

Parte nº 7. Tema III-5.

Estadística descriptiva e inferencial aplicada al entorno cotidiano

ÍNDICE

- 1) INTRODUCCIÓN.
 - 2) CONCEPTOS
 - 3) ESTUDIO ESTADÍSTICO.
 - 3.1. Recogida de datos.
 - 3.2. Organización de los datos.
 - 3.2.1. Gráficas estadísticas.
 - 3.3. Análisis de datos.
 - 3.3.1. Medidas de centralización.
 - 3.3.2. Medidas de dispersión.
 - 4) Muestras. Estudios estadísticos.
-

1) INTRODUCCIÓN

La **estadística** es una ciencia matemática especializada en el **análisis de grandes volúmenes de información** para de ella extraer conclusiones. Tras analizar los datos deduce determinadas características de dicha información.

También se podría decir que la **estadística** trata del recuento, la ordenación y clasificación de datos obtenidos por las observaciones, para poder hacer comparaciones y sacar conclusiones.

Un **estudio estadístico** consta de las siguientes fases:

- Recogida de datos
- Organización y representación de datos
- Análisis de datos
- Obtención de conclusiones

2) CONCEPTOS

En un estudio estadístico distinguimos:

- ✓ **POBLACIÓN:** Es el conjunto de todos los elementos sobre los cuales se va a estudiar una determinada característica.
Por ejemplo: si vamos a analizar la estatura media de los españoles la población sería todos los ciudadanos españoles.
- ✓ **MUESTRA:** Del total de la población se selecciona un grupo representativo que es el que vamos a estudiar.
Por ejemplo: para la analizar la estatura media de los españoles no podemos recoger esta información de todos los ciudadanos españoles sino que tenemos definir un grupo de estudio, por ejemplo seleccionar a 2000 personas. Este grupo tiene que ser representativo de la sociedad española por lo que tiene que incluir a hombres y mujeres, gente de la ciudad y del campo, gente de diversos niveles de renta, de diversas edades. Es decir, la muestra tiene que ser como una imagen "en miniatura" de la población.
- ✓ **VARIABLE ESTADÍSTICA:** el aspecto que se va a estudiar. En nuestro ejemplo, la estatura media. Si se puede medir se llama **variable cuantitativa** (por ejemplo, altura y peso). Si no se pueden medir se llama **variable cualitativa** (por ejemplo, sexo).
Si la variable estadística toma un número determinado de valores se llama **variable discreta**.
Por ejemplo, números de años en el colegio: de 1 a 15.
Si la variable estadística puede tomar cualquier valor entre dos valores dados se llama **variable continua**. Aquí la variable puede tomar un número casi ilimitado de valores. Por ejemplo, estatura; 161 cm, 162 cm, 163 cm)
- ✓ **VALOR** es cada uno de los distintos resultados que se pueden obtener en un estudio estadístico.

Ejercicio 1

Clasifica los siguientes caracteres estadísticos según sean cualitativos, variables discretas o variables continuas:

- 1) Marca de los coches.
- 2) Peso de los coches.
- 3) Número de coches vendidos de las diferentes

Ejercicio 2

Indica cuál es la población de cada uno de los siguientes estudios estadísticos y si es conveniente tomar una muestra.

- 1) Altura y peso de los alumnos de una clase.
- 2) Marca de los coches de una ciudad.

3) ESTUDIO ESTADÍSTICO

Una vez definida las variables que vamos a estudiar y la muestra que vamos a analizar, hay que comentar por obtener la información. Para realizarlo, debemos seguir los siguientes pasos.

3.1) RECOGIDA DE DATOS

Planteado el test o encuesta oportuno, una vez elegido el tema al que se quiere hacer el estudio estadístico, y recogidos los datos que correspondan, el primer análisis que realizaremos es el del tipo de variable que pretendemos estudiar (**Cualitativa o Cuantitativa; Discreta o Continua**).

Esto condicionará en gran medida su posterior tratamiento.

3.2) ORGANIZACIÓN DE LOS DATOS

Determinado el modo de agrupamiento de las observaciones, procedemos a su **RECuento**. Los datos obtenidos en el punto anterior hay que ordenarlos y recogerlos en una tabla que se denomina **tabla estadística o tabal de frecuencias**.

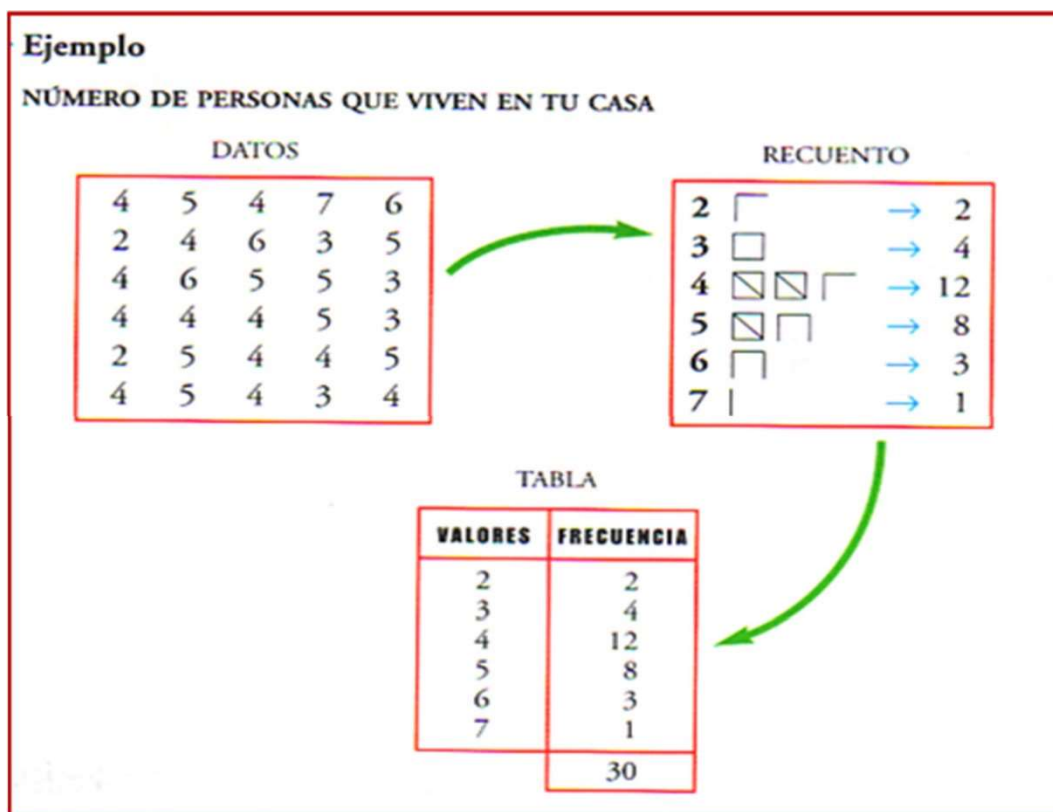


Imagen 1. Recogida de datos, recuento y tabla de frecuencias. Fuente: Desconocida. Autor: Desconocido. Licencia: Desconocida.

