

Tabla de contenido

1. Objetivo de la unidad 1.....	pág 2
2. Tema 1: Conceptos fundamentales.....	pág 2
2.1 Conceptos Fundamentales.....	pág 3
2.2 Autoevaluación.....	pág 7
3. Bibliografía.....	pág 8

1. Objetivo de la unidad 1

Apropiar conceptos básicos de la Estadística Descriptiva, para realizar análisis estadístico más sofisticados que facilite la toma de decisiones basadas en datos.

2. Tema 1: Conceptos fundamentales

¿Qué es la estadística?

La estadística es la disciplina encargada de recopilar, organizar, analizar, interpretar y presentar datos numéricos para describir las características de la población o fenómeno que se estudia. El objetivo principal es sacar conclusiones válidas a partir de los datos y tomar decisiones informadas.

Estadística Descriptiva

Se refiere al proceso de resumir y describir datos utilizando medidas como la media, la mediana, la moda y la desviación estándar. Las estadísticas descriptivas proporcionan una visión general de las características básicas de los datos, tales como: su distribución, tendencia central y dispersión.

Estadística Inferencial

Se ocupa de hacer inferencias o generalizaciones sobre una población basándose en una muestra de datos. Esto implica estimar parámetros poblacionales utilizando métodos probabilísticos, realizar pruebas de hipótesis, construir intervalos de confianza y tomar decisiones basadas en la incertidumbre asociada con los datos.

La estadística en la Ciencia de Datos

La estadística proporciona métodos y herramientas para analizar datos, sacar conclusiones válidas, modelar relaciones entre

variables, validar modelos y comunicar resultados de manera efectiva.

Desde la recopilación inicial de datos hasta la toma de decisiones informada, la estadística guía todo el proceso de análisis de datos y permiten que los datos se transformen en información significativa y útil en una variedad de campos y aplicaciones.

2.1 Conceptos Fundamentales

Población

La población se refiere al conjunto completo de elementos o individuos que comparten una característica específica y son objeto de estudio. Esta característica puede variar según el contexto de la investigación. La población puede ser finita o infinita, dependiendo de si se puede enumerar o no todos los elementos.

Ejemplo 1:

En un estudio sobre el rendimiento académico de todos los estudiantes de una universidad, la **población** sería el conjunto completo de estudiantes matriculados en esa universidad.

Ejemplo 2:

En un estudio sobre el nivel de satisfacción entre los clientes de un restaurante para mejorar sus servicios, la **población** sería todos los clientes que han visitado el restaurante en el último mes.

Ejemplo 3:

En un estudio de una empresa de servicios públicos, sobre el consumo de energía en una comuna ubicada en la ciudad de Santiago de Cali, la **población** sería todos los hogares ubicados en la comuna.

Muestra

La muestra es un subconjunto representativo de la población que se selecciona para realizar observaciones o mediciones. La elección de la muestra debe realizarse de manera que refleje de manera precisa las características de la población de interés. Se utiliza cuando es impráctico o costoso analizar todos los elementos de la población.

Ejemplo 1:

Un hospital desea conocer la satisfacción de sus pacientes con respecto a los servicios médicos recibidos. La Población son todos los pacientes que fueron atendidos en el hospital durante el último año. Una **muestra** serían 300 pacientes seleccionados al azar de la lista de todos los pacientes atendidos durante el último año.

Ejemplo 2:

Una empresa de alimentos quiere analizar las preferencias alimenticias de los habitantes de la ciudad de Santiago de Cali. La Población son todos los habitantes de la ciudad. Una **muestra** serían 500 habitantes seleccionados al azar de la lista de todos los habitantes de la ciudad.

Ejemplo 3:

Un emprendedor lanza un nuevo producto al mercado y desea conocer la opinión de los consumidores sobre el mismo. La Población son todos los consumidores que han adquirido el nuevo producto. Una **muestra** sería 200 consumidores seleccionados al azar de la lista de todos los consumidores que han adquirido el nuevo producto.

