



**Mi Universidad**

**NOMBRE: CARLOS EDUARDO NAJERA LEON**

**TEMA: SUPER NOTA**

**PROFESOR: ANDRES ALEJANDRO REYES MOLINA**

**CARRERA: ADMINISTRACION DE EMPRESAS Y  
ESTRATEGIAS DE NEGOCIOS**

**SEMESTRE: 2do CUATRIMESTRE**

## APUNTAMIENTO

## ASIMETRIA

La asimetría, en términos generales, se refiere a la falta de simetría, es decir, la ausencia de correspondencia o equilibrio entre las partes de un todo. Este concepto puede aplicarse en diversos contextos, como en la forma de objetos, en la distribución de datos, en la estructura social, entre otros.

La definición es sencilla, el apuntamiento es la diferencia entre la media ARITMÉTICA del mercado OMIE, y la media PONDERADA que tenga cada consumidor. Así, si un consumidor, tiene más demanda en horas más caras, su apuntamiento será positivo, porque su media ponderada es más alta que la media aritmética del mes.

## ESTADÍSTICOS DE POSICIÓN INDIVIDUAL

### Primer Cuartil

El primer cuartil (Q1) es el valor que divide a un conjunto de datos ordenados en orden creciente en cuatro partes iguales, siendo el valor por debajo del cual se encuentra el 25% de los datos. En otras palabras, el primer cuartil es el percentil 25, lo que significa que el 25% de los datos son menores o iguales a ese valor.

Interpretación:

El primer cuartil indica el punto en el que el 25% de los datos son menores o iguales al valor del primer cuartil. Por ejemplo, si el primer cuartil de un conjunto de edades es 20 años, significa que el 25% de las personas son menores o iguales a 20 años.

### Segundo Cuartil

El segundo cuartil, también conocido como Q2 o mediana, es el valor que divide un conjunto de datos ordenados en dos partes iguales. En otras palabras, el 50% de los datos se encuentra por debajo de la mediana y el otro 50% por encima.

### Tercer Cuartil

El tercer cuartil (Q3), también conocido como cuartil superior, es el valor que divide a una muestra de datos ordenada en cuatro partes iguales, donde el 75% de los datos se encuentra por debajo de él. En otras palabras, el tercer cuartil es el percentil 75.

## Los porcentajes acumulados

Los porcentajes acumulados representan la suma progresiva de porcentajes de un conjunto de datos. En lugar de mostrar el porcentaje de cada valor individual, muestran el porcentaje total hasta un punto específico, acumulándose a medida que se consideran más valores. Este concepto es útil para analizar distribuciones y tendencias en datos, como en gráficos de Pareto o al calcular el porcentaje completado de una tarea con el tiempo.

Cálculo y ejemplos:

- Acumulación:

Para obtener un porcentaje acumulado, se suman los porcentajes individuales de cada valor hasta un punto determinado.

- Gráficos de Pareto

En un gráfico de Pareto, la línea acumulada muestra la suma de los porcentajes de las categorías hasta cada punto.

- Ejemplo

Si en una encuesta, el 30% de los encuestados están de acuerdo con la opción A y el 25% con la opción B, el porcentaje acumulado hasta la opción B sería del 55% (30% + 25%).

## Las puntuaciones típicas

Las puntuaciones típicas, también conocidas como puntuaciones z, son una forma de expresar la posición de una puntuación individual dentro de un conjunto de datos en relación con la media y la desviación estándar de ese conjunto. Indican cuántas desviaciones estándar se encuentra una puntuación por encima o por debajo de la media.

¿Qué son las puntuaciones típicas?

Puntuación directa:

Es la puntuación original obtenida por un individuo en una prueba o medición.

Media:

Es el valor promedio de todas las puntuaciones en un conjunto de datos.

Desviación estándar:

Mide la dispersión o variabilidad de las puntuaciones alrededor de la media.

Puntuación típica (z):

Se calcula restando la media de la puntuación directa y luego dividiendo el resultado por la desviación estándar. La fórmula es:  $z = (X - \mu) / \sigma$ , donde X es la puntuación directa,  $\mu$  es la media y  $\sigma$  es la desviación estándar.

## Las escalas derivadas

En estadística, las escalas derivadas son formas de representar puntuaciones que no son las originales (puntuaciones directas) en una nueva escala con características específicas. Estas escalas derivadas se utilizan para facilitar la interpretación de los resultados, especialmente en pruebas psicológicas y educativas.

## Medidas de Posición

Las medidas de posición, también llamadas medidas de tendencia central, son estadísticas que se usan para resumir y representar los valores centrales de un conjunto de datos. Estas medidas proporcionan información sobre el valor típico o central alrededor del cual se agrupan los datos.