Posteriormente podremos visualizar tales frecuencias de forma gráfica con el diagrama estadístico apropiado.

TABLA DE FRECUENCIAS:

Tabla de frecuencias es una **ordenación** en forma de **tabla** de los **datos estadísticos**, asignando a cada **dato** su **frecuencia correspondiente**.

✓ **Frecuencia absoluta** es el **número de veces** que aparece un determinado **valor** en un estudio estadístico. Se representa por **f_i**.

La **suma de las frecuencias absolutas** es igual al número total de datos, que se representa por **N**. Es decir: **$N = \sum f_i$**

✓ **Frecuencia relativa** es el **cociente** entre la **frecuencia absoluta** de un determinado valor y el **número total de datos**.

Se representa por **n_i**.
$$n_i = \frac{f_i}{N}$$

La **suma de las frecuencias relativas** es igual a **1**.

✓ **Frecuencia relativa porcentual** es la frecuencia relativa multiplicada por 100. Se puede expresar por %

La **suma de las frecuencias relativas porcentuales** es **100**.

✓ **Frecuencias acumuladas** es la **suma de las frecuencias correspondientes** de todos los **valores inferiores o iguales** al **valor** considerado.

Se representa por:

F_i la frecuencia absoluta acumulada

N_i la frecuencia relativa acumulada

N_i% la frecuencia relativa porcentual acumulada

Para que **tengan sentido las frecuencias acumuladas**, los **valores de la variable deben estar ordenados**.

DATOS X_i	Frecuencia absoluta f_i	Frecuencia absoluta acumulada F_i	Frecuencia relativa		Frecuencia relativa acumulada	
			Decimal	Porcentual $\%_i = 100 \times n_i$	Decimal N_i	Porcentual $N_i\%$
5	6	6	0,0150	1,5 %	0,0150	1,5 %
6	48	54	0,1200	12 %	0,1350	13,5 %
7	95	149	0,2375	23,75 %	0,3725	37,25 %
8	105	254	0,2625	26,25 %	0,6350	63,50 %
9	87	341	0,2175	21,75 %	0,8525	85,25 %
10	59	400	0,1475	14,75 %	1	100 %
	$N = \sum f_i$ $N = 400$		$\sum n_i = 1$	$\sum \%_i = 100\%$		

Imagen 2. Tabla de frecuencias con variable cuantitativa.
Autor: Desconocido. Fuente: Desconocida. Licencia: Desconocida.

Ejercicio 3

El número de veces que han ido al cine en el último mes los alumnos de una clase es:

{2,3,0,1,5,3,2,1,0,0,2,1,2,3,5,0,5,4,1,1,1,2,0,1,2}

Forma la tabla de frecuencias absolutas y relativas, y las acumuladas.

Ejercicio 4

Se realiza un trabajo en la asignatura de lengua en una clase formada por 40 alumnos. 2 alumnos realizan el trabajo en un folio, 5 en 2 folios, 6 en 3 folios, y el resto en 4 folios. Forma la tabla de frecuencias absolutas y relativas. ¿Crees que el profesor ha recomendado un número determinado de folios?

Ejercicio 5

Los goles que se han marcado en la última jornada de liga han sido en los siguientes minutos de juego:

{20,11,89,3,20,4,2,35,50,29,59,30,90,33,78,54,21,19,60,34,56,63,45,31,26,32,5,78,88,85,34}

Realiza la tabla de frecuencias absolutas y relativas agrupándolos en clase por cuarto de hora.

3.2.1) GRÁFICAS ESTADÍSTICAS

Las gráficas estadísticas permiten visualizar la información contenida en las tablas de manera rápida y sencilla.

Existen muchos tipos de gráficas estadísticas. Unas se emplean con variables cuantitativas y otras con variables cualitativas.

DIAGRAMA DE BARRAS:

Se utiliza para presentar **datos cualitativos** o **datos cuantitativos de tipo discreto**.

Se representan sobre unos ejes de coordenadas, en el **eje de abscisas** se colocan los **valores de la variable**, y sobre el **eje de ordenadas** las **frecuencias absolutas, relativas, porcentajes o frecuencias acumuladas**.

Los **datos** se representan mediante **barras** de una **altura proporcional** a la **frecuencia**.

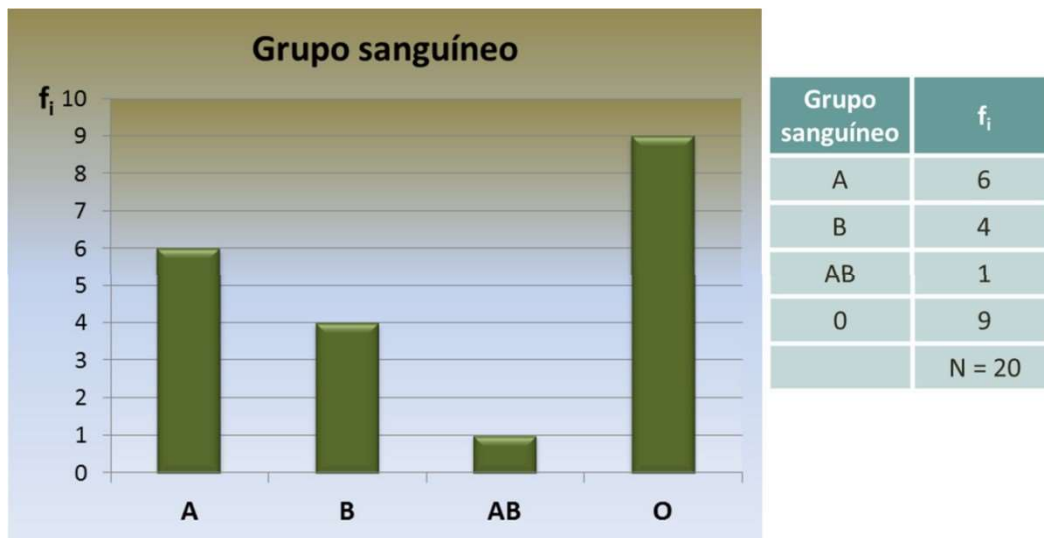


Imagen 3. Ejemplo de diagrama de barras (1/2). Autor: Desconocido. Fuente: Desconocida. Licencia: Desconocida.

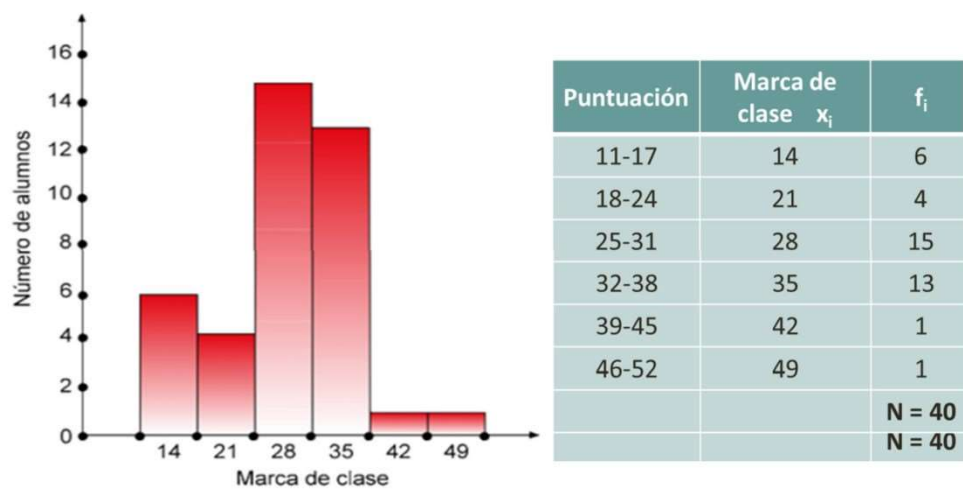


Imagen 4. Ejemplo de diagrama de barras (2/2). Autor: Desconocido. Fuente: Desconocida. Licencia: Desconocida

POLÍGONO DE FRECUENCIAS:

Se realiza para cualquier tipo de variable. Es el polígono que se forma al unir los puntos medios de las barras tanto en histogramas como en diagramas de barras.

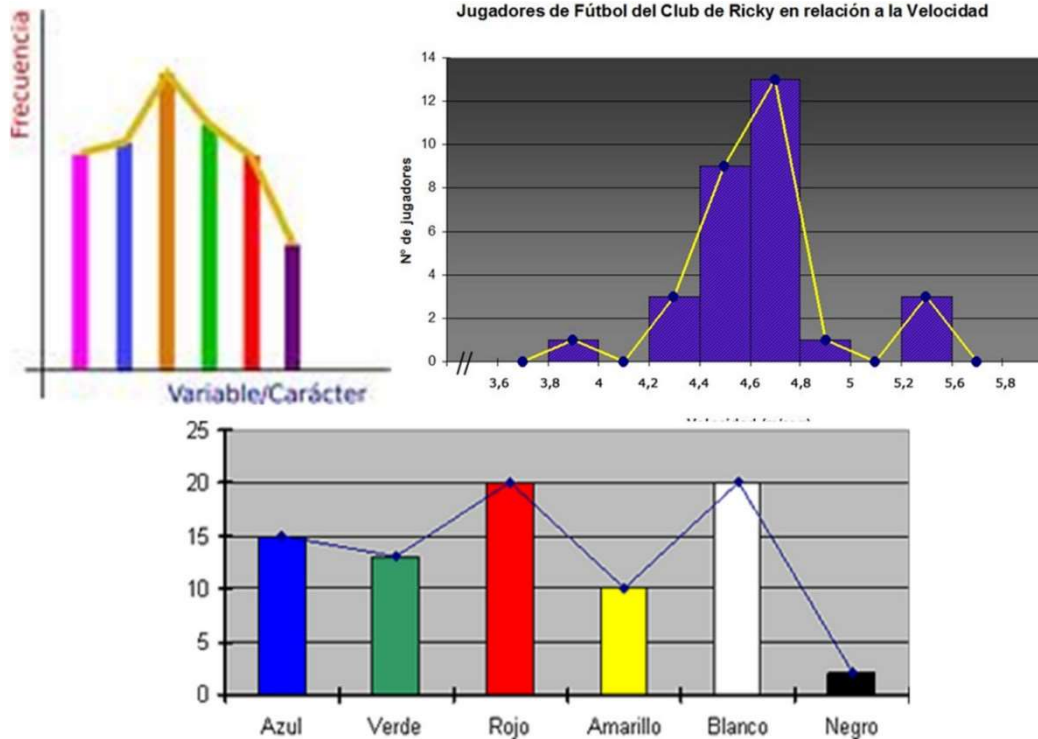


Imagen 5. Ejemplo de polígono de frecuencias. Autor: Desconocido.
Fuente: Desconocida. Licencia: Desconocida.

DIAGRAMA DE SECTORES: Es un gráfico donde se suele representar los porcentajes.

Cada sector es proporcional al porcentaje que representa.
Los grados de cada sector son:

Grados = 360 x ni

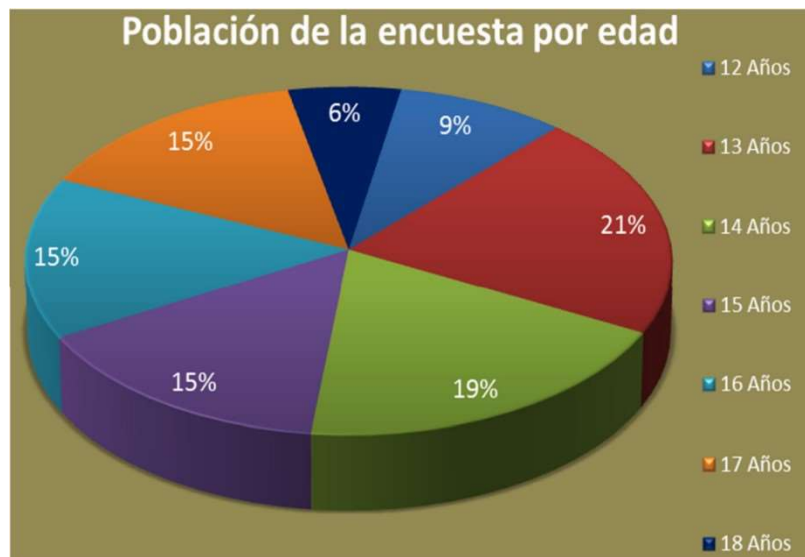


Imagen 6. Ejemplo de diagrama de sectores. Autor: Desconocido.
Fuente: Desconocida. Licencia: Desconocida.

PICTOGRAMA: Es un gráfico con figuras.

