

ESTADÍSTICA

Variable Estadística Unidimensional (X)

Frecuencia	<ul style="list-style-type: none"> - Absoluta (f_i): número de veces que aparece cada dato x_i - Relativa (h_i): frecuencia absoluta dividida por el número de datos $\rightarrow h_i = \frac{f_i}{n}$
Medidas de Posición	<ul style="list-style-type: none"> - Percentil de orden k (P_k): dato que ocupa la posición tal que el $k\%$ de los datos son menores que él - Mediana: P_{50} - Cuartiles: $Q_1=P_{25}$ y $Q_3=P_{75}$ - Deciles: $P_{10}, P_{20}, P_{30}, P_{40}, P_{60}, P_{70}, P_{80}$ y P_{90} - Valores atípicos: los que no pertenecen al intervalo $(Q_1 - 1.5 \cdot RI, Q_3 + 1.5 \cdot RI)$
Medidas de Centralización	<ul style="list-style-type: none"> - Moda: dato más repetido - Media (\bar{x}): suma de datos dividida por el número de datos $\rightarrow \bar{x} = \frac{\sum x_i f_i}{n}$ - Mediana: valor central de la distribución = P_{50}
Medidas de Dispersión	<ul style="list-style-type: none"> - Rango o Recorrido: diferencia entre los valores máximo y mínimo - Desviación media: media de los valores absolutos de las desviaciones de la media $DM(X) = \frac{\sum x_i - \bar{x} \cdot f_i}{n}$ - Varianza: media de los cuadrados de las desviaciones de la media $Var(X) = \frac{\sum (x_i - \bar{x})^2 \cdot f_i}{n} = \frac{\sum x_i^2 \cdot f_i}{n} - \bar{x}^2$ - Desviación Típica (σ_x): raíz cuadrada de la varianza $\sigma_x = \sqrt{Var(X)}$ - Coeficiente de Variación: $CV = \frac{Desviación\ Típica}{Media}$ - Recorrido Inter cuartílico: $RI = Q_3 - Q_1$
Clases	<p>Intervalos en los que se agrupan los datos de una variable cuantitativa continua.</p> <p>Al valor central del intervalo $[a, b)$ se le llama marca de clase: $x_i = \frac{a+b}{2}$</p>
Gráficos	<ul style="list-style-type: none"> - de Sectores - de Barras - Histograma - Polígono de Frecuencias - de Caja y Bigotes

DISTRIBUCIONES UNIDIMENSIONALES CON GEOGEBRA

A la hora de realizar un análisis estadístico podemos partir bien de una lista de datos brutos, o bien de dos listas: una de valores y otra de frecuencias.

Cálculo de Parámetros:

Media	media(Datos)	media(Valores, Frecuencias)
Varianza	Varianza(Datos)	Varianza(Valores, Frecuencias)
Desviación Típica	DE(Datos)	DE(Valores, Frecuencias)
Coefficiente de Variación	DE(Datos)/media(Datos)	DE(Valores, Frecuencias)/media(Valores, Frecuencias)
Mínimo	Mínimo(Datos)	Mínimo(Valores, Frecuencias)
Q ₁	Q1(Datos)	Q1(Valores, Frecuencias)
Mediana	Mediana(Datos)	Mediana(Valores, Frecuencias)
Q ₃	Q3(Datos)	Q3(Valores, Frecuencias)
Máximo	Máximo(Datos)	Máximo(Valores, Frecuencias)
Recorrido	Máximo(Datos)- Mínimo(Datos)	Máximo(Valores, Frecuencias)-Mínimo(Valores, Frecuencias)
Rec. Intercuartílico	Q3(Datos)-Q1(Datos)	Q3(Valores, Frecuencias)-Q1(Valores, Frecuencias)
Moda	Moda(Datos)	
Percentil k-ésimo	Percentil(Datos,k/100)	
Percentiles	Secuencia(Percentil(Datos, i / 100), i, 1, 100)	
Deciles	Secuencia(Percentil(Datos, i / 10), i, 1, 10)	
Desviación Media	media(Secuencia(abs(Datos(i)-media(Datos)), i, 1, N))	
TablaFrecuencias=TablaFrecuencias(Datos)		

Gráficos Estadísticos:

GBarras= Barras(Datos, a) GBarras= Barras(Valores, Frecuencias, a)	a = anchura de las barras
DCaja = DiagramaCaja(h,a,Datos,true) DCaja = DiagramaCaja(h,a, Valores, Frecuencias,true)	h = altura a la que se dibuja el gráfico a = anchura de la caja true = considerar valores atípicos
Histograma(Clases,Datos,true/false) Histograma(Clases,Valores,Frecuencias,true/false)	True → Las frecuencias quedan representadas por el área de las barras (Altura = Frecuencia/AnchoClase) False → Las frecuencias quedan representadas por la altura de las barras (Altura = Frecuencia)

En el caso de los Histogramas, los datos se agrupan en "clases": $[a_1, a_2), [a_2, a_3), \dots, [a_{n-1}, a_n]$. Hay 3 formas de hacerlo:

Clases = $\{a_1, a_2, \dots, a_n\}$	Establecemos "manualmente" las clases: $[a_1, a_2), [a_2, a_3), \dots, [a_{n-1}, a_n]$ Los intervalos pueden tener distinta anchura.
Clases = Clases(Datos, inicio, anchoclase)	Divide los datos en varias clases de la misma anchura (anchoclase) empezando por "inicio": $a_1 = \text{inicio}; a_2 = \text{inicio} + \text{anchoclase}; \dots$
Clases = Clases(Datos, numclases)	Divide los datos en "numclases" de la misma anchura (automática) desde el valor mínimo al valor máximo: inicio = mínimo anchoclase = $(\text{máximo} - \text{mínimo}) / \text{numclases}$
TablaClases = TablaFrecuencias(Clases, Datos)	

Parámetros Estadísticos en Excel y Geogebra

Variables Unidimensionales	Fórmula	EXCEL	GEOGEBRA
Media	$= \frac{\sum x_i f_i}{N}$	PROMEDIO(Datos)	Media(Datos)
Moda	Dato más repetido	MODA(Datos)	Moda(Datos)
Percentil (k%)	P _k = Valor que verifica que el k% de los datos es menor	PERCENTIL(Datos; k/100)	Percentil(Datos; k%)
Mediana	P ₅₀	MEDIANA(Datos)	Mediana(Datos)
Cuartil Primero = Q ₁	P ₂₅	CUARTIL(Datos;1)	Q1(Datos)
Cuartil Tercero = Q ₃	P ₇₅	CUARTIL(Datos;3)	Q3(Datos)
Desviación Media	$= \frac{\sum x_i - \bar{x} f_i}{N}$	DESVPROM(Datos)	Media(Secuencia(abs(Datos(i) - Media(Datos)), i, 1, N))
Varianza = σ^2	$= \frac{\sum (x_i - \bar{x})^2 f_i}{N} = \frac{\sum x_i^2 f_i}{N} - \bar{x}^2$	VARP(Datos)	Varianza(Datos)
Desviación Típica = σ	$= \sqrt{\sigma^2}$	DESVESTP(Datos)	DE(Datos)
Varianza Muestral = s^2	$\frac{\sum (x_i - \bar{x})^2 f_i}{N - 1} = \frac{\sum x_i^2 f_i - \frac{(\sum x_i f_i)^2}{N}}{N - 1}$	VAR(d)	VarianzaMuestral(Datos)
Desviación Típica Muestral = s	$= \sqrt{s^2}$	DESVEST(Datos)	DEmuestral(Datos)