

Índice

Índice	1
UD1 Estadística descriptiva unidimensional	3
INTRODUCCIÓN	3
INTRODUCCIÓN	4
OBJETIVOS	4
CONOCIMIENTOS PREVIOS	5
1. INTRODUCCIÓN A LA ESTADÍSTICA	5
2. ESTADÍSTICA DESCRIPTIVA	6
2.1 CONCEPTOS BÁSICOS	7
2.2 VARIABLES	9
3. TABLAS DE FRECUENCIA	12
3. TABLAS DE FRECUENCIA	14
3. TABLAS DE FRECUENCIA	16
4. PARÁMETROS ESTADÍSTICOS	17
4.1 MEDIDAS DE POSICIÓN	17
4.1.1 MEDIDAS CENTRALES	17
4.1.2 MEDIA ARITMÉTICA	18
4.1.3 MEDIA ARITMÉTICA PONDERADA	20
4.1.4 MEDIA GEOMÉTRICA	20

The logo for Cartagena99 features the text 'Cartagena99' in a stylized, blue, serif font. The '99' is significantly larger and more prominent than the rest of the text. The logo is set against a light blue background with a subtle gradient and a soft shadow effect.

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

UD1 Estadística descriptiva unidimensional

4.2 MEDIDAS NO CENTRALES	26
4.3 MEDIDAS DE DISPERSIÓN	28
4.3.1 MEDIDAS DE DISPERSIÓN ABSOLUTAS	28
4.3.2 MEDIDAS DE DISPERSIÓN RELATIVAS	30
4.4 MEDIDAS DE FORMA	31
5. REPRESENTACIONES GRÁFICAS	35
5.1 HISTOGRAMA	35
5.2 CAJAS Y BIGOTES (BOX-PLOT)	35
5.3 DIAGRAMA DE BARRAS	36
5.4 DIAGRAMA EN ESCALERA	36
CONCLUSIONES	37
RECAPITULACIÓN	37
BIBLIOGRAFÍA	38



Cartagena99

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

UD1 Estadística descriptiva unidimensional

UD1 Estadística descriptiva unidimensional

INTRODUCCIÓN

La estadística se puede entender como el estudio de los métodos de tratamiento de los datos que podamos obtener de cualquier tema que deseemos estudiar. En concreto se centrará en la recolección, descripción y análisis de los datos ofreciendo unos resultados que nos permitirán interpretar la información obtenida.

Si esta definición la aplicamos a la Bioestadística, podemos entender que los procesos previamente descritos se asocian a cualquier estudio que nos proporcione datos procedentes de seres vivos, algo en lo que se engloba la Nutrición Humana y la dietética ya que toda su actividad se centrará en el ser humanos y los datos que se obtengan procederán de éstos.

Sin embargo resulta importante destacar que la Bioestadística no es un bloque único sino que se subdivide, habitualmente, en 3 áreas de estudio.:

- Estadística descriptiva o deductiva
- Estadística probabilística o de probabilidad
- Estadística Inferencial o inductiva.

A la hora de desarrollar cualquier estudio estadístico, éste se define a través de diferente etapas muy claras,¹ y a lo largo del presente curso iremos profundizando en cómo abordar dichos pasos y cuáles de los anteriores bloques encajan en cada uno de esos pasos.

En esta primera Unidad, comenzaremos a estudiar la Estadística descriptiva. Entendiendo ésta como la rama de la estadística que se centrará en los métodos para organizar, resumir y presentar datos de una forma sencilla y que proporcionen al observador la mayor información posible.

Para ello, en esta unidad se introducirán una serie de términos esenciales

**CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70**

**ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70**

- Ordenación y representación gráfica de éstos
- Descripción de las características de los datos
- Análisis formal de los datos.”

INTRODUCCIÓN

Todo ello resulta esencial para desarrollar las actividades de organización, clasificación y construcción de elementos que nos permitan presentarlos de forma sencilla, clara y entendible para los observadores.

Pero la información que nos puedan proporcionar estas medidas no es suficiente y deben introducirse más elementos que nos permitan conocer más cosas de los parámetros utilizados en el estudio. En concreto se abordarán las denominadas como Medidas de dispersión y centralización que proporcionarán información acerca de la distribución de los datos sin que sea necesario que los conozcamos todos.

Para finalizar en esta unidad se abordarán las formas de presentación de estos datos, de modo que sea más sencillo entender los datos obtenidos y podamos interpretarlos más fácilmente. Estos medios de representación se clasificarán principalmente en:

- Las tablas de frecuencia
- Las representaciones gráficas

En resumen, a lo largo de esta unidad se explorarán las principales técnicas, tanto numéricas como de representación gráfica, que nos van a permitir describir los datos recogidos de tal modo que conozcamos sus características más importantes.

OBJETIVOS

Con el estudio de esta unidad se pretende establecer unas bases sólidas sobre la estadística descriptiva, enfocándonos inicialmente en:

**CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70**

**ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70**

Cartagena99

CONOCIMIENTOS PREVIOS

Para estudiar esta unidad serán necesarias nociones de matemáticas elementales.

1. INTRODUCCIÓN A LA ESTADÍSTICA

La estadística Si escuchamos la palabra estadística, lo primero que nos viene a la cabeza son series de datos numéricos presentados de forma ordenada y sistemática.

Son numerosas las oportunidades en las que podemos observarlas en nuestra vida habitual, desde datos de población activa al número de balones recuperados del jugador de nuestro equipo de futbol favorito.

La Estadística se ocupa de los métodos y procedimientos para recoger, clasificar, resumir, hallar regularidades y analizar los datos, siempre y cuando la variabilidad e incertidumbre sea una causa intrínseca de los mismos; así? como de realizar inferencias a partir de ellos, con la finalidad de ayudar a la toma de decisiones y en su caso formular predicciones.

La palabra estadística proviene de la forma femenina del término alemán Statistik, y puede tener dos acepciones:

- estadística (en minúscula): Es una colección de datos numéricos presentados de manera ordenada y sistemática.
- Estadística: Ciencia que estudia la realidad utilizando grandes conjuntos de datos, obteniendo las regularidades que se presentan para describir su comportamiento y predecir su evolución futura. Esta es la que se desarrolla en la etapa de i) ordenación y desarrollo de las representaciones gráficas de los datos y ii) descripción de las características de los datos.



**CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70**

**ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70**

Cartagena99

Por otro lado, tenemos la **Estadística Matemática, Inferencia Estadística o Estadística Inductiva**, que es aquella rama de la Estadística, que se apoya en el cálculo de probabilidades, a partir del estudio descriptivo de datos muestrales, para efectuar estimaciones, predicciones y otras generalizaciones sobre la población de partida, requiriendo verificaciones para tener un grado razonable de seguridad en las inferencias.

La Estadística sirve de base para la toma de decisiones, tanto a nivel institucional como en el seno de organizaciones, ciencias sociales o en las ciencias de la salud, entre otras.

De esta manera, la Bioestadística es la rama de la Estadística que estudia la utilización de métodos estadísticos en problemas relacionados con las ciencias de la salud (Medicina, Enfermería, Biología, Psicología, etc.) en las que la variabilidad (distribución de valores) y la incertidumbre (por ejemplo debida al desconocimiento de los mecanismos fisiológicos) forman parte de su naturaleza.

Un estudio así realizado nos proporcionará resultados que aporten información al tema estudiado, así como ruido de fondo residual que no aporten información, siendo la estadística la responsable de extraer la máxima información útil de los datos recogidos.

2. ESTADÍSTICA DESCRIPTIVA

La estadística descriptiva es la parte de la estadística que se enfoca en la descripción y análisis de un conjunto de datos que se hayan recogido previamente pero que no se encarga de la obtención de conclusiones.

Su principal misión, por tanto, será la de servir como una herramienta que nos permita, a los investigadores, manejar los datos obtenidos para expresarlos de un modo sencillo que nos permita describir sus propiedades así como poder

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99



campo de acción sobre 2 de las 4 etapas previamente descritas anteriormente en las que se divide un análisis estadístico. Esta es la que se desarrolla en la etapa de i) ordenación y desarrollo de las representaciones gráficas de los datos y ii) descripción de las características de los datos.

Cómo ya se ha mencionado, **la estadística descriptiva se enfoca en la ordenación, registro y presentación de los datos** (Los datos son medidas y/o números recopilados a partir de la observación. Es decir, puede entenderse como la información recopilada que puede ayudar en la toma de decisiones ante una situación determinada.) recogidos, pero antes deben definirse una serie de términos esenciales para entender la estadística descriptiva y de dónde se obtienen esos datos a los que nos referimos.

2.1 CONCEPTOS BÁSICOS

Lo primero es entender que los datos recogidos mediante nuestras observaciones proceden de algún lugar y para ello debemos entender de algunos términos asociados a ello:

Población

Colectivo de **individuos o elementos que poseen ciertas características comunes y que son el objeto de la observación y estudio estadístico, con la idea de obtener un resultado.** Ejemplos: Hoteles de Madrid, Centros de salud de Barcelona, etc.

El tamaño que tiene una población es un factor muy importante para el proceso investigación estadística, y aunque se desarrollará con mayor nivel de detalle en próximas unidades, es importante destacar que este tamaño viene dado por el número de elementos que constituyen la población. **Según el número de elementos la población puede ser finita** (Población finita, cuando la población está formada por un número limitado de elementos.) o **infinita.** (Población infinita, sería cuando el número de elementos que confirman la población es

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99

Unidad estadística, individuo o elemento

Cada uno de los entes o fenómenos que contienen cierta información que se desea estudiar y que pertenecen a la población en estudio. Se pueden entender también como “Cada uno de los elementos de la población. Ejemplos: Cada uno de los hoteles de Madrid, cada uno de los Centros de Salud de Barcelona.

Caracteres o variables

Propiedades, rasgos o cualidades que presentan los elementos de una población y que varían entre los diferentes individuos de esta.

Parámetro

Función definida sobre los valores numéricos de características medibles de una población. Los parámetros poblacionales se denotan con letras del alfabeto griego.

Estadístico

Es una medida descriptiva (función) de la muestra y que estima el parámetro de la población.

Estimador

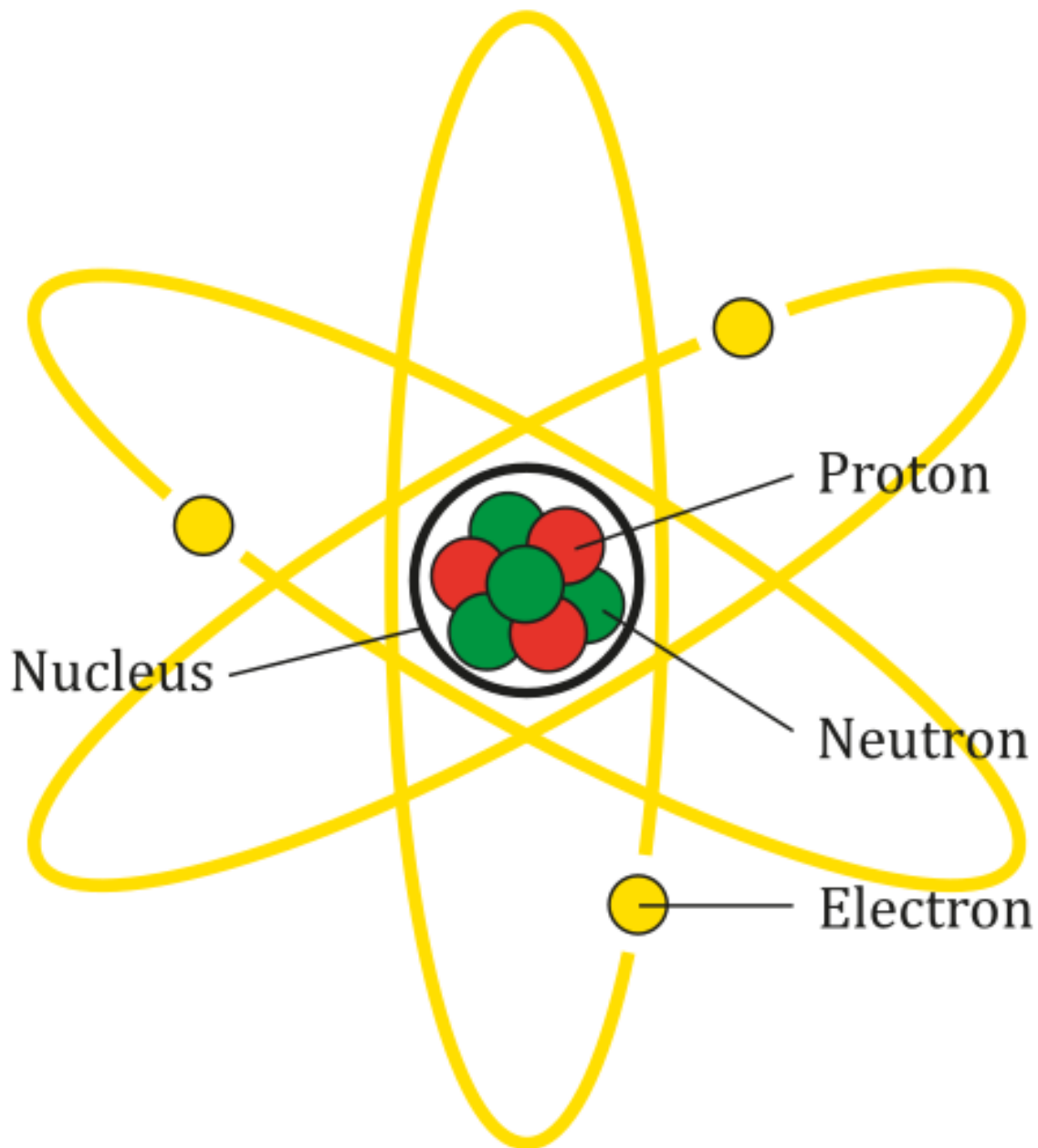
Función de los valores de una muestra que se elabora para indagar el valor de un parámetro de la población de la que procede la muestra. Los estimadores muestrales se denotan con letras de nuestro alfabeto.

Representación gráfica de Universo, población y muestra

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99



Elaboración propia
Representación gráfica de Universo, población y muestra

2.2 VARIABLES

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99

representan categorías que son mutuamente excluyentes, aunque se utilicen números para cada categoría no representan cantidades, no admitiendo gradación. También las conocemos como **variables cualitativas**.

Para su medición usamos escalas de tipo nominal u ordinal.

Escala nominal: Cuyos valores numéricos representan una categoría e identifican un grupo de pertenencia. La asignación de los valores se realiza en forma aleatoria porque NO cuenta con un orden lógico y solo permite clasificar (no jerarquizar ni ordenar). Ejemplos. (Los dorsales de los futbolistas, nacionalidad, religión, género, color favorito, estado civil, marca de coche, calificación de un examen (no numérica), etc.) x.

