



UDS

UNIDAD 3.-
ESTADISTICA



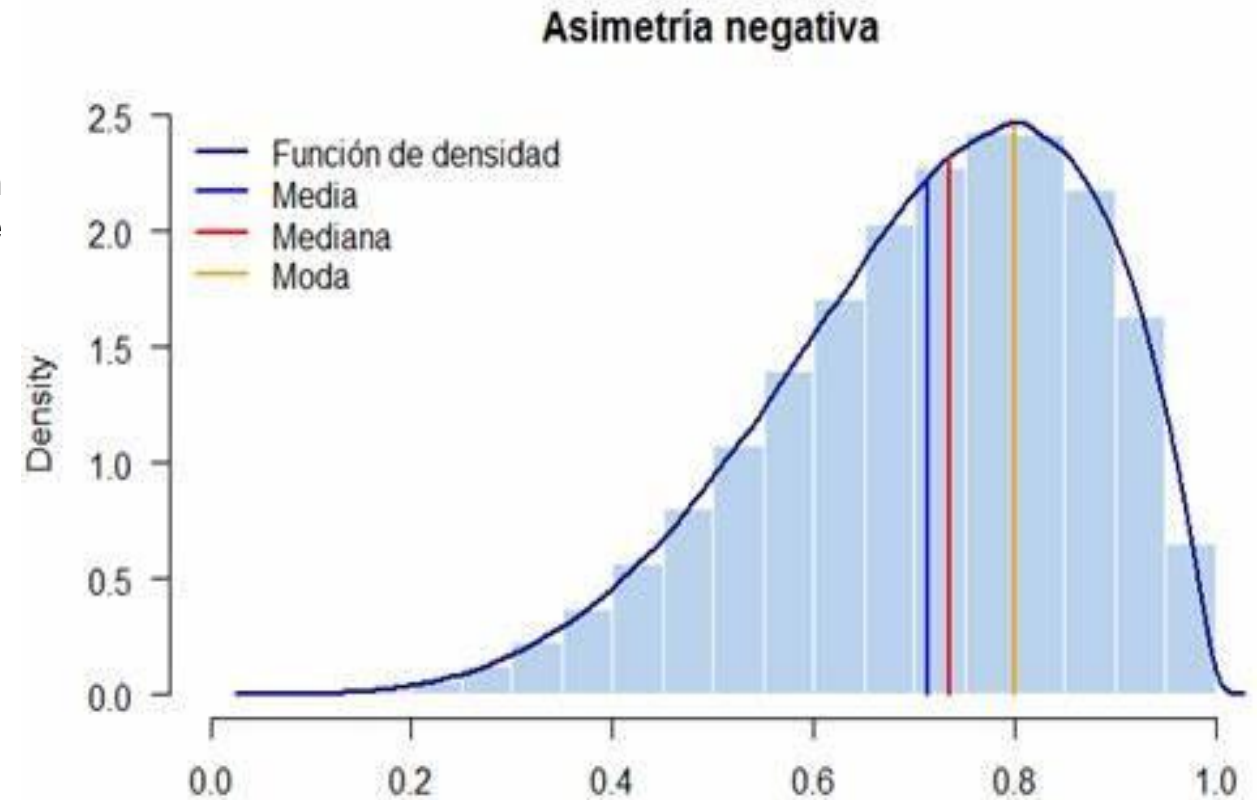
**ALEJANDRO
GUZMAN
ARROYO**

Unidad 3: Estadísticos de Forma de la Distribucion

3.1.-Asimetría

La asimetría es una propiedad de determinados cuerpos, dibujos, funciones matemáticas y otros tipos de elementos en los que, al aplicarles una regla de transformación efectiva, se observan cambios respecto al elemento original.

Los indicadores gracias a los cuales es posible identificar el grado de simetría o de asimetría presente en una distribución de probabilidad (una función que asigna a una serie de eventos determinados la probabilidad de que ocurran) de una variable aleatoria (también conocida con el nombre de variable estocástica, es una función que asigna sucesos a números reales) sin necesidad de representarla gráficamente se denominan medidas de asimetría.



Apuntamiento

La curtosis es una medida estadística que determina el grado de concentración que presentan los valores de una variable alrededor de la zona central de la distribución de frecuencias. También es conocida como medida de apuntamiento.

