

La estadística

Tipos y ciclo de vida del dato



tech*h*



CONTENIDO

1. Objetivos

2. Conceptos básicos

Estadística
Estadística descriptiva
Estadística inferencial

3. Población

Tipos de población estadística

4. Muestra

Tipos de muestra estadística

5. Bibliografía



OBJETIVOS

- Conocer los conceptos básicos de estadística y sus tipos.
- Establecer que se conoce como población estadística.
- Presentar que es una muestra estadística y cuáles son sus tipos.
- Conocer las escalas de medidas aplicadas a los estudios estadísticos.

CONCEPTOS BÁSICOS

ESTADÍSTICA

La estadística es una ciencia que se ocupa de la obtención, ordenamiento y análisis de datos para conocer determinados fenómenos y eventos con la aplicación de diversos métodos que permiten describir los hechos a partir de su observación científica, así como realizar deducciones a partir de dicha observación para anticipar su comportamiento futuro con niveles razonables de confianza. Es la técnica matemática que obtiene, organiza, presenta y describe un conjunto de datos con el propósito de facilitar el uso, generalmente con el apoyo de tablas, medidas numéricas o gráficas.

La estadística se divide tradicionalmente en dos ramas:

ESTADÍSTICA DESCRIPTIVA

También llamada deductiva, se refiere a 'describir' y analizar algunas características de los elementos de un grupo dado con el fin de explicar dicho grupo, sin extraer conclusiones para un grupo mayor. No hace uso del cálculo de probabilidades.

La estadística descriptiva proporciona resúmenes sencillos sobre la muestra y sobre las observaciones que se han hecho. Esos resúmenes pueden ser cuantitativos, es decir, estadísticas resumidas, o visuales, es decir, gráficos de fácil comprensión. Dichos resúmenes también pueden constituir la base de la descripción inicial de los datos como parte de un análisis estadístico más amplio, o pueden ser suficientes por sí mismos para una investigación determinada.

El uso de la estadística descriptiva tiene una larga historia y, de hecho, el simple cálculo de las poblaciones y de los datos económicos fue la primera forma en que apareció el tema de las estadísticas. Hoy en día en el campo de la ciencia de datos, se ha formulado una colección de técnicas de resumen bajo el título de análisis exploratorio.

ESTADÍSTICA INFERENCIAL

Es aquella que permite extraer conocimiento de una información existente, y partir de ello sacar conclusiones generales para toda la población con un nivel de confianza establecido y unos márgenes de error calculados y conocidos. Sus principales herramientas son el muestreo, la estimación de parámetros y el contraste de hipótesis. Se desarrolla a partir de las muestras representativas, que permitan proyectar la información y comportamiento presentado por estas a una población de mayor tamaño, así mismo, permite estimar los valores aproximados de los parámetros estudiados para dicha población, al tiempo que aportará la información necesaria y suficiente para hacer estimaciones del comportamiento futuro de ciertos fenómenos o eventos con márgenes de error conocidos y admisibles.

Algunos de los métodos aplicados en inferencia estadística y que permiten llegar a conclusiones y deducciones son:

- Pruebas de estimación puntual.
- Pruebas de hipótesis.
- Pruebas paramétricas.
- Pruebas no paramétricas.
- Análisis de correlación.
- Regresión.
- Análisis de varianza, por mencionar algunos de los más utilizados.

Algunos ejemplos de la aplicación de la estadística inferencial son:

Uno de los ejemplos clásicos de estadística inferencial, son los sondeos de intención de voto electoral. Se suelen hacer con relativa frecuencia y permiten establecer con proximidad el comportamiento que en materia de resultado electoral se puede dar a partir de una muestra representativa, es importante destacar que para la fiabilidad de estos estudios es necesario estudiar y definir la muestra con base a criterios de selección científicos bien establecidos.

Los estudios de mercado también son un buen ejemplo de la utilidad de la estadística inferencial.

En el área de salud, su uso es de gran valor y aplicabilidad tanto para la morbilidad como en la epidemiología y en general para hacer proyecciones que permitan conocer la frecuencia con la que se presentarán los eventos en un horizonte temporal establecido y poder por tanto construir las condiciones que permiten a los sistemas sanitarios su abordaje, establecer necesidades de recursos de toda índole para su atención oportuna.