Dentro de estos atributos podemos distinguir:

- **Dicotómicas o binarias,** con solo dos categorías. Un **ejemplo sería el Sexo:** hombre/mujer.
- **Policotómicas:** Aquellas variables que pueden adoptar más de dos modalidades o categorías. Un **ejemplo sería el grupo sanguíneo.**

Escala ordinal: Cuyos valores numéricos representan una categoría o identifican un grupo de pertenencia contando con un orden lógico. Podemos identificar si una categoría es mayor o menor que otra, por lo que se encuentran jerarquizadas y ordenadas. Ejemplos. (el grado de satisfacción, calidad de servicios, percepción del dolor, medallas en las olimpiadas,, etc.)

Variables cuantitativas

Las **variables** son caracteres cuantitativos, también llamadas **variables cuantitativas**, y se expresan mediante cantidades numéricas que permiten **hacer operaciones matemáticas**, de forma que la medición de una variable da lugar a un valor. Pueden medirse mediante una escala de intervalo o de razón.

- **Variables discretas:** Las que sólo pueden tomar un valor entero, es decir, determinados valores y no un valor intermedio entre dos de aquellos (el número de valores es finito o infinito. Se diferencian de las continuas en que, dadas dos observaciones suficientemente próximas, no se puede encontrar

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99

número con decimales ya que un ser vivo (para estar vivo) debe tomarse como una unidad.

Es decir Si el promedio fuese de 14,8 gatos, implicaría que hay 14 gatos enteros (vivos) y el 0.8 de un gato (evidentemente no estaría vivo).

En este caso se redondearía al número entero “más cercano”, siendo 15.

Por lo que el promedio de gatos vivos en Villanueva de la Cañada serían 15 gatos.

- **Variables continuas:** Las variables pueden tomar infinitos valores dentro de un intervalo. Ejemplos. ²

La variable estadística se denota con mayúsculas. Asimismo, cada una de estas variables puede tomar distintos valores siendo su notación la siguiente:

$$X = (x_1, x_2, x_3, \dots, x_{k-2}, x_{k-1}, x_k)$$

Tipos de variables ³

Diferencias entre variables cualitativas y cuantitativas

Variables Cuantitativas

¿Cómo se define una variable?

Especificidad

Relación entre variables

Variables Cualitativas

Numéricamente

Suele ser específica, no da lugar a duda alguna.

Determinan la correlación o asociación en ellas

2

Longitud de un palo

Altura de la clase

Tiempo que tarda en entregar la pizza el repartidor

Precio de un producto

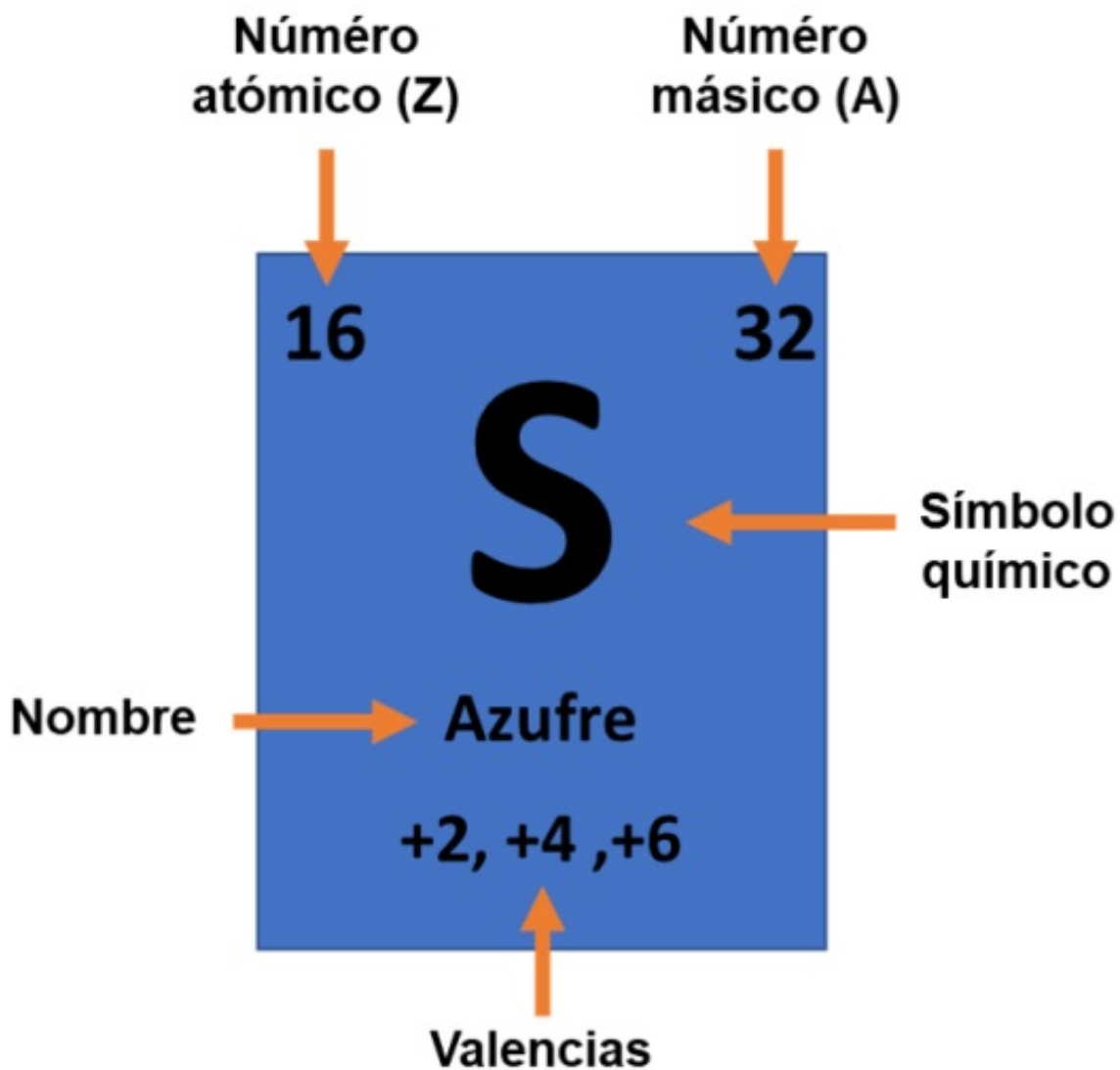
Distancia entre dos puntos

3

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99



Elaboración propia
Tipos de variables

3. TABLAS DE FRECUENCIA

Tenemos que valorar si los datos están agrupados o no.

estadísticas con datos sin agrupar

**CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70**

**ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70**

Es el conjunto de valores que ha tomado una variable con sus frecuencias correspondientes. Simbólicamente viene dada por los pares (x_i, n_i) donde x_i son los valores de la variable y n_i son sus frecuencias.

Es necesario que las variables por estudiar sean ordenables, es decir, debe ser posible establecer una relación de orden entre las variables.

frecuencia absoluta

La **frecuencia absoluta** (n_i ($n_i \geq 0$)) es el número de observaciones, siendo N el número de observaciones, que presentan el valor x_i . El subíndice i puede tomar los valores $i=1, 2, \dots, k$ donde

$$\sum_{i=1}^k n_i = N$$

frecuencia relativa

La **frecuencia relativa** f_i es la proporción de observaciones que presentan el valor x_i . Se calcula como el cociente entre la frecuencia absoluta n_i y el número total de observaciones, N , es decir,

$$f_i = \frac{n_i}{N}$$

frecuencias relativas porcentuales

Si calculamos el porcentaje de las frecuencias relativas tendremos las **frecuencias relativas porcentuales**, que definimos como el porcentaje de observaciones que presentan el valor x_i . Para su cálculo se multiplican las frecuencias relativas por 100 y la suma de todas las frecuencias relativas porcentuales es igual a 100.

$$p_i = f_i * 100$$

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99



frecuencias relativas acumuladas, denominadas F_i , y frecuencias relativas porcentuales acumuladas, denominadas P_i .