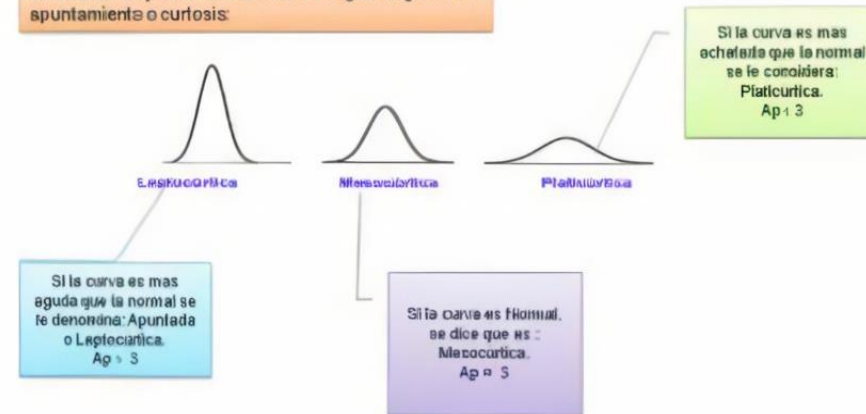
Según su concepción clásica, una curtosis grande implica una mayor concentración de valores de la variable tanto muy cerca de la media de la distribución (pico) como muy lejos de ella (colas), al tiempo que existe una relativamente menor frecuencia de valores intermedios. Esto explica una forma de la distribución de frecuencias/probabilidad con colas más gruesas, con un centro más apuntado y una menor proporción de valores intermedios entre el pico y colas.

Una mayor curtosis no implica una mayor varianza, ni viceversa.

APUNTAMIENTO O CURTOSIS (A_p)

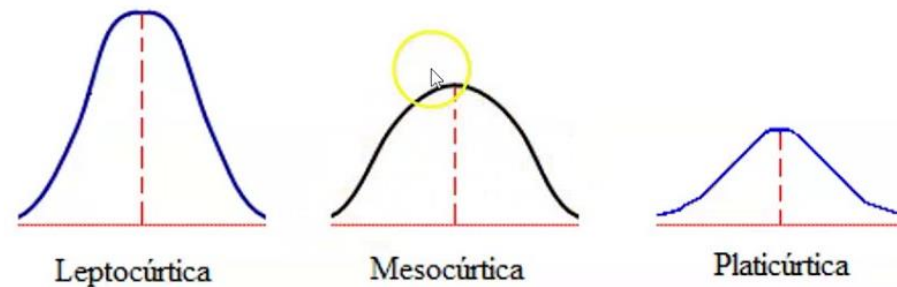
Mide la mayor o menor cantidad de datos que se agrupan en torno a la moda.

Se definen 3 tipos de distribuciones según su grado de apuntamiento o curtosis:



Medidas de apuntamiento (Curtosis)

Representa la medida de concentración de datos alrededor de la media.



3.3.- Estadísticos de Posición Individual

3.1.1.- Porcentajes Acumulados



El porcentaje acumulado (%a) de un valor concreto de una variable es el porcentaje de casos que obtienen un valor inferior o igual a ese en la variable en cuestión, información que puede obtenerse directamente a partir de la distribución de frecuencias correspondiente a esa variable.

- A estos porcentajes acumulados se les suele denominar más habitualmente, aunque de un modo equívoco, percentiles (término ya utilizado en el contexto de los estadísticos de posición grupal). En referencia al porcentaje acumulado de una determinada puntuación es la de rango centil, si bien, su uso no está muy extendido.

3.1.2.- Puntuaciones Típicas

Una alternativa al problema planteado en la utilización de las puntuaciones diferenciales son las puntuaciones típicas (estándar o z), una transformación de las puntuaciones directas que tiene en cuenta tanto la tendencia central (media) como la dispersión (desviación típica) de la distribución de frecuencias de la variable.

La puntuación z correspondiente a un determinado valor expresa el número de desviaciones típicas que ese valor dista de la media del conjunto de las observaciones.

11.1. Puntuaciones típicas o normalizadas

Para poder comparar dos datos correspondientes a dos distribuciones distintas, hay que tipificar —o normalizar— dichos valores, es decir, calcular los valores $z = \frac{x - \bar{x}}{\sigma}$, después, comparar los resultados.

Las puntuaciones típicas o normalizadas, también llamadas puntuaciones z, tienen las siguientes propiedades:

- Si se transforma una distribución en puntuaciones típicas, no varía la forma de la distribución original.
- La media aritmética de las puntuaciones normalizadas es cero, es decir, $\bar{z} = 0$
- La desviación típica de las puntuaciones típicas es la unidad, es decir, $\sigma_z = 1$

3.1.3.- Las Escalas Derivadas



Una dificultad con las puntuaciones típicas se puede plantear a la hora de comunicar resultados debido a las posiciones decimales y valores negativos inherentes a las mismas. Es por ello que se han propuesto algunas transformaciones lineales de las puntuaciones típicas que pretenden hacerlas más intuitivamente interpretables.

- Todas estas escalas derivadas de la escala de las puntuaciones típicas se basan en una transformación genérica del tipo:

$$D_i = a \cdot z_i + b,$$

3.4.- Organización Y Representación Grafica de Datos Multivariados

3.4.1.- La Distribucion Conjunta Multivariada

En el caso de tener más variables, se denominaría multivariada. El nombre largo de distribución conjunta es distribución de probabilidad conjunta. El nombre se abrevia dado que ya se conoce que esas distribuciones son de probabilidad.

