



Ethan Alberto Mendoza Ruiz.

Ing. ANDRES ALEJANDRO REYES MOLINA

Estadística descriptiva

Tercer cuatrimestre.

Administración y gestión de negocios– A.

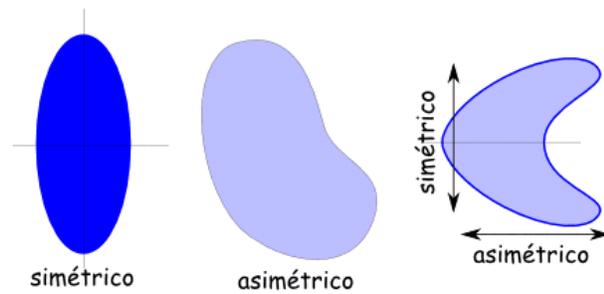
16 de junio de 2025.

## ASIMETRIA

Las medidas de asimetría son indicadores que permiten establecer el grado de simetría (o

asimetría) que presenta una distribución de probabilidad de una variable aleatoria sin tener que

hacer su representación gráfica.

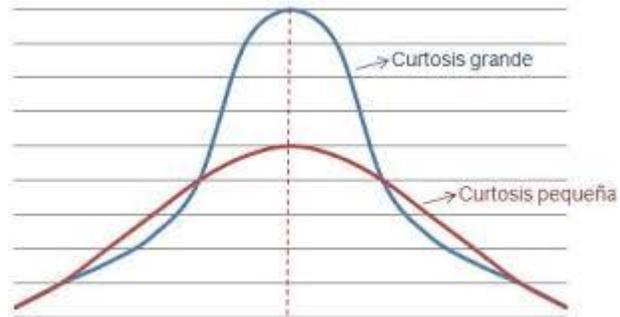


## APUNTAMIENTO

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.



## ESTADÍSTICOS DE POSICIÓN INDIVIDUAL

Se centra la atención en la descripción de observaciones particulares, en concreto, en estadísticos que nos van a ofrecer información

sobre un valor concreto en relación a la posición que ocupa dentro de un conjunto de valores observados.

## LOS 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.

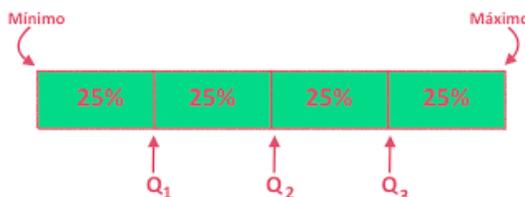


## LAS PUNTUACIONES TÍPICAS

Son los primeros datos de los que habitualmente disponemos, pero la comparación de las puntuaciones directas de un mismo sujeto en dos variables puede llevarnos a confusión. De hecho, conocida una puntuación directa no sabemos si se trata de un valor alto o bajo porque esto depende del promedio del grupo.

## MEDIDAS DE POSICIÓN

Nos facilitan información sobre la serie de datos que estamos analizando. La descripción de un conjunto de datos, incluye como un elemento de importancia la ubicación de éstos dentro de un contexto de valores posible



Medidas de posición		
Cuartiles	Deciles	Percentiles
$\frac{k \cdot N}{4}$	$\frac{k \cdot N}{10}$	$\frac{k \cdot N}{100}$
$Q_k = L_i + \left( \frac{\frac{k \cdot N}{4} - F_{i-1}}{f_i} \right) c$	$D_k = L_i + \left( \frac{\frac{k \cdot N}{10} - F_{i-1}}{f_i} \right) c$	$P_k = L_i + \left( \frac{\frac{k \cdot N}{100} - F_{i-1}}{f_i} \right) c$

$$T = z \cdot S + M_T$$

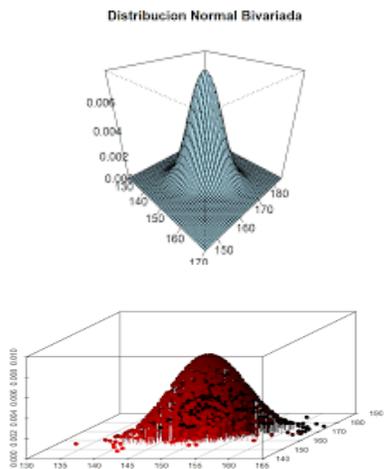
donde

S: Desviación Típica de la distribución derivada.

$M_T$ : Media de la distribución derivada.

## ORGANIZACIÓN Y REPRESENTACIÓN GRÁFICA DE DATOS MULTIVARIADOS

Se describen una serie de procedimientos asociados al tratamiento conjunto de dos o más variables, los cuales van a permitir extraer diversas facetas de la información compartida por esas variables. En bastantes momentos se va a ceñir esta exposición al caso bivariado (dos variables) por ser más sencillo en su presentación y por tratarse, con frecuencia, del caso particular más simple del modo general de abordar el problema a nivel multivariado (dos o más variables)

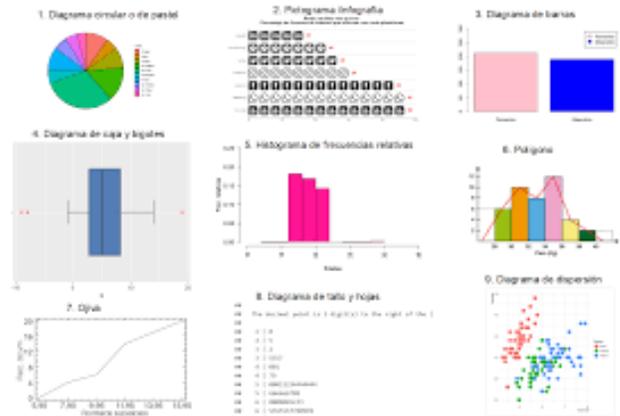


## LA TABLA DE CONTINGENCIA

una tabla de doble entrada en que cada lado de la tabla contiene las modalidades de una variable. En las casillas interiores de la tabla aparecen las frecuencias conjuntas (ya sean absolutas, relativas o porcentajes) de la combinación de los valores fila y columna correspondientes.