Estas definiciones nos permiten crear la tabla de frecuencias, que es la manera más adecuada para sintetizar los datos.

Tabla2. Ejemplo de tabla de frecuencia y distribución de los datos según las frecuencias

Valores de la variable X_i	Frecuencias absolutas n_i	Frecuencias relativas $f_i = n_i/N$	Frecuencias relativas porcentuales $p_i = f_i * 100$	Frecuencias absolutas acumuladas N_i	Frecuencias relativas acumuladas $F_i = N_i/N$	Frecuencias relativas porcentuales acumuladas $P_i = F_i * 100$
X_1	n_1	f_1	p_1	$N_i = n_i$	$F_i = f_i$	$P_i = p_i$
X_2	n_2	f_2	p_2	$N_2 = N_1 + n_2$	$F_2 = F_1 + f_2$	$P_2 = P_1 + p_2$
...
X_i	n_i	f_i	p_i	$N_i = N_{i-1} + n_i$	$F_i = F_{i-1} + f_i$	$P_i = P_{i-1} + p_i$
...
X_k	n_k	f_k	p_k	$N_k = N$	$F_k = 1$	$P_k = 100$
	$\sum_{i=1}^k n_i = N$	$\sum_{i=1}^k f_i = 1$	$\sum_{i=1}^k p_i = 100$			

3. TABLAS DE FRECUENCIA

Las distribuciones de frecuencias absolutas se conocen como estadísticas primarias, por obtenerse directamente de la observación de un carácter.

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Tabla de frecuencias para datos no agrupados

Variable	n_i	$f_i=n_i/N$	$P_i=f_i*100$	N_i	$F_i=N_i/N$	$P_i=F_i*100$
X_i						
3	20	0,22	22	20	0,22	22
4	25	0,28	28	45	0,5	50
5	14	0,16	16	59	0,66	66
6	12	0,13	13	71	0,79	79
9	10	0,11	11	81	0,9	90
10	9	0,1	10	90	1	100
	90	1	100			

En ocasiones el número de datos diferentes que se está estudiando es muy numeroso. La solución para crear una tabla de frecuencias es agrupar los datos en intervalos o clases, de modo que cada dato pertenezca a uno y solo un intervalo.

En consecuencia, los conceptos relativos a la frecuencia que hasta ahora se referían a los valores diferentes de los datos, al realizar la agrupación, deben hacer referencia a los intervalos.



Un **intervalo** se suele representar por $(L_{i-1}, L_i]$ y se define como el conjunto formado por todos los valores reales que son mayores o iguales que L_{i-1} (extremo inferior) y menores que L_i (extremo superior).

En distribuciones con agrupación en intervalos nos encontramos con otros conceptos adicionales:

Amplitud del intervalo

Es la diferencia entre el extremo superior y el inferior del mismo. Se representa por a_i y para el intervalo i -ésimo sería: $a_i = L_i - L_{i-1}$

$$a_i = L_i - L_{i-1}$$

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99

$$X_i = \frac{L_i + L_{i-1}}{2}$$

Densidad de frecuencia

Si las amplitudes de los intervalos son distintas, es necesario calcular las densidades de frecuencia, que son el resultado de dividir las frecuencias absolutas por la amplitud del intervalo. Se denota por h_i y será

$$h_i = \frac{n_i}{a_i}$$

recorrido

Se llama **recorrido** de un conjunto de datos a la diferencia entre el valor más grande y el más pequeño del conjunto. Se denota por R_e .

3. TABLAS DE FRECUENCIA

En el caso de las distribuciones con agrupación por intervalos aparece el error de agrupamiento (Error debido a la pérdida de información provocada al incluir un conjunto de observaciones en un intervalo.) por lo que, a pesar de que ayuda a resumir y clarificar la información es importante destacar la aparición de este tipo de error.

Ejemplo:

En las pacientes de una planta de ginecología se observa el número de mujeres que han dado a luz. Los resultados de la observación son:

a) Cuál es la población: Las pacientes de una planta de ginecología	Pacientes de ginecología	Mujeres que han dado a luz
---	--------------------------	----------------------------

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99

- f) Diga si es una distribución de datos agrupados o no agrupados: Es de datos no agrupados
- g) Calcule todas las frecuencias.

X_i	n_i	f_i	P_i	N_i	F_i	P_i
1	6	0,3	30	6	0,3	30
2	8	0,4	40	14	0,7	70
3	4	0,2	20	18	0,9	70
4	2	0,1	10	20	1	90
Total	20	1	100			

4. PARÁMETROS ESTADÍSTICOS

- 4.1 Medidas de posición
- 4.2 Medidas no centrales
- 4.3 Medidas de dispersión
- 4.4 Medidas de forma

4.1 MEDIDAS DE POSICIÓN

Las **medidas de posición** son coeficientes que tratan de representar una determinada distribución y pueden ser de dos tipos, centrales y no centrales.

- 4.1.1 Medidas centrales
- 4.1.2 Media aritmética
- 4.1.3 Media aritmética ponderada
- 4.1.4 Media geométrica
- 4.1.5 Media harmónica
- 4.1.6 Mediana
- 4.1.7 Moda

4.1.1 MEDIDAS CENTRALES

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99

Los promedios que vamos a estudiar son la media aritmética, la moda y la mediana.

4.1.2 MEDIA ARITMÉTICA

4.1.2 Media aritmética

La **media aritmética** será la suma de todos los valores de la variable dividida entre el número total de elementos, por lo que es el valor que habitualmente se toma como representación de los datos.

$$\bar{x} = \frac{\sum_{i=1}^k x_i n_i}{\sum_{i=1}^k n_i} = \frac{\sum_{i=1}^k x_i n_i}{N}$$

$$\bar{x} = \sum_{i=1}^k x_i f_i$$

En el caso en que los datos están agrupados, se toma la marca de la clase (Es el punto medio de cada intervalo. Será el valor que representa todo el intervalo para calcular algunos parámetros como la media o desviación típica, entre otros.) como representante del intervalo y se realizan todos los cálculos como si los valores de la variable fueran las marcas de las clases.

La unidad de medida será la misma que la de la variable estudiada y su valor tiene que estar comprendido en el recorrido de la variable, aunque no tiene por qué coincidir con un valor de la variable. Además, en el caso de variables discretas, puede coincidir con un valor no observable de la variable.

Ejemplo 4

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Si todos los valores de la variable se multiplican por una misma constante, la

media aritmética queda multiplicada por dicha constante.

4. Si una variable Y es transformación lineal de otra variable X ($Y = a * X + b$, siendo a y b números reales), la media aritmética de Y sigue la misma transformación lineal respecto a la media aritmética de X .

Ventajas

1. Utiliza toda la información: hace uso de todos los valores para su cálculo.
2. Es única.
3. Es sencilla de calcular y tiene una interpretación intuitiva.

Inconvenientes

1. Es muy sensible a los valores extremos de la variable.
2. En el caso de variable agrupada en intervalos, su valor depende de los intervalos elegidos.
3. No puede calcularse en distribuciones abiertas (si el límite inferior del primer intervalo o el límite superior del último intervalo están sin determinar), a no ser que se disponga de información adicional.