Variables

Una variable es una característica o propiedad que puede medirse, cuantificarse o clasificarse en diferentes categorías. Estas categorías pueden representar diferentes valores o niveles de la variable. Las variables son esenciales para recopilar datos y realizar análisis estadísticos porque permiten explorar relaciones y patrones que existen dentro de un conjunto de datos.

Ejemplo 1:

Se realiza una encuesta para conocer las preferencias de género cinematográfico entre los estudiantes de una escuela. Las variables podrían ser:

1. Género de la película (por ejemplo, Acción, Comedia, Drama).
2. Sexo del encuestado (Masculino, Femenino, otro).
3. Número de películas vistas en el último mes.

4. Puntuación de satisfacción con la última película vista en una escala del 1 al 10.

Ejemplo 2:

Se desea investigar los hábitos de consumo de café entre los empleados de una empresa. Las variables podrían ser:

1. Tipo de café preferido (Espresso, Latte, Americano).
2. Endulzante utilizado (Azúcar, Edulcorante, Sin endulzante).
3. Número de tazas de café consumidas por día.
4. Precio máximo que están dispuestos a pagar por una taza de café.

Ejemplo 3:

Un restaurante quiere evaluar la satisfacción de sus clientes con el fin de tomar decisiones informadas. Las variables podrían ser:

1. Satisfacción con la comida (Muy satisfecho, Satisfecho, Neutral, Insatisfecho, Muy insatisfecho).
2. Edad del cliente (18-25, 26-35, 36-45, 46 o más).
3. Puntuación de satisfacción general en una escala del 1 al 10.
4. Monto gastado en la última visita al restaurante.

Variables Cualitativas

También conocidas como variables categóricas, representan características o atributos que no se pueden medir numéricamente. Se dividen en dos subtipos: Nominales u Ordinales.

Variables cualitativas nominales: Son aquellas que representan categorías sin un orden inherente. Ejemplos de variables nominales incluyen el género, el estado civil, el tipo de sangre, el color favorito, etc.

Variables cualitativas ordinales: A diferencia de las nominales, estas variables tienen un orden natural o jerarquía entre las categorías. Ejemplos de variables ordinales incluyen la clasificación de productos (bueno, regular, malo), niveles de satisfacción (alto, medio, bajo), entre otros.

Variables Cuantitativas

Son aquellas que se pueden medir y expresar numéricamente. Se dividen en dos subtipos: Continuas o discretas.

Variables cuantitativas continuas: Representan medidas que pueden tomar cualquier valor dentro de un rango determinado. Ejemplos incluyen la altura, el peso, la temperatura, el tiempo de reacción, etc.

Variables cuantitativas discretas: Representan conteos o recuentos de objetos y tienen valores que son números enteros no negativos. Ejemplos incluyen el número de hijos en una familia, el número de clientes que visitan una tienda en un día, el número de veces que se lanza una moneda hasta obtener cara, etc.

Ejemplo 1:

Se realiza una encuesta para conocer los hábitos de ejercicio de los residentes de un barrio de Cali.

Variables Cualitativas:

1. Tipo de ejercicio preferido (Correr, Nadar, Yoga).
2. Nivel de frecuencia de ejercicio (Nunca, Ocasionalmente, Regularmente, Diariamente).

Variables Cuantitativas:

1. Duración promedio de la sesión de ejercicio en minutos.
2. Número de veces que se ejercita por semana.

Ejemplo 2:

Se desea investigar los hábitos de uso de redes sociales entre los estudiantes de una universidad.

Variables Cualitativas:

1. Red social favorita (Facebook, Instagram, Twitter).
2. Frecuencia de uso (Nunca, Ocasionalmente, Diariamente).

Variables Cuantitativas:

1. Número de horas promedio de uso diario de redes sociales.
2. Cantidad de publicaciones realizadas por día.

Ejemplo 3:

Se realiza una encuesta para conocer las preferencias de transporte de los residentes en Bogotá.

Variables Cualitativas:

1. Medio de transporte preferido (Carro, Bicicleta, Transmilenio).
2. Motivo de elección del medio de transporte (Comodidad, Costo, Rapidez).