Una distribución conjunta es la distribución de probabilidad de la intersección de las realizaciones de dos o más variables aleatorias cualesquiera. En otras palabras, una distribución conjunta es la distribución de probabilidad que forman dos o más variables aleatorias cuando sus realizaciones se producen simultáneamente.



3.4.2.- La Tabla de Contingencia

Una tabla de contingencia es una herramienta utilizada en la rama de la estadística, la cual consiste en crear al menos dos filas y dos columnas para representar datos categóricos en términos de conteos de frecuencia.

Esta herramienta, que también se conoce como tabla cruzada o como tabla de dos vías, tiene el objetivo de representar en un resumen, la relación entre diferentes variables categóricas.

La tabla de contingencia es una de las herramientas analíticas más útiles y un pilar de la industria de la investigación de mercados.

La tabla permite medir la interacción entre dos variables para conocer una serie de información “oculta” de gran utilidad para comprender con mayor claridad los resultados de una investigación.



TABLAS DE CONTINGENCIA

Estadística bidimensional y Probabilidad

		Itinerario (y)		
		Ciencias	Letras	
Sexo (x)	Chicas	13	6	19
	Chicos	9	12	21
		22	18	40



Ventajas de realizar una tabla de contingencia:

Entre los principales beneficios de realizar una tabla de contingencia se encuentran los siguientes:

Facilita la lectura de los datos recolectados, ya que permite agruparlos cuando aún se encuentran sin procesar, lo que disminuye el margen de error al realizar un informe de investigación.

Gracias a la tabla de contingencia es posible realizar gráficas que permitan visualizar la información fácilmente para su comprensión.

A diferencia de otros métodos estadísticos de análisis de datos, la tabla de contingencia permite ahorrar tiempo durante la correlación de variables.

Las tablas ofrecen resultados claros y precisos que permiten tomar mejores decisiones y crear estrategias basadas en datos.

Organización de datos

Tablas de contingencia

Tabla 1. Estado civil y sexo de los encuestados

Estado Civil \ Sexo	Sexo		Total
	Masculino	Femenino	
Soltero	2	1	3
Casado	2	2	4
Divorciado	2	1	3
Total	6	4	10

Fuente: Datos imaginarios



3.4.3.- Representación Graficas

3.4.3.1.- Dos Variables Categóricas

Las variables categóricas también se denominan variables cualitativas o variables de atributos. Los valores de una variable categórica son categorías o grupos mutuamente excluyentes. Los datos categóricos pueden tener o no tener un orden lógico.

Se puede hacer estadística con las variables categóricas, a pesar de no ser variables numérica o cuantitativas. Por ejemplo, para conocer la tendencia o valor más probable de una variable categórica se toma la moda. La moda es, en este caso, el resultado o valor más repetido de una variable categórica

Gráficos para variables cualitativas o categóricas

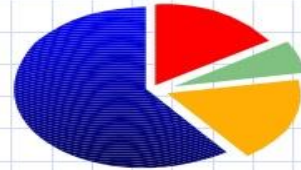


Diagrama circular



Gráfico de barras

Estas son algunas de las representaciones más utilizadas para variables categóricas.

LAS VARIABLES CATEGÓRICAS Ejemplo:

Al conjunto constituido por un grupo de individuos que comparten una característica ("estudiantes de enfermería") le asignamos el símbolo "1", o "I", o "A1"; también al conjunto constituido por individuos varones le asignamos "1" y al de los individuos pertenecientes al sexo femenino, "0".

En todo caso, la idea es evitar asignar el mismo número a dos conjuntos.

3.4.3.2.- Dos Variables Cuantitativas

Lo cualitativo se centra en la calidad o cualidad de las cosas, mientras que lo cuantitativo se refiere a la cantidad o a aquello que es susceptible de ser representado numéricamente.

Las variables cuantitativas son aquellas variables estadísticas que otorgan, como resultado, un valor numérico.





Tipos de variables

VARIABLES CUALITATIVAS ó CATEGÓRICAS

Expresa una cualidad, característica o atributo que solo se pueden clasificar o categorizar mediante el conteo.

Dependiendo del número de categorías pueden ser dicotómicas o politómicas

Ejemplos:

Sexo, Estado Civil, Dolor, Grupo Sanguíneo, Métodos anticonceptivos, Profesionales de la Salud.

VARIABLES CUANTITATIVAS ó NUMÉRICAS

Son variables que se expresan numéricamente. Son susceptibles de medirse

Ejemplos:

Número de hijos, Número de camas, Peso, Talla, Edad, Glucosa, Colesterol, etc.