4 Media en horas de ocupación de quirófanos durante junio

Duración en horas de
quirófanos

	3
	4
	5
	6
	9
	10
Total	37

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99

Resultado: 6,17 horas de media de ocupación de quirófanos durante junio.
Observamos como el valor no coincide con valor observable de alguna variable.

4.1.3 MEDIA ARITMÉTICA PONDERADA

4.1.3 Media aritmética ponderada

Puede utilizarse para mostrar cambios porcentuales en una serie de números positivos.

La media geométrica proporciona una medida precisa de un cambio porcentual medio en una serie de números. Se representa por *MG* y su cálculo – efectuando la notación habitual– sigue la siguiente expresión.

$$MP = \bar{x} = \frac{\sum_{i=1}^n x_i w_i}{N}$$

4.1.4 MEDIA GEOMÉTRICA

4.1.4 Media geométrica

Puede utilizarse para mostrar cambios porcentuales en una serie de números positivos.

La media geométrica proporciona una medida precisa de un cambio porcentual medio en una serie de números. Se representa por *MG* y su cálculo – efectuando la notación habitual– sigue la siguiente expresión.

$$MG = \bar{x} = \sqrt[n]{x_1 * x_2 * \dots * x_n}$$

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99

$$H = \bar{x} = \frac{N}{\frac{1}{X_1} + \frac{1}{X_2} + \dots + \frac{1}{X_N}}$$

Se utiliza para calcular el valor medio de magnitudes expresadas en términos relativos como velocidades, tiempos, rendimiento, tipo de cambio monetario, etc., y su principal contrariedad es que cuando algún valor de la variable es cero o próximo a cero no se puede calcular.

4.1.6 MEDIANA

4.1.6 Mediana

La mediana (Tiene la misma unidad de medida que la variable.) es el valor de la variable que divide las observaciones en dos grupos de igual número de elementos, de modo que en el primer grupo todos los datos sean menores o iguales que la mediana, y en el otro grupo, todos los datos sean mayores o iguales.

Por lo tanto, es una cantidad que indica orden dentro de la ordenación.

Para su cálculo hay que ordenar las variables de menor a mayor y ver si los datos se encuentran agrupados o no.

Para distribuciones de datos no agrupados, (Serán aquellas distribuciones en las que las frecuencias con las que aparece un dato estadístico no se ha modificado, es decir se expresa del mismo modo en que han sido recogidas.) con todas las frecuencias unitarias se pueden presentar dos casos:

1. Número de observaciones impar: la mediana es el valor central una vez ordenados de menor a mayor.
2. Número de observaciones par: la mediana es la semisuma de los dos

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

no existe una frecuencia acumulada igual a $N/2$.

UD1 Estadística descriptiva unidimensional

La mediana es el valor de la variable que corresponde a la primera frecuencia acumulada inmediatamente superior a $N/2$.

Personas Ocupadas	Nº de familias	Ni
1	6	6
2	8	14
3	4	18
4	2	20
Total	20	

$$N/2 = 20/2 = 10$$

Me=2 personas ocupadas

Existe una frecuencia acumulada igual a $N/2$.

La mediana será en ese caso la media aritmética entre el valor de la variable al que le corresponde $N/2$ y el siguiente.

Personas Ocupadas	Nº de familias	Ni
1	6	6
2	5	11
3	7	18
4	4	22
Total	22	

$$N/2 = 22/2 = 11$$

Me= $(2+3) / 2 = 2,5$ personas ocupadas.

Para distribuciones de datos agrupados (Serán aquellas distribuciones en las que los datos se agrupan en los que denominamos como “intervalos o clases”. Se hace esto cuando tenemos gran cantidad de elementos y necesitamos facilitar el proceso de análisis.) se busca el intervalo mediano. El intervalo mediano es el primero cuya frecuencia acumulada sea igual o inmediatamente superior a $N/2$ y siempre suponemos que los datos se distribuyen uniformemente en el intervalo.

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99

Ejemplo 6

Ventajas

1. Es única.
2. La mediana no se ve influida por los valores extremos de la variable, ya que no depende de los valores de la variable, sino del orden de los mismos. Por ello, su uso es más adecuado que la media aritmética en distribuciones campaniformes fuertemente asimétricas, o distribuciones con forma de L o de J.
3. En caso de intervalos abiertos, puede calcularse sin problema ya que en su cálculo no intervienen los valores extremos, sino tan sólo se hace uso del intervalo mediano.

Inconvenientes

1. En su determinación no intervienen todos los valores de la variable, sólo es preciso saber el valor de la observación central.
2. No es adaptable al cálculo algebraico.
3. En el caso de distribuciones agrupadas en intervalos, su valor depende de los intervalos elegidos.

5

Mediana de horas de estudio de un alumno a lo largo de mes de enero.

Horas de estudio: 23, 26, 29, 31 y 3

Me= 29 horas de estudio

6

Salario €/hora (intervalo)	Nº de trabajadores	Ni
15-25	a	a

**CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70**

**ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70**

Cartagena99

$$N/2 = 54/2 = 27$$

$$Me = 55 + ((27 - 21/14) * 20) = 63,57 \text{ €/hora}$$

4.1.7 MODA

4.1.7 Moda

Es el valor (También posee la misma unidad de medida que la variable.) de la variable que más veces se repite, es decir, el valor que tiene mayor frecuencia absoluta.

Para distribuciones de datos sin agrupar, la obtención de la moda es inmediata, siendo el valor de la variable que tenga mayor frecuencia absoluta, n_j .

Ejemplo 7

En los supuestos que la distribución venga dada en intervalos, se pueden producir dos casos: que tengan la misma amplitud, o que esta sea distinta. En ambos casos el objetivo es encontrar un valor que represente la moda.

En los intervalos con la misma amplitud es evidente que una vez determinada la mayor frecuencia a esta no le corresponde un valor sino un intervalo, por lo que tenemos un intervalo modal.

Para calcular el representado del intervalo tendremos que observar si los intervalos son de amplitud constante o variable. (Recuerda que cuando los datos estadísticos que tenemos que manejar son numerosos, se puede organizar en una distribución de frecuencias de valores agrupados en intervalos de amplitud constante y variable.)

En los intervalos de amplitud constante, el intervalo modal es aquel que presenta mayor n_j .

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP: 689 45 44 70

Cartagena99

7

Nº personas ocupadas	Nº de familias
1	6
2	8
3	4
4	2
Total	20

Me= 2 personas ocupadas

8

Salario €/hora (intervalo)	Nº de trabajadores	ai
15-35	9	20
35-55	12	20
55-75	14	20
75-95	10	20
95-115	9	20
	54	

Mo= 55+ ((10/12+10)*20) = 64,1 €/hora

4.1.7 MODA

En el caso de Intervalos de amplitud variable: el intervalo modal es aquel que presenta mayor h_i . La moda se calcula, de forma aproximada, mediante la fórmula:

$$Mo = L_{i-1} + \frac{h_{i+1}}{h_{i-1} + h_{i+1}} a_i$$

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP:689 45 44 70

observar que la moda estará más cerca de aquel



intervalo adyacente con mayor frecuencia absoluta.

Ventajas

1. Tiene una interpretación y cálculo sencillos.
2. No está influida por los valores anormalmente grandes o pequeños de la variable.
3. Puede no importar que los intervalos sean abiertos, si se puede determinar cuál es el intervalo modal y se pueden calcular las alturas de los intervalos adyacentes a éste.

