

Principios generales de la estadística

Presentación de datos

Los datos estadísticos se presentan generalmente expresando el valor de la frecuencia absoluta que toman las variables significativas de un estudio, ya correspondan a una población o a una muestra.

La frecuencia absoluta de un valor o de una modalidad de una variable estadística es el número de datos observados que presentan ese valor o modalidad.

El cociente entre la frecuencia absoluta de un valor o modalidad y el número total de datos es llamado frecuencia relativa.

También suelen presentarse los datos en forma de porcentaje (es decir, en forma de razón de denominador 100).

Estadística descriptiva

La estadística descriptiva es una disciplina que se encarga de recoger, almacenar, ordenar, realizar tablas o gráficos y calcular parámetros básicos sobre el conjunto de datos.

Tipos

Distribución de frecuencias

Es básicamente una presentación o un resumen de datos agrupados que se han clasificado en función de clases mutuamente excluyentes y del número de ocurrencias en cada clase respectiva.

Permite una forma más estructurada y organizada de presentar los datos en bruto.

Tendencia central

Se refiere al resumen descriptivo de un conjunto de datos utilizando un único valor que refleja el centro de la distribución de los datos.

La media

Que se considera la medida de tendencia central más popular, es el valor medio o más común en un conjunto de datos

La mediana

Se refiere a la puntuación media de un conjunto de datos en orden ascendente.

La moda

Se refiere a la puntuación o valor más frecuente en un conjunto de datos.

Variabilidad

Refleja el grado de dispersión de una muestra. Las medidas de variabilidad determinan la distancia que los puntos de datos parecen tener con respecto al centro.

Desviación estándar

Se utiliza para determinar la varianza media de un conjunto de datos

Proporciona una idea de la distancia o la diferencia entre un valor de un conjunto de datos y el valor medio del mismo conjunto de datos.

La varianza

Refleja el grado de dispersión y es esencialmente una media de las desviaciones al cuadrado.

Tipos de gráficas

Gráfico o diagrama de barras

Suele expresar mediante la elevación de barras de diferente color aquella información que intenta dilucidar un sólo aspecto entre un grupo de personas encuestadas.

Pueden ser horizontales.

Gráfico o diagrama de Sectores

Este es el tipo de gráfico ideal para representar porcentajes en una situación similar a la anterior.

Histograma

Representación gráfica de una variable determinada a través de barras, en las cuales su superficie es proporcional a la frecuencia de los valores hallados.

El eje vertical marca las frecuencias, y el horizontal los valores posibles de las variables.

Polígono de frecuencias

Tipo de gráfico lineal que utilizamos para la representación de la incidencia de respuesta de una variable cuantitativa.

El polígono surge de unir los puntos medios de las bases superiores de las barras de un diagrama de barras, e incluso también de un histograma

Pictograma

se sustituyen los elementos abstractos (como las barras) por dibujos relativos a la temática de lo que se está graficando

Su tamaño debe ser proporcional a la frecuencia que representen; para una mayor claridad se sugiere indicarla.

Diagramas de cajas

¿Qué es?

Los diagramas de caja muestran la distribución de datos para una variable continua.

¿Cómo se utilizan los diagramas de caja?

Los diagramas de caja ayudan a ver el centro y la extensión de los datos. También se pueden utilizar como herramienta visual para comprobar normalidad o identificar puntos que podrían ser valores atípicos.

¿Cómo se crean?

Calcular la mediana, el percentil 25 y el percentil 75.

Calcular el rango intercuartílico (IQR) como la diferencia entre el percentil 75 y el 25.

Calcular la longitud máxima de las patillas multiplicando el IQR por 1,5.

Identificar los valores atípicos.

Usar las estadísticas calculadas para representar los resultados y trazar un diagrama de caja.