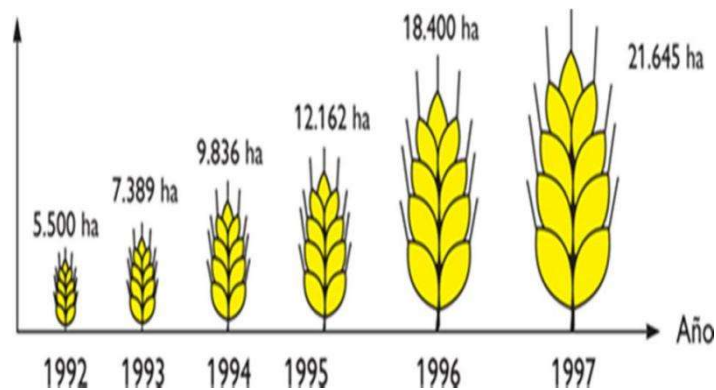
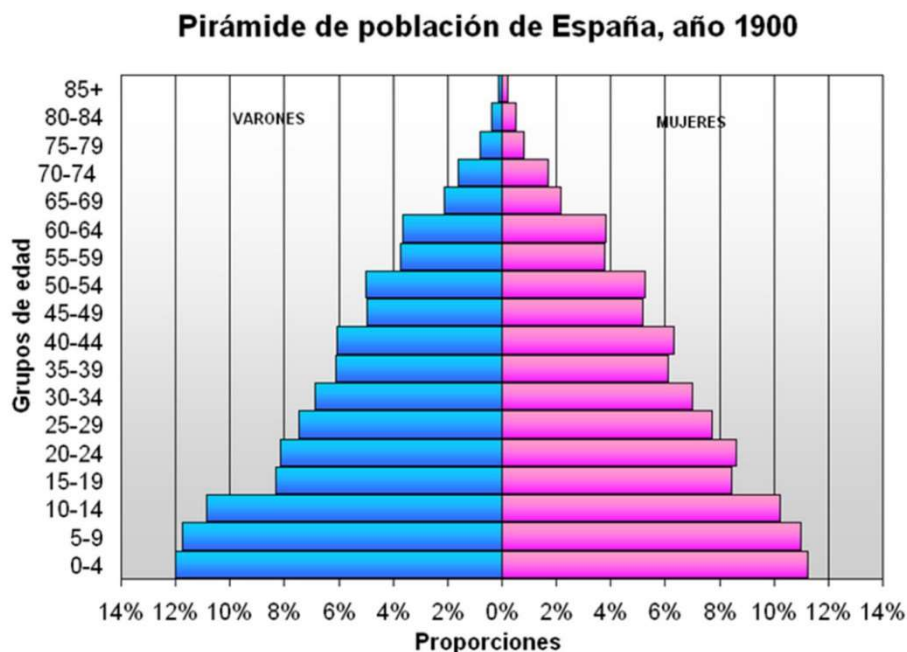


Imagen 7. Ejemplo de pictograma. Autor: Desconocido.
Fuente: Desconocida. Licencia: Desconocida.

PIRÁMIDE DE POBLACIÓN:

Consiste en dos histogramas, uno para hombres y otro para mujeres, correspondientes a habitantes de una misma comunidad más o menos extensa, repartidos por edades.

Es útil para estudiar su situación demográfica y buscar explicaciones a situaciones presentes, pasadas y futuras.



Fuente: Instituto Nacional de Estadística. Censo de 1900

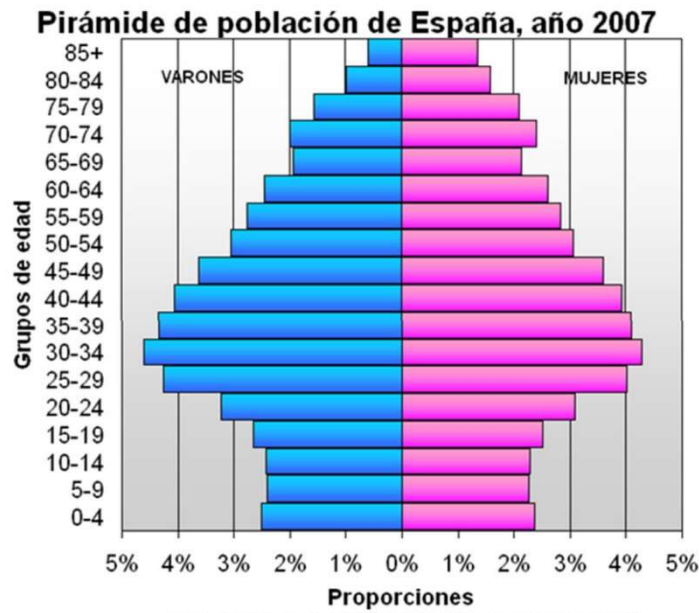
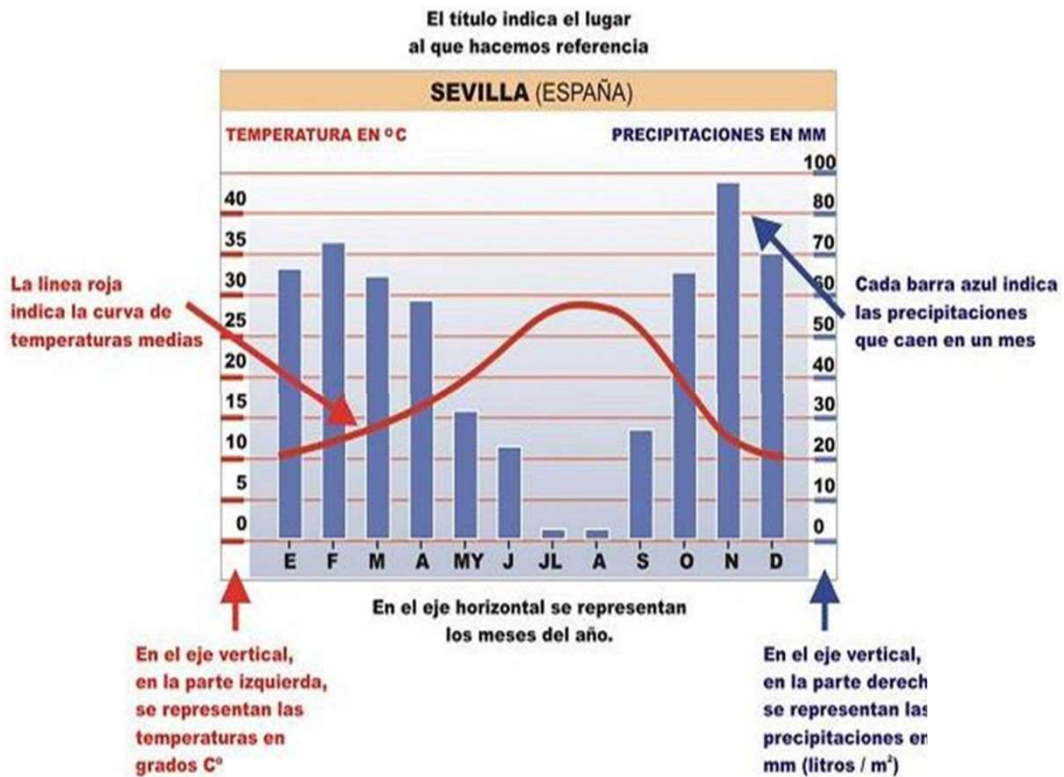


Imagen 8. Ejemplo de pirámide de población. Autor: Desconocido. Fuente: Desconocida. Licencia: Desconocida.