Inconvenientes

1. En su determinación no intervienen todos los valores de la variable.
2. Sólo se debe aplicar en distribuciones con un número elevado de observaciones.
3. No está definida algebraicamente.
4. Puede no existir y no tiene por qué ser única. La distribución puede ser bimodal. (Será aquella distribución en la que dos de los valores representados aparecen con más frecuencia que el resto en el conjunto total de datos.) 1. trimodal, etc., siendo necesario utilizar otro promedio.
5. Depende de los intervalos elegidos en el caso de variable agrupada en intervalos.
6. En ocasiones las distribuciones abiertas, por la indefinición de los límites, puede impedir el cálculo de la moda.

9

Salario €/hora (intervalo)	Nº de trabajadores	a_j	$h_j = n_j / a_j$
15-45	9	30	0,3
45-55	12	10	1,2
55-95	14	40	0,35
95-115	0	0	0,15

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99

Percentiles o cuantiles (Podemos decir que los cuantiles son unas medidas que dividen la distribución en un cierto número de partes y que como característica principal, cada parte contiene el mismo número de frecuencias.)

Son medidas de localización similares a la mediana. Su función es informar del valor de la variable que ocupará la posición (en tanto por ciento) que nos interese respecto de todo el conjunto de observaciones.

Los cuantiles de uso más frecuente y conocidas son:

- **Cuartiles**, dividen la distribución en cuatro partes iguales (tres divisiones). C1 (25%), C2 (50%), C3 (75%).
- **Deciles**, dividen la distribución en 10 partes iguales (9 divisiones). D1 (10%), D2 (20%), ..., D9 (90 %).
- **Percentiles**, dividen a la distribución en 100 partes (99 divisiones). P1 (1%), P2 (2%), ..., P99 (99%). Lo interpretamos como el valor correspondiente al percentil 78, tiene un 78 % de los datos menores o iguales a él.

Existe un valor en el que coinciden los cuartiles, los deciles y percentiles, y corresponde con la mediana, ya que: $P_{50} = C_2 = D_5$.

El cálculo de los cuantiles sigue el mismo procedimiento que el que se ha utilizado en la mediana, tanto para los datos agrupados como para los datos sin agrupar, con la única diferencia que en lugar de utilizar $N/2$ se usará $N/100$, $2N/100$, ..., ó $99N/100$, según corresponda.

Para distribuciones con datos agrupados, se busca el intervalo cuya frecuencia absoluta acumulada sea mayor o igual que $N_i/100$:

$$P_j = L_{i-1} + \frac{\frac{jN}{100} - N_{i-1}}{n_i} a_i \quad \text{siendo } j = 1, 2, 3, \dots, 99$$

Ejemplo ¹⁰

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99

$$M_o = 45 + ((0,35/0,3+0,35)*100) = 50,38 \text{ €/hora}$$

4.3 MEDIDAS DE DISPERSIÓN

Son complementarias de las de posición, en el sentido que señalan la dispersión del conjunto de todos los datos de la distribución, respecto de la medida o medidas de localización adoptadas, es decir, nos mostraran si los valores de las observaciones están próximos entre sí o están muy separados.

Distinguiremos entre medidas absolutas y relativas de dispersión.

4.3.1 MEDIDAS DE DISPERSIÓN ABSOLUTAS

4.3.1 Medidas de dispersión absolutas

Recorrido

Viene expresada en la misma unidad de medida que la variable.

Es la diferencia entre el mayor y menor valor de una distribución:

$$R = \max(x_i) - \min(x_i)$$

Observaciones:

- Se trata de una primera aproximación a la medida de la dispersión de una distribución, es muy burda y no está referida a ninguna medida de tendencia central.
- Si los datos están agrupados en intervalos el recorrido será la diferencia entre el valor del límite superior del último intervalo y el valor del límite inferior del primer intervalo.

Ejemplo **11**

Ventajas:

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99

La unidad de medida de la varianza será el cuadrado de la unidad en la que se mide la variable de interés.

Se define como la media aritmética de los cuadrados de las desviaciones de los valores de la variable respecto de la media aritmética de la distribución y viene expresada en la unidad de medida de la variable, *elevada al cuadrado*.

Responde a la fórmula:

$$S^2 = \frac{\sum_{i=1}^k x_i^2 n_i}{N} - \bar{x}^2$$

Como se puede observar en la definición, la varianza es un promedio del cuadrado de los errores que se cometen al considerar la media aritmética como «el representante» de todos y cada uno de los datos.

Ejemplo 12

Ventajas:

1. Utiliza toda la información.
2. Se presta al cálculo algebraico.

Inconvenientes:

1. No sirve para comparar dispersiones en el caso de que varias distribuciones posean distintas medias o bien vengán expresadas en distintas unidades de medida.
2. No puede obtenerse en distribuciones abiertas a no ser que se dé información complementaria.
3. Su valor depende del criterio de agrupación en intervalos.
4. Es muy sensible a los valores anormales de la variable.

Desviación típica o desviación estándar

Se define como la raíz cuadrada, con signo positivo, de la varianza, tomando el mismo valor que la variable y responde a la siguiente fórmula:

$$S = \sqrt{S^2}$$

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99

tendencia central entre dos distribuciones de datos diferentes. Por ello, a veces se recurre a medidas de dispersión relativas.

11

Horas de estudio de un grupo de 7 alumnos: 3, 5, 6, 8, 10, 13, 11,

R= máximo (13) – Mínimo (3)

R= 10 horas.

12

Ejemplo de intervalos de amplitud variable:

Salario €/hora (intervalo)	Nº trabajadores	X_i	$X_i n_i$	$X_i^2 n_i$
15-35	9	25	225	5.625
35-55	12	45	540	24.300
55-75	14	65	910	59.150
75-95	10	85	850	72.250
95-115	9	105	945	99.225
	54		3.470	260.550

=3.470/54= 64,26 €/hora

$S^2 = (260.550/54) - (64,26)^2$

13

Horas de estudio X_i	X_i^2
3	9
4	16
5	25
6	36
9	81
10	100
27	267

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99

Se enfocan a determinar la distribución estadística de forma independiente a las unidades en que se exprese la variable de interés.

Con las medidas de dispersión que hemos visto, las absolutas, no pueden compararse las dispersiones de dos distribuciones, si las variables no tienen la misma unidad de medida, o bien, si aún teniendo la misma unidad de medida, las medias aritméticas son muy distintas.

El coeficiente de variación de Pearson

Es una de las más significativas y determina el grado de significación de un conjunto de datos relativo a su media aritmética.

Se define como el cociente entre la desviación típica y la media aritmética de la distribución de datos y representa el número de veces que la desviación típica contiene a la media, luego, a mayor coeficiente, menor representatividad de la media. Es adimensional, no tiene unidad de medida.

$$CV = \frac{S}{\bar{x}} \geq 0$$

El coeficiente de variación de Pearson tiene dos utilidades:

Determinar si una media aritmética es representativa de la distribución, usando los los siguientes límites:

- Si $CV < 0,5$ (50%) se considera que la media es representativa.
- Si $CV > 0,5$ (50%) hay que cuestionarse su representatividad.
- Si $CV > 1$ (100%) la media no es representativa.

Además, el coeficiente de variación va a servir además para comparar la representatividad de la media de dos distribuciones, siendo más representativa la media en aquella distribución para la que el CV es menor. (Implica que tiene menor dispersión y es más homogénea.)