Las medidas de posición más comunes son:

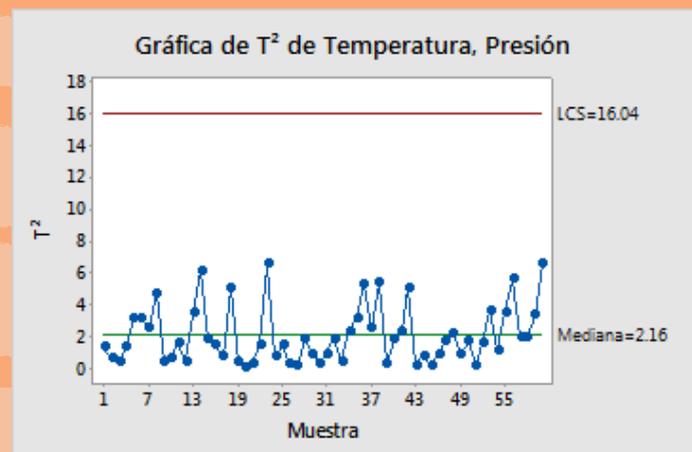
Media: El promedio de todos los valores en el conjunto de datos.

Mediana: El valor que se encuentra en el centro de los datos ordenados, dividiendo la distribución en dos mitades iguales.

Moda: El valor que aparece con mayor frecuencia en el conjunto de datos.

## Organización y representación gráfica de datos multivariados

La organización y representación gráfica de datos multivariados implica visualizar relaciones entre múltiples variables simultáneamente para identificar patrones, tendencias y estructuras. Existen diversas técnicas gráficas para este fin, como diagramas de dispersión, gráficos de coordenadas paralelas, diagramas de estrellas, entre otros, que ayudan a explorar las interacciones complejas entre las variables.



## La distribución conjunta multivariada

La distribución conjunta multivariada describe la probabilidad de múltiples variables aleatorias, mostrando cómo están relacionadas entre sí. En otras palabras, especifica la probabilidad de que varias variables tomen ciertos valores simultáneamente. Este concepto es fundamental en estadística y probabilidad, especialmente cuando se analizan conjuntos de datos con múltiples variables interrelacionadas.

## La tabla de contingencia

Una tabla de contingencia, también llamada tabla de doble entrada o tabla cruzada, es una herramienta estadística que se utiliza para resumir y analizar datos categóricos. Permite visualizar la relación entre dos variables categóricas y calcular probabilidades condicionales.

¿Qué muestra una tabla de contingencia?

- Filas y columnas:

La tabla se organiza en filas y columnas, donde cada fila representa una categoría de una variable y cada columna representa una categoría de la otra variable.

- Frecuencias:

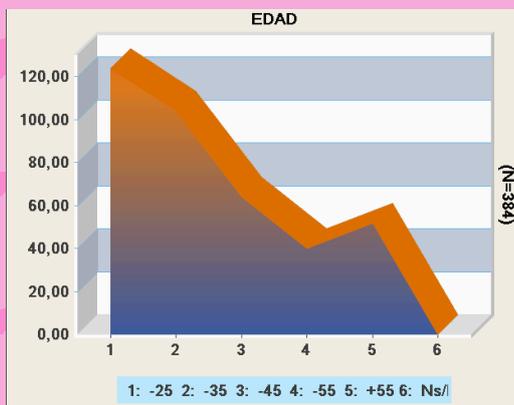
Cada celda de la tabla muestra la frecuencia (o número) de observaciones que pertenecen a la combinación específica de categorías de las dos variables.

- Totales marginales:

Los totales de cada fila y columna (llamados totales marginales) indican la frecuencia total de cada categoría de una variable, independientemente de la otra.

- Gran total:

La suma de todos los valores de la tabla, o el total general, representa el número total de observaciones.



- Variables aleatorias:

Son variables cuyos valores son resultados de un fenómeno aleatorio.

- Distribución conjunta:

Describe la probabilidad de que múltiples variables aleatorias tomen valores específicos simultáneamente.

- Multivariada:

Se refiere a la situación donde hay más de dos variables aleatorias involucradas.

- Distribución de probabilidad:

Describe cómo se distribuyen las probabilidades de los posibles resultados de una variable aleatoria.