CLIMOGRAMA:

Son gráficas que representan la distribución de precipitaciones y temperaturas a largo de un año en un lugar determinado.



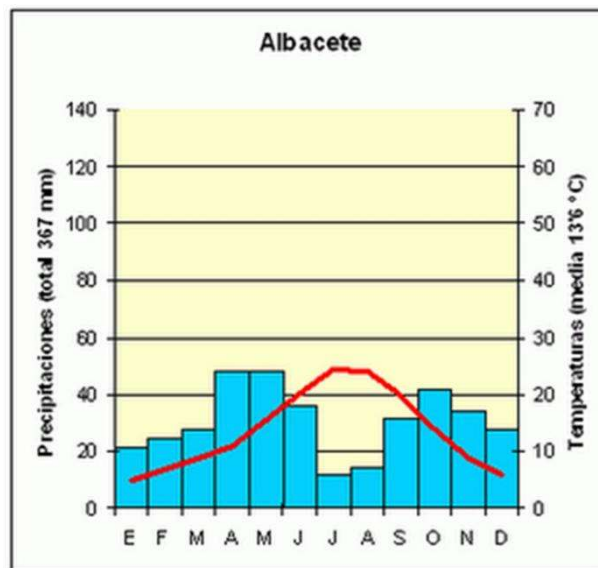


Imagen 9. Ejemplo de climograma. Autor: Desconocido.
Fuente: Desconocida. Licencia: Desconocida.

Ejercicio 6

Realiza un diagrama de sectores para los siguientes datos:

Autonomía	Nº de centros
Andalucía	30
Asturias	27
Cataluña	43
Galicia	25
Madrid	40
Navarra	15

Ejercicio 7

El número de veces que han ido al cine durante el último mes los habitantes de un pueblo es:

Nº veces	fi
0	15
1	26
2	32
3	20
4	15
5 o más	8

Realiza el polígono de frecuencias para estos datos.

Ejercicio 8

En una clase de hemos preguntado a los alumnos por las horas de estudio que dedican a la semana. Estas han sido las respuestas:

16	11	17	12	10	5	1	8	10	14
15	20	3	2	5	12	7	6	3	9
10	8	10	6	16	16	10	3	4	12

- 1) Ordena los datos en una tabla de frecuencias, agrupándolos en intervalos de la forma que creas más adecuada.**
- 2) Representa gráficamente la distribución.**

3.3) ANÁLISIS DE DATOS

Para este análisis se utilizan los parámetros estadísticos. Son los siguientes:

a) Medidas de centralización:

MEDIA, MEDIANA y MODA

b) Medidas de dispersión:

RECORRIDO, DESVIACIÓN MEDIA, VARIANZA, DESVIACIÓN TÍPICA, COEFICIENTE de VARIACIÓN.

3.3.1) MEDIDAS DE CENTRALIZACIÓN

MEDIA ARITMÉTICA:

Media aritmética es el **valor** obtenido al **sumar** todos los **datos** y **dividir** el resultado entre el **número** total de **datos**.

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{N} \qquad \bar{x} = \frac{\sum_{i=1}^n x_i}{N}$$

Imagen 10. Cálculo de media aritmética (1/2). Autor: Desconocido.

Fuente: Desconocida. Licencia: Desconocida.

Si los **datos** vienen en una tabla de frecuencias, la expresión de la **media** es:

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + x_3 f_3 + \dots + x_n f_n}{N} \qquad \bar{x} = \frac{\sum_{i=1}^n x_i f_i}{N}$$

Imagen 11. Cálculo de media aritmética (2/2). Autor: Desconocido.

Fuente: Desconocida. Licencia: Desconocida.

Evidentemente **esta medida sólo se puede hallar para variables cuantitativas**.

MODA: es el **valor** que tiene mayor frecuencia absoluta.

Se representa por **Mo**.

Se puede hallar para cualquier tipo de variable, aunque para variables cuantitativas es poco útil.

La **moda** de la distribución:

2, 3, 3, 4, 4, 4, 5, 5 es **Mo= 4**

Si en un grupo hay **dos o varias puntuaciones** con la **misma frecuencia** y esa frecuencia es la máxima, la **distribución es bimodal o multimodal**, es decir, tiene **varias modas**.

1, 1, 1, 4, 4, 5, 5, 5, 7, 8, 9, 9, 9 **Mo= 1, 5, 9**

MEDIANA: es el **valor** que ocupa el **lugar central** de todos los **datos** cuando éstos están **ordenados de menor a mayor**.

La **mediana** se representa por **Me**.

La **mediana** se puede **hallar** sólo para **variables cuantitativas**.

Cálculo de la mediana con pocos datos

- 1) **Ordenamos los datos de menor a mayor.**
- 2) Si la serie tiene un **número impar de medidas** la **mediana** es la **puntuación central** de la misma.

2, 3, 4, 4, 5, 5, 5, 6, 6 **Me= 5**

- 3) Si la serie tiene un **número par** de puntuaciones la **mediana** es la **media** entre las dos **puntuaciones centrales**.

7, 8, 9, 10, 11, 12 **Me= 9.5**

Ejemplo de cálculo de media:

En un test realizado a un grupo de 42 personas se han obtenido las puntuaciones que muestra la tabla. Calcula la puntuación media.

x_i	f_i	$x_i \cdot f_i$
15	1	15
25	8	200
35	10	350
45	9	405
55	8	440
65	4	260
75	2	150
	42	1 820

$$\bar{x} = \frac{1820}{42} = 43.33$$

Imagen 12. Ejemplo cálculo de media. Autor: Desconocido.
Fuente: Desconocida. Licencia: Desconocida.

Cálculo de la mediana para muchos datos

x_i	f_i	F_i
60	7	7
66	18	25
70	44	69
72	27	96
75	8	104
	N = 104	

Imagen 13. Ejemplo de cálculo mediana con muchos datos. Autor: Desconocido.
Fuente: Desconocida. Licencia: Desconocida.

Se divide **N** entre dos para ver dónde está el centro

$$104/2 = 52$$

Se busca en la columna de **Fi**

donde estaría 52.

Luego el valor o intervalo mediano será:

70

Ejercicio 9

Las notas de inglés de una clase de 40 alumnos han sido las siguientes:

1	7	9	2	5	4	4	3	7	8
4	5	6	7	6	4	3	1	5	9
2	6	4	6	5	2	2	8	3	6
4	5	2	4	3	5	6	5	2	4

Calcula la nota media.

3.3.2) MEDIDAS DE DISPERSIÓN

Las **medidas de dispersión** sirven para comparar dos o más distribuciones y decidir cuál de ellas es más o menos dispersa.

