo dos posibles valores (éxito/fracaso), entonces la población sigue una distribución binomial.

Cada muestra de la población tiene un porcentaje de individuos que tiene esta característica. p es la proporción de éxito de esta variable aleatoria de la población.

La proporción de fracaso es $q = 1 - p$

Sean todas las muestras de tamaño n de la población. Cada muestra tiene una proporción de individuos con esa característica.

La distribución asociada a la variable aleatoria que une cada muestra con su proporción se llama distribución muestral de proporciones. Como, para poblaciones grandes, la binomial se aproxima a la normal, la distribución muestral de proporciones también sigue una distribución normal:

$$N\left(\sqrt{\frac{pq}{n}}\right)$$

sí n es suficientemente grande, $n \geq 30$, and $np \geq 5$, $nq \geq 5$

Como generalmente las proporciones de la población son desconocidas, las aproximamos por las de la muestra.

Ejemplo: Una máquina fabrica piezas de precisión. En su producción habitual fabrica un 3% de piezas defectuosas. Un cliente recibe una caja de 500 piezas procedentes de la fábrica.

a) ¿Cuál es la probabilidad de que encuentre más del 5% de piezas defectuosas en la caja?

b) ¿Cuál es la probabilidad de que encuentre menos de un 1% de piezas defectuosas?

$$\hat{p} = p = 0,03 \quad \sigma_{\hat{p}} = \sqrt{\frac{0,03 \cdot 0,97}{500}} = 0,0076 \quad N(0,03; 0,0076)$$

$$\begin{aligned} a) P(\hat{p} > 0,05) &= 1 - P(\hat{p} \leq 0,05) = 1 - P\left(\frac{\hat{p} - 0,03}{0,0076} \leq \frac{0,05 - 0,03}{0,0076}\right) \\ &= 1 - P(Z \leq 2,63) = 0,0043 \end{aligned}$$

$$b) P(\hat{p} > 0,01) = P\left(\frac{\hat{p} - 0,03}{0,0076} \leq \frac{0,01 - 0,03}{0,0076}\right) = P(Z \leq -2,63) = 0,0043$$

Conclusión

Los conceptos clave de muestreo y estadística proporcionan su importancia en la investigación y la toma de decisiones basada en datos. El muestreo, con sus diversos métodos, permite realizar estudios más manejables y económicos sin sacrificar la validez de los resultados. El Teorema de Límite Central, por su parte, se erige como una piedra angular en la teoría estadística, ofreciendo una base robusta para el análisis de datos, incluso en situaciones donde las distribuciones subyacentes no son normales.

La distribución muestral de la media y la distribución muestral de una proporción son herramientas esenciales para la inferencia estadística, permitiendo a los investigadores evaluar la variabilidad y la incertidumbre inherentes a las estimaciones derivadas de muestras. Estas distribuciones no solo facilitan la construcción de intervalos de confianza, sino que también son cruciales para la realización de pruebas de hipótesis, lo que fortalece la capacidad de los investigadores para hacer afirmaciones basadas en evidencia.

En mi opinión, una comprensión y aplicación de estos conceptos es indispensable para cualquier investigador o profesional que trabaje con datos. La capacidad de diseñar estudios de muestreo adecuados y aplicar correctamente las técnicas de inferencia estadística no solo mejora la calidad de la investigación, sino que también contribuye a la toma de decisiones informadas en una amplia gama de disciplinas, desde las ciencias sociales hasta las ciencias naturales y la ingeniería. Por lo tanto, invertir tiempo y esfuerzo en dominar estos principios es una tarea fundamental para aquellos que buscan avanzar en el campo de la investigación y la estadística aplicada.

Referencias Bibliográficas

6.1 Muestreo

Economía Simple. (s.f.). Muestreo Estadístico. Recuperado de <https://economiasimple.net/glosario/muestreo-estadistico>

6.1.1 Tipos de muestreo

GCF Global. (s.f.). Tipos de Muestreo. Recuperado de <https://edu.gcfglobal.org/es/estadistica-basica/tipos-de-muestreo/1/>

6.1.2 El teorema del límite central

Minitab. (s.f.). About the Central Limit Theorem. Recuperado de <https://support.minitab.com/es-mx/minitab/help-and-how-to/statistics/basic-statistics/supporting-topics/data-concepts/about-the-central-limit-theorem/>

6.1.3 Distribución Muestral de Medias

Hopelchen Tecnológico Nacional de México. (s.f.). Distribución Muestral de Medias. Recuperado de <https://hopelchen.tecnm.mx/principal/sylabus/fpdb/recursos/r129750.PDF>

6.1.4 Distribución muestral de una proporción

Martínez, V. (s.f.). Distribución Muestral de Proporciones. Recuperado de https://www.victormat.es/mcs2/Tema12-DistribucionesMuestrales/distribucion_muestral_de_proporciones.html