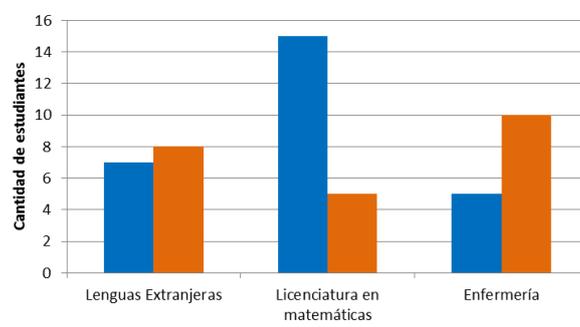
## Calculo para una distribución de Frecuencia

Para crear una tabla de distribución de frecuencias, primero se deben organizar los datos, luego calcular la frecuencia absoluta (el número de veces que aparece cada valor), la frecuencia relativa (la frecuencia absoluta dividida por el número total de datos) y opcionalmente la frecuencia acumulada y el porcentaje.

## Representación Grafica



## Diagrama de barras para las variables inclinación profesional y género



# El caso de dos variables categóricas

El análisis de dos variables categóricas implica examinar la relación entre dos variables que solo pueden tomar un número limitado de categorías, como "sí" o "no", o diferentes tipos de colores. Se utilizan principalmente tablas de contingencia, gráficos de barras agrupadas o apiladas para visualizar la relación. La prueba chi-cuadrado es una herramienta estadística común para determinar si existe una asociación significativa entre las dos variables.

El análisis de dos variables categóricas implica examinar la relación entre dos variables que solo pueden tomar un número limitado de categorías, como "sí" o "no", o diferentes tipos de colores. Se utilizan principalmente tablas de contingencia, gráficos de barras agrupadas o apiladas para visualizar la relación. La prueba chi-cuadrado es una herramienta estadística común para determinar si existe una asociación significativa entre las dos variables.

Conceptos Clave:

- Variables Categóricas:
  - Variables que representan categorías o grupos, en lugar de valores numéricos continuos.
  -
- Tabla de Contingencia:
  - Una tabla que muestra la distribución de frecuencias de dos o más variables categóricas, indicando cuántas observaciones caen en cada combinación de categorías.
  -
- Gráfico de Barras Agrupadas:
  - Un gráfico que utiliza barras para mostrar la frecuencia de cada categoría de una variable, agrupadas por las categorías de la segunda variable.
  -
- Gráfico de Barras Apiladas:
  - Un gráfico que muestra la frecuencia de cada categoría de una variable, con barras apiladas para representar las diferentes categorías de la segunda variable.

# El caso de dos variables cuantitativas

En el caso de dos variables cuantitativas, se busca analizar la relación entre ellas. Esto puede implicar determinar si existe una correlación, ya sea positiva o negativa, y la fuerza de dicha relación. Técnicas como el análisis de correlación y la regresión lineal son útiles para este propósito, permitiendo visualizar la asociación y predecir valores de una variable basados en la otra.

Análisis de la relación entre dos variables cuantitativas:

- Correlación:
  - Se utiliza para medir la fuerza y dirección de la relación lineal entre dos variables cuantitativas. El coeficiente de correlación de Pearson es una herramienta común para este análisis, variando entre -1 y +1. Un valor cercano a +1 indica una correlación positiva fuerte, un valor cercano a -1 indica una correlación negativa fuerte, y un valor cercano a 0 indica una correlación débil o inexistente.
  -
- Regresión lineal:
  - Permite modelar la relación entre las variables y predecir el valor de una variable (dependiente) en función de la otra (independiente). La ecuación de la recta de regresión,  $y = a + bx$ , describe la relación lineal entre las variables, donde "a" es la intersección con el eje y y "b" es la pendiente.

# El caso de una variable categórica y una variable cuantitativa

En estadística, una variable categórica (también llamada cualitativa) representa grupos o categorías, mientras que una variable cuantitativa representa cantidades numéricas. Al analizar datos que involucran ambas, se busca entender la relación entre las categorías y las cantidades asociadas a ellas.

Ejemplos y usos:

- Comparación de medias:
  - Se puede comparar la estatura promedio (variable cuantitativa) entre hombres y mujeres (variable categórica).
- Análisis de varianza (ANOVA):
  - Se utiliza para comparar las medias de una variable cuantitativa entre tres o más grupos definidos por una variable categórica.
- Gráficos de cajas y bigotes:
  - Permiten visualizar la distribución de una variable cuantitativa para cada categoría de una variable categórica, mostrando la mediana, cuartiles y valores atípicos.
- Tablas de contingencia:
  - Estas tablas muestran la distribución conjunta de dos variables categóricas, lo que permite analizar la relación entre ellas.
- Regresión:
  - Se pueden usar modelos de regresión para predecir una variable cuantitativa basada en una variable categórica o en una combinación de variables categóricas y cuantitativas.