## 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



## LA DISTRIBUCIÓN CONJUNTA MULTIVARIADA.

Resumen básico de la información de un grupo de 2 o más variables consiste en la distribución conjunta de frecuencias de las mismas, la cual se basa en el conteo del número de casos (frecuencias) que presentan las distintas combinaciones de valores que a nivel empírico se hayan dado para esas variables.

Sino a todas las posibles combinaciones de los valores de las variables que se consideren – excepto aquellas combinaciones que no se hayan presentado a nivel empírico y que por tanto no tiene sentido incluir en la distribución de frecuencias

## PROBABILIDAD CONDICIONAL

### Tablas de contingencia

	F	M	TOTAL
Niños	30	20	50
Adultos	52	38	90
Senior	41	19	60
TOTAL	123	77	200

$$\frac{463}{4} = 115,5 \quad L_i = 300 \quad 115,5 - 85 = 30,75$$

$$f_i = 90 \quad I_c = 100 \quad Q_1 = 300 + \frac{30,75}{90} * 100 = 334$$

$$4^{\text{ª sextil}} : \text{posición} = \frac{4(463)}{6} = \frac{1852}{6} = 308,66$$

$$308,66 - 295 = 13,66 \quad f_i = 70$$

$$4^{\text{ª S}} = 500 + \frac{13,66}{70} * 100 = 59,51$$

$$7^{\text{ª decil}} : \text{posición} = \frac{7(463)}{10} = \frac{3241}{10} = 324,1$$

$$324,1 - 295 = 29,1 \quad f_i = 70$$

$$7^{\text{ª D}} = 500 + \frac{29,1}{70} * 100 = 541,57$$

$$30^{\text{ª percentil}} : \text{posición} = \frac{30(463)}{100} = \frac{13890}{100} = 138,9$$

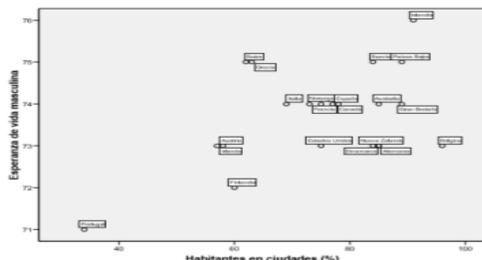
## REPRESENTACIÓN GRÁFICA EN EL ANÁLISIS DE DATOS

Se diferencia su presentación en función del tipo de variables al que van dirigidos. No se van a presentar gráficos específicos para las variables ordinales, si bien, puede ser utilizado cualquiera de los orientados a variables categóricas o, si se asume naturaleza cuantitativa para las mismas, los orientados a este tipo de variables.

### EL CASO DE DOS VARIABLES CUANTITATIVAS

El diagrama de dispersión (bivariado) Ejemplo de diagrama de dispersión obtenido a partir de los datos de una muestra de 8 fumadores en las variables "Nº de cigarrillos que, en promedio, se fuma al día" y "Puntuación en un test de ansiedad [0, ..., 30]". Se muestran también los datos a partir de los que ha sido obtenido el mismo con el programa SPSS:

Otro ejemplo de diagrama de dispersión obtenido, en este caso, a partir de los datos de esperanza de vida masculina y el porcentaje de población que viven en ciudades en un conjunto de países de la OCDE en el año 199



## CÁLCULO PARA UNA DISTRIBUCIÓN DE FRECUENCIA

Para el cálculo de esta medida en datos agrupados en una distribución de frecuencia, se

utiliza el mismo procedimiento estudiado para el cálculo de la Mediana, e; cual es:

Se efectúa la columna de las frecuencias acumuladas.

Se determina la posición del término cuyo valor se pretende calcular, en caso de ser el primer cuartil.

## EL CASO DE DOS VARIABLES CATEGÓRICAS

El diagrama de barras tridimensional o 3-D Ejemplos de diagrama de barras 3-D con la

distribución conjunta de frecuencias absolutas de "Estado de ánimo" y "Vivir residencia",

intercambiando la posición de ambas variables referencia al primero como diagrama de barras de frecuencias absolutas de la variable "Estado de ánimo" agrupada en función de "Vivir residencia", mientras que al segundo como diagrama de barras de la variable "Vivir residencia" agrupada en función de "Estado de ánimo".

## EL CASO DE UNA VARIABLE CATEGÓRICA Y UNA VARIABLE CUANTITATIVA

El diagrama de dispersión también puede ser aplicado en la representación conjunta de la

distribución de frecuencias absolutas de una variable categórica y una variable cuantitativa. A

este tipo de gráfico se le denomina en algunos textos como diagrama de puntos y es habitual

que aparezca representada la variable categórica en el eje de abscisas y la variable cuantitativa en el eje de ordenadas.