4.4 MEDIDAS DE FORMA

Nos dan información de la forma del histograma de su simetría y de la menor o

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

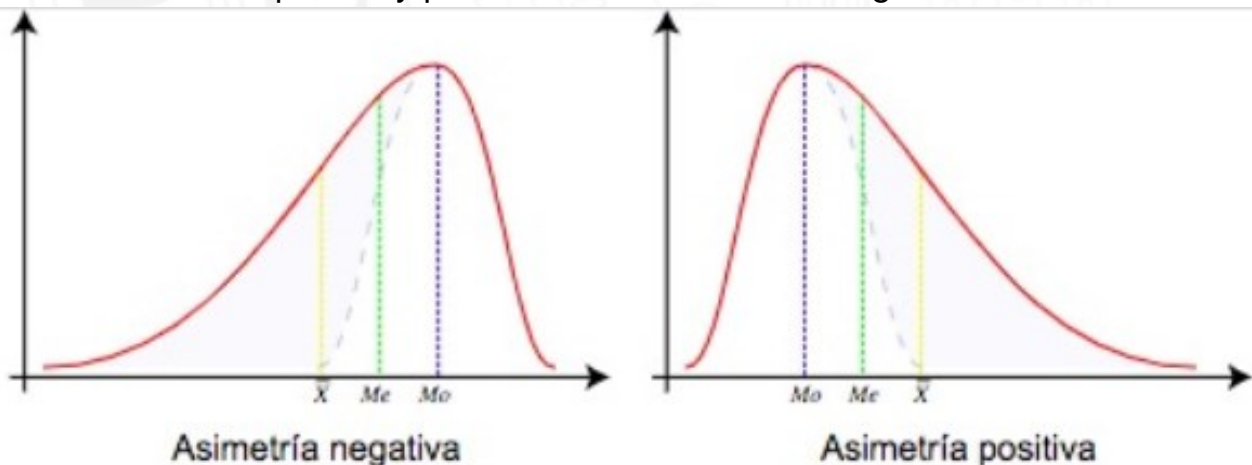
Cartagena99

aritmética.

Una distribución es simétrica cuando a la derecha y a la izquierda de la media existe el mismo número de valores, equidistantes dos a dos de la media y además con la misma frecuencia. En caso contrario la distribución es asimétrica.

Será asimétrica a positiva o la derecha si las frecuencias más altas se encuentran en el lado izquierdo de la distribución mientras que en el derecho hay frecuencias más pequeñas (cola).

Presentará asimetría negativa o la izquierda en caso contrario, quedando la cola en el lado izquierdo y puede estudiarse de forma gráfica o analítica.



https://commons.wikimedia.org/wiki/File:Posiciones_relativas_de_par%3%A1m
Asimetrías

Por tanto, esta medida debe reflejar dos aspectos:

1. Distancia de cada observación respecto a la media aritmética (es decir, la diferencia entre cada valor y la media aritmética).
2. Frecuencia de cada una de estas distancias (la que coincidirá, evidentemente, con la frecuencia de cada observación).

Coeficiente de asimetría de Fisher

Ahora pues, lo que hay que encontrar es el estadístico que determine la asimetría de la distribución de datos.

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99

$$\gamma_1 = \frac{m_3}{s^3}$$

$$\text{Donde } m_3 = \frac{\sum_{i=1}^k (x_i - \bar{x})^3 n_i}{N}$$

Para interpretar los resultados nos debemos de fijar en:

- $\gamma_1 = 0$ no puede decirse nada de la simetría de la distribución. (Condición necesaria pero no suficiente de simetría).
- $\gamma_1 > 0$ distribución asimétrica a la derecha.
- $\gamma_1 < 0$ distribución asimétrica a la izquierda.

Curtosis

Para estudiar el grado de curtosis de una distribución hay que tomar un modelo teórico como referencia, la representación gráfica tenga forma de campana simétrica y nos indica si la distribución es más o menos apuntada (estrecha y alargada) en la zona central.

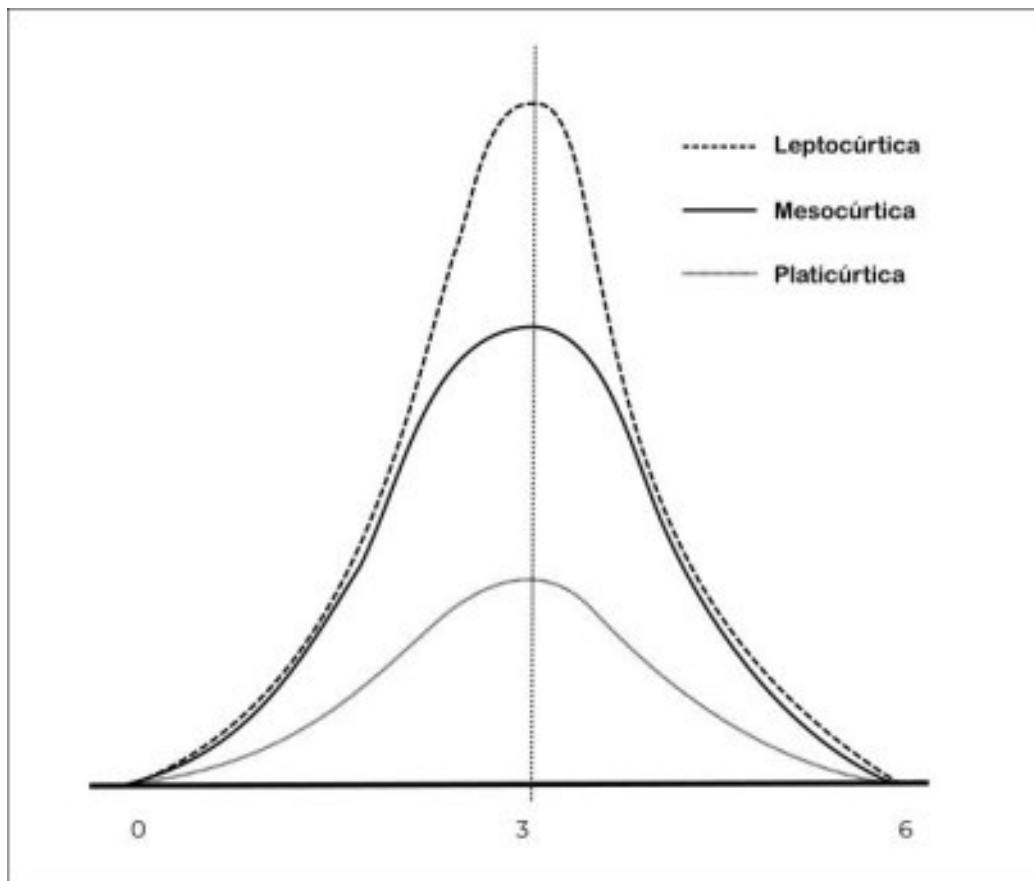
Podemos clasificarlas comparándolas con una distribución normal:

- **Leptocúrtica** (Es más apuntada que la distribución normal.)
- **Mesocúrtica** (Tiene el mismo apuntamiento que la distribución normal.)
- **Platicúrtica** (Es menos apuntada que la distribución normal.)

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99



Elaboración propia
Curtosis

Del mismo modo que en el caso del estudio de la asimetría, hay un coeficiente que permite clasificar los datos según la curtosis, denominado Coeficiente de apuntamiento de Fisher.

Como en el caso de la otra medida de forma, este indicador tampoco tiene dimensión y se calcula:

$$\gamma_2 = \frac{m_4}{s^4} - 3$$

Donde
$$m_4 = \frac{\sum_{i=1}^k (x_i - \bar{x})^4 n_i}{n}$$

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99

2.3.3 distribución platicúrtica o apriada.