Como cualquier disciplina, la estadística tiene una terminología que debe conocerse, en este caso viene heredada de sus orígenes, que fueron trabajos de tipo demográfico (tablas de mortalidad y censos):

- **Estadística en los negocios:** la estadística es fundamental para gestionar y mejorar temas o actividades relacionados con los negocios, por ejemplo:
 - Los niveles de satisfacción de los clientes y usuarios.
 - Los tipos de accidentes y sus frecuencias.
 - Ventas por clientes, vendedores, zonas y productos.
 - Predicciones de ventas por zonas, productos, servicios o sucursales.
 - Proyectos de inversión.
 - Evolución de las distintas ratios económicas-financieras y patrimoniales a lo largo del tiempo.
 - Estudios e investigación de mercado.
 - Productos más demandados, en el ámbito global, por zona y por canal de comercialización.
 - Estadística del personal (directivo y empleado).
 - El control de calidad.

POBLACIÓN

Se entiende por población estadística a el conjunto de elementos o sujetos que poseen una o más características comunes, que los convierte en potenciales objetos de estudio, por lo que permitirán sacar conclusiones específicas para determinar resultados.

TIPOS DE POBLACIÓN ESTADÍSTICA

Existen varios tipos de población estadística:

- **Población finita:** aquella población que está compuesta por un número pequeño de elementos, artículos o sujetos. Ejemplo: número de unidades vendidas de un determinado producto, número de solicitantes de un servicio en un tiempo determinado.
- **Población infinita:** aquella población que está compuesta por un número grande de elementos, no conocido. Ejemplo, la cantidad de metros cúbicos de arenales.
- **Población real:** Aquella población que se compone de un grupo de cosas existentes y numerables.
- **Población hipotética:** aquella población conformada por un grupo posible de determinadas cosas, pero difícil de comprobar.
- **Población aleatoria:** aquella población cuya aparición se considera accidental o azarosa. Ejemplo: el peso de un grupo de piezas de fruta
- **Población dependiente:** aquella población que depende de una causa conocida.

MUESTRA

La muestra estadística consiste en la porción que se extrae de una población de una dimensión mayor para un determinado estudio, con el fin de conocer, representar y determinar las características o atributos de dicha población.

Este método se utiliza cuando no se puede aplicar un estudio determinado en la totalidad de la población, ya sea por sus dimensiones y/o características, a través del muestreo se puede establecer una porción de la realidad a estudiar. Al establecer una muestra se deberá establecer con claridad la metodología usada para el establecimiento de esta, con la finalidad de permitir a los consumidores de este estudio conocer claramente los términos en los que dicha muestra fue seleccionada.

TIPOS DE MUESTRA ESTADÍSTICA

La muestra estadística puede ser tanto probabilística como no probabilística.

- **Muestra probabilística:** el muestreo probabilístico es el tipo más utilizado durante las investigaciones. En este, todos los elementos de la población o universo tienen la posibilidad de ser parte de la muestra. No existe una única fórmula para establecer una muestra probabilística, se presenta a continuación una sencilla formula, para comprender el razonamiento matemático de la misma:

$$n_i = \left(\frac{N_i}{N} \right) \cdot n$$

- **n_i :** tamaño de la muestra del estrato i (número de individuos que se seleccionarán de ese subgrupo)
- **N_i :** tamaño total de la población del estrato i (número total de individuos en ese subgrupo)
- **N :** tamaño total de la población (suma de todos los estratos).
- **n :** tamaño total de la muestra que se desea obtener.

Dentro de la muestra probabilística pueden encontrarse los siguientes tipos de muestreo:

- **Muestreo aleatorio simple:** método de selección más básico, en el que cada sujeto posee un número de identificación y, mediante un sorteo aleatorio, algunos son seleccionados para la muestra. Para realizarla se debe conocer la cantidad de individuos que serán necesarios para completar toda la muestra.
- **Muestreo sistemático:** con este método se enumera la población a trabajar y los investigadores se encargan de listar a cada individuo en grupos de 10, de forma aleatoria. Luego se elige a uno de los primeros de cada grupo al azar, formando de esta manera la muestra.
- **Muestreo estratificado:** consiste en dividir a la población en estratos o grupos que compartan características similares. Luego se seleccionan proporcionalmente algunos individuos de cada grupo.

- **Muestreo por conglomerados:** se da cuando la población ya se encuentra dividida de forma natural en grupos, por lo que se seleccionan individuos aleatoriamente de cada conglomerado para conformar la muestra.

Pasos para llevar a cabo un muestreo probabilístico:

- Selección de la población de interés.
- Determina un marco o contexto de interés para llevar a cabo la selección.
- Selección de muestra, con base a los criterios establecidos.
- Recolecta los datos de interés de cada uno de los sujetos seleccionados.

Utilidad del muestreo probabilístico:

- Cuando se pretende reducir el sesgo.
- Genera mayor nivel de confianza con relación a los datos recolectados, aporta calidad y credibilidad en los hallazgos del estudio.
- En poblaciones de grandes dimensiones la selección al azar de los individuos de estudio suele reflejar con bastante fidelidad la composición de la población general.
- Se puede diseñar con bastante exactitud la dimensión que debe tener la muestra para emular en comportamiento de la población.

