Probabilidad: Cuantificar la ocurrencia de un evento con ayuda de la estadística.

$P(A) = \frac{n}{N}$ $\frac{casos\ favorables\ del\ evento\ A}{casos\ posibles\ del\ evento\ A}$

Jamás la probabilidad es mayor a 1

Reglas Aditivas

Si un evento puede ser expresado como la unión de otros eventos entonces:

$$P(C) = P(A \cup B) = P(A) + P(B) - P(\ A \cap B\ )$$

Probabilidad condicional

Probabilidad de ocurrencia de un evento cuando se sabe que ocurrió otro relacionado con éste.

$$P(B/A): