ESTADÍSTICA

Población es el conjunto de todos los elementos objeto de nuestro estudio.
Muestra es un subconjunto, extraído de la población, cuyo estudio sirve para inferir características de toda la población.
Individuo es cada uno de los elementos que forman la población o la muestra.
Variable (**x_i**) es una característica que al ser medida en diferentes *individuos* es susceptible de adoptar diferentes valores. Puede ser cualitativa o cuantitativa (discreta o continua).
Frecuencia absoluta (**f_i**) es el número de individuos correspondiente a cada valor de la variable.
Frecuencia relativa es el cociente entre la frecuencia absoluta y el número total de individuos, multiplicado por 100 (si se da en tanto por ciento).
Ejemplo: notas obtenidas por un número de alumnos.
9, 4, 8, 5, 5, 4, 1, 7, 2, 2, 3, 9, 6, 4, 10, 8, 2, 1, 6, 7, 6, 10, 10, 8, 8, 4, 6, 5, 5, 10, 6, 7, 2, 5, 5, 3, 5, 3, 6 y 8.

TABLA DE FRECUENCIAS

x_i	f_i	$f_{relativa}$
1	2	$\frac{2}{40}\cdot 100 = 5\%$
2	4	$\frac{4}{40}\cdot 100 = 10\%$
3	3	$\frac{3}{40}\cdot 100 = 7{,}5\%$
4	4	$\frac{4}{40}\cdot 100 = 10\%$
5	7	$\frac{7}{40}\cdot 100 = 17{,}5\%$
6	6	$\frac{6}{40}\cdot 100 = 15\%$
7	3	$\frac{3}{40}\cdot 100 = 7{,}5\%$
8	5	$\frac{5}{40}\cdot 100 = 12{,}5\%$
9	2	$\frac{2}{40}\cdot 100 = 5\%$
10	4	$\frac{4}{40}\cdot 100 = 10\%$
	40	100%

Gráficos: barras, histograma, polígono y sectores.
Media ($\bar{x}$). Se calcula con la fórmula:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum x_i}{n}$$

Mediana (M_e). Es el valor que está en medio, es decir, tiene tantos individuos por debajo como por encima de él.
Moda (M_o). Es el valor con mayor frecuencia.
Rango. Diferencia entre el dato mayor y el menor.
Desviación media (DM). Es el promedio de las distancias de los daos a la media:

$$DM = \frac{|x_1 - \bar{x}| + |x_2 - \bar{x}| + \dots + |x_n - \bar{x}|}{n} = \frac{\sum |x_i - \bar{x}|}{n}$$

Varianza. Es el promedio de los cuadrados de las distancias de los datos a la media:

$$Varianza = \frac{|x_1 - \bar{x}|^2 + |x_2 - \bar{x}|^2 + \dots + |x_n - \bar{x}|^2}{n} = \frac{\sum (x_i - \bar{x})^2}{n} = \frac{\sum x_i^2}{n} - \bar{x}^2$$

Desviación típica (σ). Es la raíz cuadrada de la varianza.
Coeficiente de variación. Se define como:

$$CV = \frac{\sigma}{\bar{x}}$$ (se suele dar en tanto por ciento)

Utilizándose las tablas de frecuencia se llega a la conclusión:

Media	Desviación típica
$\bar{x} = \frac{f_1 x_1 + f_2 x_2 + \dots + f_n x_n}{\sum f_i} = \frac{\sum f_i x_i}{\sum f_i}$	$\sigma = \sqrt{\frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i}} = \sqrt{\frac{\sum f_i x_i^2}{\sum f_i} - \bar{x}^2}$

Volviendo al ejemplo:

x_i	f_i	$f_i \cdot x_i$	$f_i \cdot x_i^2$
1	2	1·2 = 2	$2 \cdot 1^2 = 2$
2	4	2·4 = 8	$4 \cdot 2^2 = 16$
3	3	3·3 = 9	$3 \cdot 3^2 = 27$
4	4	4·4 = 16	$4 \cdot 4^2 = 64$
5	7	5·7 = 35	$7 \cdot 5^2 = 175$
6	6	6·6 = 36	$6 \cdot 6^2 = 216$
7	3	7·3 = 21	$3 \cdot 7^2 = 147$
8	5	8·5 = 40	$5 \cdot 8^2 = 320$
9	2	9·2 = 18	$2 \cdot 9^2 = 162$
10	4	10·4 = 40	$4 \cdot 10^2 = 400$

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i} = \frac{225}{40} = 5{,}6$$

$$\sigma = \sqrt{\frac{\sum f_i x_i^2}{\sum f_i} - \bar{x}^2} = \sqrt{\frac{1529}{40} - 5{,}6^2} = 2{,}6$$

$$\mathbf{5{,}6 \pm 2{,}6}$$

$$\boldsymbol{CV = \frac{2{,}6}{5{,}6} \cdot 100 = 46{,}4\%}$$

40	225	1529	
$\sum f_i$	$\sum f_i \cdot x_i$	$\sum f_i \cdot {x_i}^2$	

Errores absoluto y relativo

Por lo visto anteriormente, nos hemos de conformar con el valor medio de los resultados de una medida, sin poder saber si este coincide con el valor exacto de la medida. Sin embargo, el valor exacto debe estar muy próximo a dicho valor medio.

Todo esto nos hace utilizar en lugar del valor medio, el intervalo más probable en el que se encontrará el valor exacto de la medida. Para ello, junto al valor medio de la medida se utiliza la llamada **Imprecisión Absoluta** (δ), obteniéndose como intervalo más probable: $(x - \delta, x + \delta)$.

La imprecisión absoluta se va a obtener según los criterios:

1) Determinación del valor medio de las medidas.

$$\bar{x} = \frac{medida\,1 + medida\,2 + medida\,3 + medida\,4 + ...}{número \quad de \quad medidas}$$

2) Cálculo de la desviación entre cada medida y la media.
   Imprecisión 1 (δ_1) = $|$medida 1 - valor medio$|$
   Imprecisión 2 (δ_2) = $|$medida 2 - valor medio$|$
   Imprecisión 3 (δ_3) = $|$medida 3 - valor medio$|$
   Imprecisión 4 (δ_4) = $|$medida 4 - valor medio$|$
   ...

3) Determinación del valor medio de las desviaciones.

$$Imprecisión\ media\,(\bar{\delta}) = \frac{\delta_1 + \delta_2 + \delta_3 + \delta_4 + ...}{número\ de\ medidas}$$

   La imprecisión absoluta será este valor o el dado por la sensibilidad del aparato, tomándose el mayor de ellos.

4) La medida se expresa: (medida) ± imprecisión absoluta

¿Sería suficiente conocer las imprecisiones absolutas de dos medidas para poder comparar entre ellas cuál es la más precisa?

¿Sería comparable la imprecisión de 1 m al medir 1 Km y la de 1 m al medir el radio de la Tierra? No, no son comparables. En el 2do caso, la medida realizada es mucho más precisa (o correcta). Por lo tanto, la imprecisión absoluta no es suficiente para compararme una medida, de ahí que tengamos que introducir la llamada **Imprecisión Relativa:**

$$\boldsymbol{\varepsilon = \frac{\delta}{\bar{x}} \cdot 100}$$

La imprecisión relativa me indica la calidad con que se realiza una medida. Normalmente se expresa en %, por lo que se multiplica por 100.