Muestra no probabilística: el muestreo no probabilístico es aquel en el que la selección es arbitraria por parte del investigador o recolector de datos, seleccionando de forma directa y caprichosa el orden de los sujetos de estudio. Este método de determinación de la muestra no brinda a todos los sujetos de la población la misma oportunidad de participación.

Este tipo de muestra se puede encontrar en los siguientes tipos:

- **Muestreo por cuotas:** los investigadores forman la muestra partiendo de determinadas características, con el fin de lograr la misma distribución de características en la muestra que en la población. Por ejemplo, si en una población hay 70 % hombres y 30 % mujeres, los individuos de la muestra deberán estar distribuidos de la misma manera.
- **Muestreo por conveniencia:** aquí el investigador suele elegir a los individuos de su muestra solo por su proximidad. Generalmente, el investigador no reconoce esta muestra como una representación de toda una población, sino que la realiza para conocer opiniones, datos e información de manera rápida.
- **Muestreo por bola de nieve:** se utiliza cuando el investigador requiere que un sujeto de su muestra ayude a identificar a otro con las mismas características y, a su vez, estos a otros, para formar la muestra deseada.

- **Muestra discrecional:** también conocida como muestreo por juicio o criterio, se da cuando el investigador selecciona los individuos de su muestra en base a determinado conocimiento de la población. Por ejemplo, si desea realizar una investigación de personas con problemas psicológicos en el embarazo, seleccionará aquellas mujeres que hayan pasado por la experiencia del embarazo.

Utilidad de muestreo no probabilístico

- Se utiliza para establecer características relevantes a estudiar en las etapas de preparación de un estudio determinado, estudios piloto.
- En estudios con limitaciones importantes de recursos físicos y de tiempo.

Individuo: son las personas, elementos o sujetos de estudio que poseen la información que se pretende estudiar.

Escalas de medida: de esta manera, la estadística contempla cuatro niveles de medición de datos, conocidos como escalas de medición estadística, que son:

- **Medición nominal:** este nivel de medición establece categorías distintivas que no implican un orden específico, en la escala nominal se asignan números a eventos para identificarlos. Ejemplo: categoría sexo con dos niveles, masculino (M) y femenino (F).
- **Medición ordinal:** establece categorías con dos o más niveles que poseen entre sí. Se puede usar una escala numérica tipo *liker*, por lo que será cuantitativa. También se pueden usar categóricas, siempre que la cualidad descrita mantenga un orden conocido. Ejemplo: nivel de estudio.
- **Medición de intervalo:** la medición de intervalo posee las características de la medición nominal y ordinal. Establece la distancia entre una medida y otra, estableciendo intervalos. La escala de intervalo se aplica a variables continuas, pero carece de un punto cero absolutos. Ejemplo: un ejemplo clásico de ello es la temperatura, esta puede contener datos menores a negativos.
- **Medición de razón:** esta incluye las características de los tres anteriores niveles de medición (nominal, ordinal e intervalo). Determina la distancia exacta entre los intervalos de una categoría. El cero representa la ausencia de la cualidad. Ejemplo: número de hijos, nivel de ingresos.

BIBLIOGRAFÍA

- [1] E. R. González. *Estadística general*. Universidad Central de Venezuela, Ediciones de la Biblioteca; 2014.
- [2] Equipo editorial, Etecé. Estadística inferencial [En línea]. Concepto.de. 2018. Disponible en: <https://concepto.de/estadistica-inferencial/> [Accedido el 14 de octubre de 2021].
- [3] J. Gorgas, N. Cardiel, J. Zamorano. *Estadística básica para estudiantes de ciencias*. Departamento de Astrofísica y Ciencias de la Atmosfera, Facultad de Ciencias Físicas, Universidad Complutense de Madrid; 2011.
- [4] Muestreo no probabilístico: definición, tipos y ejemplos [En línea]. Questionpro.com. 2018. Disponible en: <https://www.questionpro.com/blog/es/muestreo-no-probabilistico/> [Accedido el 14 de octubre de 2021].
- [5] Muestreo probabilístico: Qué es y cómo utilizarlo en tu investigación [En línea]. Questionpro.com. 2018. Disponible en: <https://www.questionpro.com/blog/es/como-realizar-un-muestreo-probabilistico/> [Accedido el 14 de octubre de 2021].
- [6] Estadística descriptiva - Estadística para Sociología [En línea]. Google.com. Disponible en: <https://sites.google.com/site/estadisticaparasociologia/tarea> [Accedido el 14 de octubre de 2021].