Variables Cuantitativas:

1. Distancia promedio del trayecto diario en kilómetros.
2. Tiempo promedio de desplazamiento diario en minutos.

2.2 Autoevaluación

Pregunta 1:

La muestra incluye a todas las unidades individuales de interés, mientras que la población es un subconjunto de la muestra.

Opción 1: Verdadero

Opción 2: Falso

Pregunta 2:

Las Variables en estadística son características o atributos que pueden tomar diferentes valores y que se pueden medir o categorizar.

Opción 1: Verdadero

Opción 2: Falso

Pregunta 3:

¿Cuál de las siguientes opciones describe correctamente los tipos de variables en estadística?

Opción 1: Variables cuantitativas y cualitativas

Opción 2: Variables numéricas y variables no numéricas

Opción 3: Variables continuas y variables discretas

Opción 4: Todas las anteriores

3. Bibliografía

- ✓ Romero Ramos, E. (2016). Estadística para todos: análisis de datos: estadística descriptiva, teoría de la probabilidad e inferencia. Difusora Larousse - Ediciones Pirámide, (pp. 22-30). <https://elibro-net.bibliotecavirtual.unad.edu.co/es/ereader/unad/49136?page=1>
- ✓ Romero Ramos, E. (2016). Estadística para todos: análisis de datos: estadística descriptiva, teoría de la probabilidad e inferencia. Difusora Larousse - Ediciones Pirámide, (pp. 31-39). <https://elibro-net.bibliotecavirtual.unad.edu.co/es/ereader/unad/49136?page=1>
- ✓ Romero Ramos, E. (2016). Estadística para todos: análisis de datos: estadística descriptiva, teoría de la probabilidad e inferencia. Difusora Larousse - Ediciones Pirámide, (pp. 41-55). <https://elibro-net.bibliotecavirtual.unad.edu.co/es/ereader/unad/49136?page=1>
- ✓ Romero Ramos, E. (2016). Estadística para todos: análisis de datos: estadística descriptiva, teoría de la probabilidad e inferencia. Difusora Larousse - Ediciones Pirámide, (pp. 57-69). <https://elibro-net.bibliotecavirtual.unad.edu.co/es/ereader/unad/49136?page=1>
- ✓ Romero Ramos, E. (2016). Estadística para todos: análisis de datos: estadística descriptiva, teoría de la probabilidad e inferencia. Difusora Larousse - Ediciones Pirámide, (pp. 93-100). <https://elibro-net.bibliotecavirtual.unad.edu.co/es/ereader/unad/49136?page=1>
- ✓ Cely, L. A. (2020). Tipos de muestreo y error muestral. [Objeto_virtual_de_Informacion_OVI]. Repositorio Institucional UNAD. <https://repository.unad.edu.co/handle/10596/38474>

- ✓ Ortegon, M. F. & Cabrera, F. (2018). Tablas de Frecuencia. [Objeto_virtual_de_Informacion_OVI]. Repositorio Institucional UNAD. <https://repository.unad.edu.co/handle/10596/20458>
- ✓ Ortegon, M. F. (2019). Medidas de Tendencia Central. [Objeto_virtual_de_aprendizaje_OVA]. Repositorio Institucional UNAD. <https://repository.unad.edu.co/handle/10596/33842>
- ✓ Camargo, I. (2022). *Medidas estadísticas univariantes: Muestreo*. [Objeto_virtual_de_Informacion_OVI]. Repositorio Institucional UNAD. <https://repository.unad.edu.co/handle/10596/59205>
- ✓ Camargo, I. (2022). Medidas univariantes para variables cuantitativas. [Objeto_virtual_de_aprendizaje_OVA]. Repositorio Institucional UNAD. <https://repository.unad.edu.co/handle/10596/50133>
- ✓ Camargo, I. (2023). Muestreo y caracterización de variables cualitativas. [Objeto_virtual_de_aprendizaje_OVA]. Repositorio Institucional UNAD. <https://repository.unad.edu.co/handle/10596/59206>