RECORRIDO O RANGO es la diferencia entre los valores extremos, es decir, entre el mayor valor y el menor.

$$\text{Recorrido (Rango)} = \text{Valor mayor} - \text{Menor valor}$$

DESVIACIÓN MEDIA es un parámetro asociado a la media; y es el promedio (o media) de las distancias de los valores de todos los individuos a la media.

$$DM = \frac{\text{suma de las distancias a } \bar{x}}{N}$$

Para muchos valores \longrightarrow $DM = \frac{\sum |x_i - \bar{x}| \cdot f_i}{N}$

VARIANZA:

Sirve para identificar si los datos están cercanos a la media o no, se calcula sumando los valores que se obtienen de elevar al cuadrado la diferencia de cada dato con la media, y dividiendo este valor entre el número de datos, para representar este parámetro se utiliza el símbolo σ^2 .

$$\sigma^2 = \sum \frac{x_i^2 \cdot f_i}{N} - \bar{x}^2$$

DESVIACIÓN TÍPICA:

Da un valor de las diferencias de los valores con respecto a la Media que se obtiene haciendo la raíz cuadrada de la varianza.

$$\sigma = \sqrt{\sum \frac{x_i^2 \cdot f_i}{N} - \bar{x}^2}$$

COEFICIENTE DE VARIACIÓN:

Se usa para comparar las dispersiones de dos distribuciones heterogéneas.

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

Al calcular el coeficiente de variación *estamos relativizando la dispersión*. El resultado se da, a veces, en tantos por ciento.

Cálculo de desviación media para pocos valores:

Distribución 5, 7, 8, 9, 11, 13, 13, 15, 16, 18.

$$\bar{x} = \frac{5+7+8+9+11+13+13+15+16+18}{10} = 11,5$$

Datos	5	7	8	9	11	13	13	15	16	18	
Distancia a	6,5	4,5	3,5	2,5	0,5	1,5	1,5	3,5	4,5	6,5	SUMA → 35

$$DM = \frac{\text{suma de las distancias a } \bar{x}}{N} = \frac{35}{10} = 3,5$$

Cálculo de desviación media para muchos valores:

x_i	f_i	$x \cdot f_i$	Distancia a	$f_i \cdot \text{distancia}$
1	5	5	$2,5 - 1 = 1,5$	7,5
2	15	30	$2,5 - 2 = 0,5$	7,5
3	11	33	$3 - 2,5 = 0,5$	5,5
4	4	16	$4 - 2,5 = 1,5$	6
6	1	6	$6 - 2,5 = 3,5$	3,5
	N= 36	$\Sigma = 90$		$\Sigma = 30$

$$\bar{x} = \frac{90}{36} = 2,5$$

$$DM = \frac{\sum |x_i - \bar{x}| \times f_i}{N} = \frac{30}{36} = 0,83$$

Obtener las medidas de dispersión de la siguiente distribución de notas: 2, 4, 4, 4, 5, 7, 9, 9, 10.

RECORRIDO: $10 - 2 = 8$ MEDIA: $\bar{x} = 6$

DESVIACIÓN MEDIA: $DM = \frac{|2-6| + |4-6| + |4-6| + \dots}{9} = \frac{22}{9} = 2,44$

VARIANZA: $Var = \frac{(2-6)^2 + (4-6)^2 + (4-6)^2 + \dots}{9} = \frac{64}{9} = 7,11$

o bien: $Var = \frac{2^2 + 4^2 + 4^2 + 4^2 + 5^2 + \dots}{9} - 6^2 = \frac{388}{9} - 36 = 7,11$

DESVIACIÓN TÍPICA: $\sigma = \sqrt{\text{varianza}} = \sqrt{7,11} = 2,67$

Ejercicio 10

Calcula la varianza y la desviación típica de los siguientes datos:

4, 7, 5, 3, 6.

Ejercicio 11

En un examen de matemáticas los 30 alumnos de una clase han obtenido las puntuaciones recogidas en la siguiente tabla:

Calificaciones	Nº alumnos
[0,1)	2
[1,2)	2
[2,3)	3
[3,4)	6
[4,5)	7
[5,6)	6
[6,7)	1
[7,8)	1
[8,9)	1
[9,10)	1

Halla la varianza y la desviación típica.

4) Muestras. Estudios estadísticos

Si queremos hacer un estudio estadístico tenemos que:

- Recoger los datos
- Describir esos datos con tablas y gráficas, cálculo de parámetros estadísticos....

c) Extraer conclusiones.

Para recoger los datos y determinar los valores de la variable se puede utilizar a toda la población, todo el universo sobre el que se realiza el estudio, o hacer una muestra. En muchas ocasiones no es conveniente recoger valores de toda la población, porque es complicado o demasiado costoso, o incluso porque es imposible como en el caso de un control de calidad en que se destruya el objeto a analizar.

La parte de la Estadística que se ocupa de cómo seleccionar adecuadamente las muestras se denomina Teoría de Muestras, cuyos conceptos básicos son:

- Población o universo es todo el conjunto de individuos sobre el que se realiza el estudio.
- Una muestra es un subconjunto representativo de esa población.
- Cada uno de los elementos de la población es un individuo.

Las características de la población que se estudian se denominan variables estadísticas, que se clasifican en cuantitativas y cualitativas según que los valores que tomen sean o no numéricos.

- Las variables cuantitativas que toman valores aislados se denominan variables discretas, un ejemplo es el número de calzado.
- Las variables que pueden tomar cualquier valor de un intervalo de la recta real, se denominan variables continuas, por ejemplo la altura de una persona.

La parte de la Estadística que ordena, analiza y representa un conjunto de datos para describir sus características se denomina Estadística Descriptiva.

Para extraer conclusiones se utilizan las probabilidades y la parte de la Estadística que se ocupa de ello es la Inferencia Estadística.

Ejemplos:

Si queremos conocer las preferencias en deportes de las personas adultas, es posible preguntar a toda la población (persona mayor de 18 años), pero este sería un trabajo muy costoso de realizar, por lo que, parece adecuado elegir una muestra representativa, seleccionando a algunas personas adultas. En este estudio sobre preferencias deportivas, la variable utilizada es cualitativa, porque la respuesta no será un número.

Para conocer la intención de voto ante unas elecciones europeas, municipales, autonómicas... se utilizan muestras, pues preguntar a toda la población sería muy costoso (esto se hace en las elecciones). La variable en este caso también es cualitativa.

Si una fábrica de bombillas quiere conocer las horas de vida útil de una bombilla, no puede encender todas las bombillas que fabrique para comprobar su duración, eso sería la ruina de la fábrica. En este caso es imprescindible seleccionar una muestra, es decir, elegir de manera aleatoria una cantidad de bombillas, a esta cantidad de bombilla elegida se le denomina tamaño de la muestra. La variable en este caso es cuantitativa, y el tiempo toma cualquier valor, por lo que es una variable cuantitativa continua.

Si en una ciudad queremos hacer un estudio sobre el número de hermanos que hay en una familia, se toma una muestra. Para ello se define el tamaño de la muestra, se eligen familias al azar y se les pregunta por el número de hermanos. En este caso se trata de una variable cuantitativa discreta.

En todos estos ejemplos, una vez obtenidos los datos de la muestra, se realiza su estudio,

utilizando las técnicas de la Estadística descriptiva, se calculan sus parámetros, y en muchos casos, estos valores se pueden extrapolar de la muestra a la población, siempre considerando la existencia de un error, que, en caso de ser asumible, permite tomar decisiones:

- En el caso de las preferencias deportivas, un ayuntamiento puede decidir la oferta de actividades para un año.
- En el caso de la intención de voto se avanzan los posibles resultados electorales, ayudando a los partidos políticos a definir sus estrategias de campaña.
- En el caso de la fábrica, permite decidir si se debe invertir en estudios para mejorar la durabilidad de las bombillas, o, en caso de ser una durabilidad aceptable, publicitar el resultado para mejorar las ventas.
- En el caso del número de hermanos, un supermercado puede decidir hacer ofertas en sus productos a familias en función del número de hermanos, eligiendo que este número sea dos, tres o cuatro, dependiendo del resultado del estudio.

Ejercicio 12

1. Queremos realizar un estudio estadístico sobre el tiempo dedicado al estudio por el alumnado de enseñanzas para adultos en Castilla la Mancha. Para ello se seleccionan adecuadamente 80 alumnos. Indica cuál es la población, cuál la muestra, qué tamaño tiene la muestra y quién sería un individuo.

2. Quieres pasar una encuesta para conocer, lo mismo que en el problema anterior, el tiempo dedicado al estudio, en este caso el de los compañeros y compañeras de tu centro escolar. ¿Se la pasarías sólo a las chicas? ¿Sólo a los chicos? ¿Preguntarías a los mejores de la clase? ¿A los de peores notas? Indica el criterio que seguirías para seleccionar la muestra a la que preguntar.

Ejercicios resueltos

Ejercicio 1

Clasifica los siguientes caracteres estadísticos según sean cualitativos, variables discretas o variables continuas:

- | | |
|--|--------------------------|
| 1) Marca de los coches. | <u>Cualitativo</u> |
| 2) Peso de los coches. | <u>Variable discreta</u> |
| 3) Número de coches vendidos de las diferentes | <u>Variable continua</u> |

Ejercicio 2

Indica cuál es la población de cada uno de los siguientes estudios estadísticos y si es conveniente tomar una muestra.

- 1) Altura y peso de los alumnos de una clase.
- 2) Marca de los coches de una ciudad.

- 1) La población son los alumnos de la clase, no es necesario realizar una muestra.
- 2) Los coches de la ciudad, es necesario realizar una muestra.

Ejercicio 3

El número de veces que han ido al cine en el último mes los alumnos de una clase es:

{2,3,0,1,5,3,2,1,0,0,2,1,2,3,5,0,5,4,1,1,1,2,0,1,2}

Forma la tabla de frecuencias absolutas y relativas, y las acumuladas.

xi	ni	fi	Ni	Fi
0	5	0,20	5	0,20
1	7	0,28	12	0,48
2	6	0,24	18	0,72
3	3	0,12	21	0,84
4	1	0,04	22	0,88
5	3	0,12	25	1
Σ	25	1		

Ejercicio 4

Se realiza un trabajo en la asignatura de lengua en una clase formada por 40 alumnos. 2 alumnos realizan el trabajo en un folio, 5 en 2 folios, 6 en 3 folios, y el resto en 4 folios. Forma la tabla de frecuencias absolutas y relativas. ¿Crees que el profesor ha recomendado un número determinado de folios?

Variable	Frecuencia absoluta	Frecuencia relativa
1	2	0,050
2	5	0,125
3	6	0,150
4	27	0,675
Σ	40	1

Ejercicio 5

Los goles que se han marcado en la última jornada de liga han sido en los siguientes minutos de juego:

{20,11,89,3,20,4,2,35,50,29,59,30,90,33,78,54,21,19,60,34,56,63,45,31,26,32,5,78,88,85,34}

Realiza la tabla de frecuencias absolutas y relativas agrupándolos en clase por cuarto de hora.

Clase	Frecuencia absoluta	Frecuencia relativa
[0,15)	5	0,161
[15,30)	6	0,194
[30,45)	7	0,223
[45,60)	5	0,161
[60,75)	2	0,065
[75,90]	6	0,194
Σ	31	1

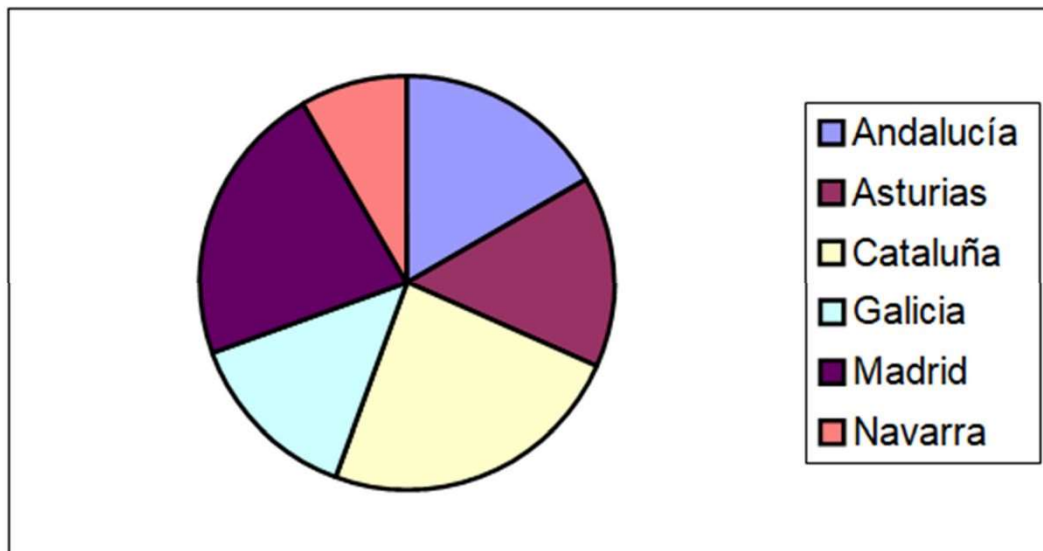
Ejercicio 6

Realiza un diagrama de sectores para los siguientes datos:

Autonomía	Nº de centros	Porcentaje	Ángulo central
Andalucía	30	16,7 %	60°
Asturias	27	15,0 %	54°
Cataluña	43	23,9 %	86°
Galicia	25	13,9 %	50°
Madrid	40	22,2 %	80°
Navarra	15	8,3 %	30°
Total Σ	180	1	360

El porcentaje de Galicia se ha calculado: $f_i \cdot 100 = (25/180) \cdot 100 = 13,9\%$ El ángulo central correspondiente a Galicia es: $f_i \cdot 360 = (25/180) \cdot 360 = 50^\circ$.

Por tanto, el diagrama de sectores es:

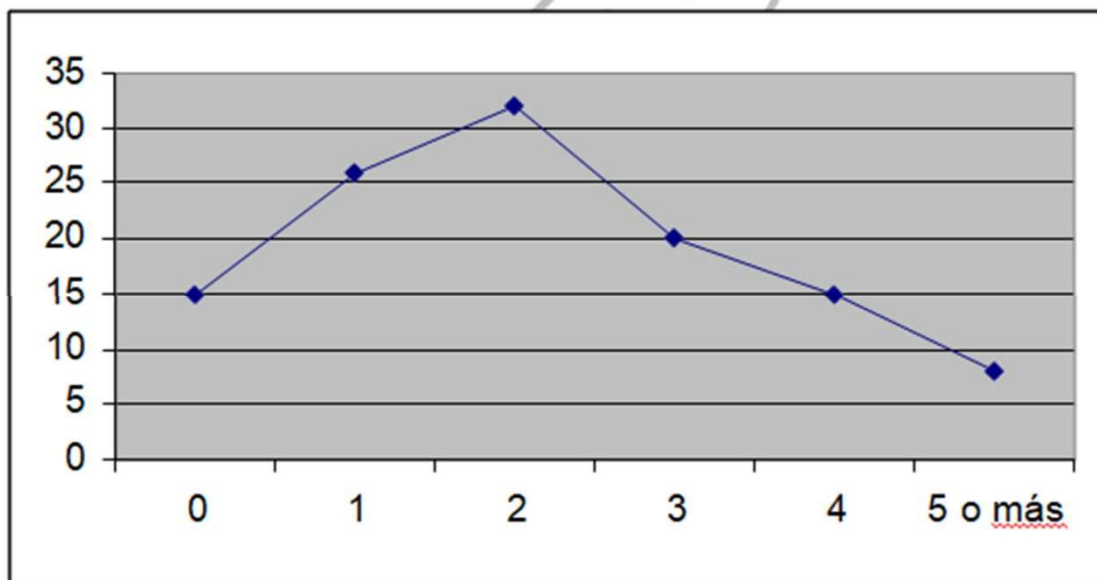


Ejercicio 7

El número de veces que han ido al cine durante el último mes los habitantes de un pueblo es:

Nº veces	fi
0	15
1	26
2	32
3	20
4	15
5 o más	8

Realiza el polígono de frecuencias para estos datos.



Ejercicio 8

En una clase de hemos preguntado a los alumnos por las horas de estudio que dedican a la semana. Estas han sido las respuestas:

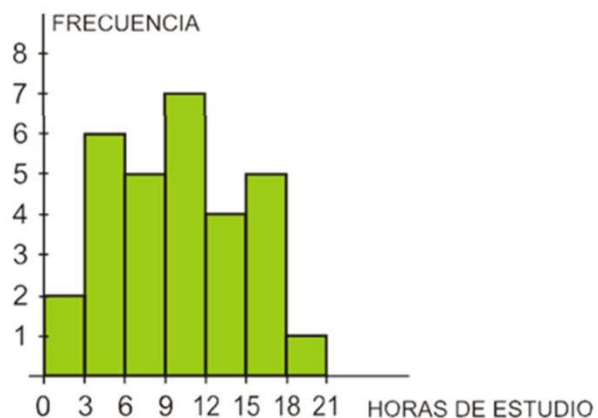
16	11	17	12	10	5	1	8	10	14
15	20	3	2	5	12	7	6	3	9
10	8	10	6	16	16	10	3	4	12

- 1) Ordena los datos en una tabla de frecuencias, agrupándolos en intervalos de la forma que creas más adecuada.
- 2) Representa gráficamente la distribución.

Por una parte, la variable que estamos estudiando (horas de estudio) es continua. Además, entre los datos que tenemos hay una gran variedad. Por tanto, debemos agrupar los datos en intervalos.

El menor valor es 1 y el mayor es 20; su diferencia es $20 - 1 = 19$. Por tanto, podemos tomar 7 intervalos de longitud 3, empezando en 0:

INTERVALO	FRECUENCIA
[0, 3)	2
[3, 6)	6
[6, 9)	5
[9, 12)	7
[12, 15)	4
[15, 18)	5
[18, 21)	1
	30



Ejercicio 9

Las notas de inglés de una clase de 40 alumnos han sido las siguientes:

1	7	9	2	5	4	4	3	7	8
4	5	6	7	6	4	3	1	5	9
2	6	4	6	5	2	2	8	3	6
4	5	2	4	3	5	6	5	2	4

Calcula la nota media.

Solución: 4.6

Ejercicio 10

Calcula la varianza y la desviación típica de los siguientes datos:

4, 7, 5, 3, 6.

Varianza = 2

Desviación típica = 1,41.

Ejercicio 11

En un examen de matemáticas los 30 alumnos de una clase han obtenido las

puntuaciones recogidas en la siguiente tabla:

Calificaciones	Nº alumnos
[0,1)	2
[1,2)	2
[2,3)	3
[3,4)	6
[4,5)	7
[5,6)	6
[6,7)	1
[7,8)	1
[8,9)	1
[9,10)	1

Halla la varianza y la desviación típica.

Solución: Varianza = 4,23 Desviación típica = 2,06.

Ejercicio 12

1. Queremos realizar un estudio estadístico sobre el tiempo dedicado al estudio por el alumnado de enseñanzas para adultos en Castilla la Mancha. Para ello se seleccionan adecuadamente 80 alumnos. Indica cuál es la población, cuál la muestra, qué tamaño tiene la muestra y quién sería un individuo.

Solución:

La población es “todas las personas matriculadas en una enseñanza para adultos”.

La muestra será un grupo de personas matriculadas seleccionadas al azar.

El tamaño de la muestra es de 80 personas

Individuo es cualquier persona matriculada en una enseñanza para adultos

2. Quieres pasar una encuesta para conocer, lo mismo que en el problema anterior, el tiempo dedicado al estudio, en este caso el de los compañeros y compañeras de tu centro escolar. ¿Se la pasarías sólo a las chicas? ¿Sólo a los chicos? ¿Preguntarías a los mejores de la clase? ¿A los de peores notas? Indica el criterio que seguirías para seleccionar la muestra a la que preguntar

Solución:

Se eligen alumnos y alumnas, de manera que haya personas matriculadas en los diferentes módulos, suponiendo que el tamaño de la muestra fuese de 40 personas, tomaremos 10 personas por nivel, y dentro de cada nivel, se realiza un sorteo para elegir a los componentes de la muestra.