UD1 Estadística descriptiva unidimensional

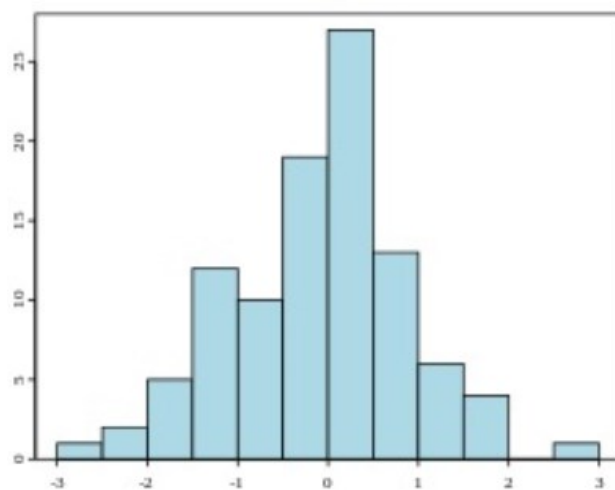
5. REPRESENTACIONES GRÁFICAS

- 5.1 Histograma
- 5.2 Cajas y Bigotes (Box-Plot)
- 5.3 Diagrama de barras
- 5.4 Diagrama en escalera

5.1 HISTOGRAMA

Lo usamos para la representación gráfica de una variable agrupada en intervalos y con frecuencias absolutas o relativas, pero no acumuladas.

Se utiliza un diagrama cartesiano donde en el eje de abscisas se ponen los límites de los intervalos y en el de ordenadas las frecuencias, absolutas o relativas. Sobre los intervalos se levantan rectángulos que tienen por base la amplitud del intervalo y por altura su frecuencia o densidades de frecuencias en el caso de amplitudes distintas.



<https://commons.wikimedia.org/wiki/File:Histograma>

5.2 CAJAS Y BIGOTES (BOX-PLOT)

Un diagrama de cajas y bigote (También conocido también

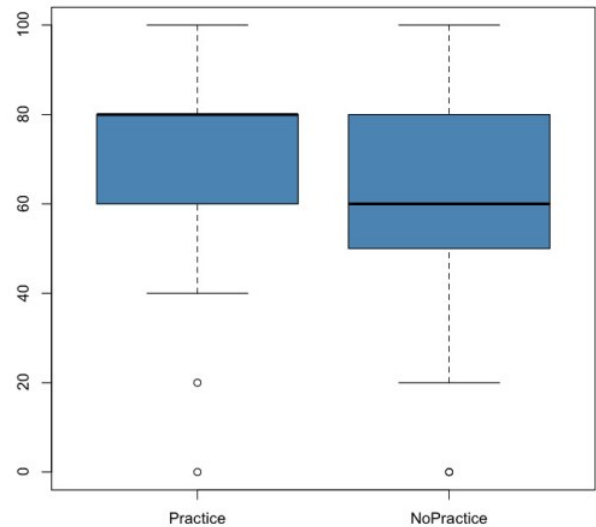
CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

EXISTENCIA DE VALORES ANÓMALOS DE UN

UD1 Estadística descriptiva unidimensional

conjunto de observaciones (*outliers*).



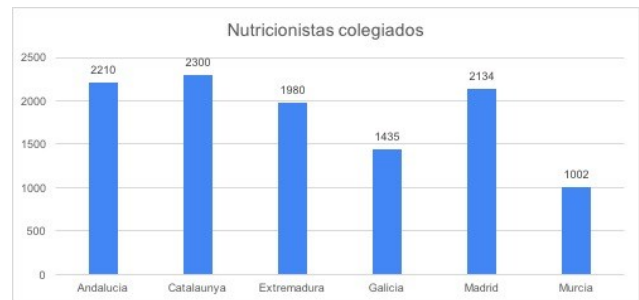
<https://live.staticflickr.com/8513/848551>
Cajas y Bigotes

5.3 DIAGRAMA DE BARRAS

Es importante destacar que se utiliza para representar frecuencias absolutas o relativas, pero **nunca** acumuladas.

Utilizamos un diagrama cartesiano donde en el eje de abscisas se ponen los valores de la variable y en el de ordenadas las frecuencias, absolutas o relativas, pero sin acumular.

En cada valor de la variable se eleva un segmento de altura igual a la frecuencia que le corresponde.



Elaboración propia
Diagrama de Barras

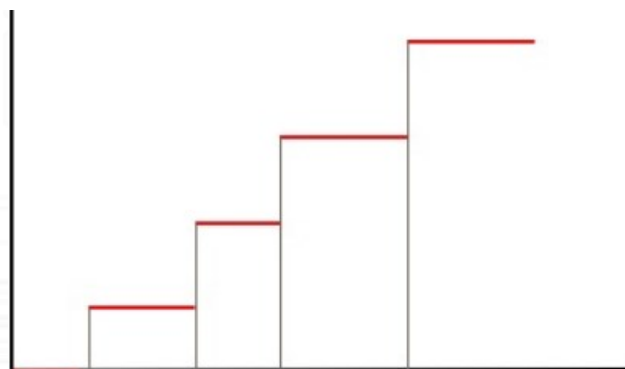
CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

abscisas se sitúan los valores de la variable y en el de ordenadas las frecuencias acumuladas (absolutas o relativas).

En cada valor de la variable se dibuja una línea horizontal (paralela al eje de abscisas) a la altura de su frecuencia acumulada correspondiente, que va hasta el siguiente valor de la variable.

A partir del último valor de la variable se prolonga una paralela al eje de abscisas hacia la derecha, a la altura de la última frecuencia acumulada, y a partir del menor valor de la variable se prolonga un trazo horizontal sobre el eje de abscisas hacia la izquierda, indicando que no hay acumulación de observaciones antes del menor valor ni después del mayor valor observado.



Elaboración propia
Histograma

CONCLUSIONES

- En esta unidad hemos visto como clasificar y ordenar los datos obtenidos.
- Entender qué son las variables estadísticas y cómo clasificarlas
- Conocer qué son las tablas de frecuencia, su utilidad y aplicabilidad.
- Conocer los diferentes estadísticos y su definición:
 - i) Posición,
 - ii) centralización,

**CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70**

**ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70**

En esta unidad didáctica se han tratado los siguientes puntos:

- Qué es la estadística descriptiva
- Qué son las variables estadísticas
- Conocer qué son las tablas de frecuencia, su utilidad y aplicabilidad.
- Conocer los diferentes estadísticos y su utilización a la hora de entender la distribución de los datos obtenidos. Sirviéndonos para hacer una primera descripción de los mismos.
- Cómo desarrollar las principales tipos de representaciones gráficas que nos ayuden a interpretar los datos obtenidos.

BIBLIOGRAFÍA

- Almiñana, P., & Daniel, V. (2016). Procedimientos de muestreo y preparación de la muestra. Síntesis.
- Nolasco Bonmatí A., Moncho Vasallo J. (2016). Estadística básica en Ciencias de la Salud. Universidad de Alicante.
- Ruiz Camacho M., Morcillo Aixelá M.C., García Galisteo J., Castillo Vázquez C. (2000). "curso de probabilidad y estadística", ed. universidad de Málaga.
- Zylberberg A.D. (2006). Probabilidad y estadística. Nueva Librería (Argentina).

The logo for Cartagena99 features the text 'Cartagena99' in a stylized, blue, serif font. The '99' is significantly larger and more prominent than the 'Cartagena' part. The text is set against a light blue background with a white arrow pointing to the right, and a yellow and orange gradient bar at the bottom.

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

UD1 Estadística descriptiva unidimensional