



NOMBRE: CARLOS EDUARDO NAJERA LEON

TEMA: MAPA CONCEPTUAL

PROFESOR: ANDRES ALEJANDRO REYES MOLINA

CARRERA: ADMINISTRACION DE EMPRESAS Y
ESTRATEGIAS DE NEGOCIOS

ACTIVIDAD 1

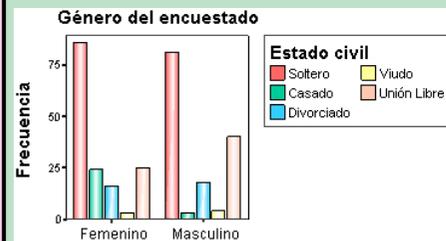
DISTRIBUCION DE FRECUENCIAS

Es una forma organizada de presentar datos, mostrando cuántas veces se repite cada valor o categoría dentro de un conjunto de datos. Es una tabla o gráfica que clasifica los datos en diferentes grupos (o clases) y registra la frecuencia con la que cada grupo aparece.

TIPOS DE FRECUENCIA

- **Frecuencia absoluta:** El número de veces que aparece un determinado valor o categoría en un estudio estadístico.
- **Frecuencia relativa:** La frecuencia absoluta dividida por el número total de datos, a menudo expresada como porcentaje.
- **Frecuencia acumulada:** La suma de las frecuencias absolutas hasta un cierto valor.
- **Frecuencia relativa acumulada:** La suma de las frecuencias relativas hasta un cierto valor.

REPRESENTACION GRAFICA DE UN TIPO DE FRECUENCIA



PROPIEDADES DE LAS DISTRIBUCIONES DE FRECUENCIAS

- **Organización:** Las distribuciones de frecuencias agrupan datos en categorías o intervalos, lo que facilita la comprensión de la distribución general.
- **Frecuencia:** Indican la cantidad de veces que cada categoría o intervalo aparece en los datos.
- **Clases o intervalos:** Los datos se dividen en clases o intervalos para facilitar la agrupación y la visualización.
- **Frecuencia absoluta:** Representa el número de observaciones que pertenecen a cada categoría o intervalo.
- **Frecuencia relativa:** Indica la proporción o porcentaje de observaciones que pertenecen a cada categoría o intervalo.
- **Frecuencia acumulada:** Representa la suma de las frecuencias absolutas hasta una categoría o intervalo determinado.
- **Frecuencia relativa acumulada:** Indica la proporción o porcentaje acumulado de observaciones hasta una categoría o intervalo determinado.
- **Interpretación visual:** Las distribuciones de frecuencias pueden representarse gráficamente como histogramas, polígonos de frecuencia, etc., lo que facilita la interpretación visual de la distribución.
- **Calculo de medidas estadísticas:** Permiten calcular medidas descriptivas como la media, mediana, moda, percentiles, etc., y visualizar la forma de la distribución.

ESTADÍSTICOS DE POSICIÓN GRUPAL

Cuartiles:

Dividen los datos en cuatro partes iguales.

Q1 (primer cuartil): Representa el valor por debajo del cual se encuentra el 25% de los datos.

Q2 (segundo cuartil): Representa la mediana (el valor central) y el 50% de los datos están por debajo de él.

Q3 (tercer cuartil): Representa el valor por debajo del cual se encuentra el 75% de los datos.

Ejemplo: Si las calificaciones de un examen son 5, 6, 6, 7, 7, 7, 8, 8, 8, 8, 9, 9, 10, 10, 10, entonces Q1 es 7, Q2 es 8 y Q3 es 9.

Deciles:

Dividen los datos en diez partes iguales.

D1, D2, ..., D9: Representan el 10%, 20%, ..., 90% de los datos.

Ejemplo: Si se tienen 20 estudiantes y se ordenan sus notas de menor a mayor, el decil 5 (D5) representa la nota de la posición 10 ($20 + 1/10 * 5$).

Percentiles:

Dividen los datos en 100 partes iguales.

P1, P2, ..., P99: Representan el 1%, 2%, ..., 99% de los datos.

Ejemplo: El percentil 75 (P75) representa el valor por debajo del cual se encuentra el 75% de los datos.

ACTIVIDAD 1

DISTRIBUCION DE FRECUENCIAS

La moda en variables categóricas representa la categoría o valor que ocurre con mayor frecuencia en un conjunto de datos. Es una medida de tendencia central que se usa para identificar la categoría más común.

- Variables categóricas: Son variables que representan categorías o grupos distintos, como colores, tipos de productos, o respuestas a preguntas de opción múltiple
- Moda: Es el valor o categoría que aparece con mayor frecuencia en un conjunto de datos.
- Ejemplo: Si se realiza una encuesta y se pregunta por el color favorito, la moda sería el color que más personas eligen.
- Cálculo: Para encontrar la moda en datos categóricos, simplemente se cuenta la frecuencia de cada categoría y se identifica la categoría con la frecuencia más alta.
- Variables nominales: La moda es la única medida de tendencia central útil para variables nominales, que son variables categóricas sin orden inherente.
- Variables ordinales:

VARIABLES ORDINALES

El mínimo y el máximo:

El mínimo es el valor más bajo observado en la variable ordinal.

El máximo es el valor más alto observado en la variable ordinal.

La mediana:

La mediana es el valor que divide la distribución ordenada en dos partes iguales, es decir, el 50% de los valores están por debajo y el 50% están por encima.

Para calcular la mediana, primero se ordenan los datos de menor a mayor. Si hay un número impar de datos, la mediana es el valor central. Si hay un número par de datos, la mediana es la media entre los dos valores centrales.

Los cuartiles:

Los cuartiles son valores que dividen la distribución en cuatro partes iguales.

El primer cuartil (Q1) es el valor que deja el 25% de los datos por debajo.

El segundo cuartil (Q2) es la mediana.

El tercer cuartil (Q3) es el valor que deja el 75% de los datos por encima.

Los cuartiles, junto con el mínimo y el máximo, proporcionan un rango intercuartílico (IQR), que es la diferencia entre Q3 y Q1.

Otros cuantiles:

Además de los cuartiles, pueden definirse otros cuantiles como los percentiles (que dividen la distribución en 100 partes iguales) o los deciles (que dividen la distribución en 10 partes iguales).

VARIABLES CUANTITATIVAS

La media se calcula sumando todos los valores de un conjunto de datos y dividiendo por el número total de valores.

Es muy útil para describir el centro de un conjunto de datos cuando no hay outliers significativos.

Sin embargo, los outliers pueden elevar o reducir la media de manera considerable, distorsionando la representación de la tendencia central.

Alternativas Robustas a la Media:

La Mediana:

Es el valor central en un conjunto de datos ordenados. Es mucho más robusta a los outliers que la media, ya que no se ve afectada por valores extremos.

La Media Recortada:

Se calcula excluyendo un porcentaje de los valores más bajos y más altos del conjunto de datos, y luego se calcula la media de los valores restantes.

La Media Winsorizada:

Se reemplazan los valores atípicos más extremos por valores cercanos a los valores no atípicos del conjunto, y luego se calcula la media.

ESTADÍSTICAS DE DISPERSION

En estadística, las medidas de dispersión es el grado en que una distribución se estira o se comprime. Ejemplos comunes de medidas de dispersión estadística son la varianza, la desviación estándar y el rango intercuartil.

Medidas de dispersión

Rango

$$R = L_m - L_i$$

Desviación media

$$D_{\bar{x}} = \frac{\sum_{i=1}^N |x_i - \bar{x}| f_i}{N}$$

Coefficiente de variación

$$CV = \frac{\sigma}{\bar{x}}$$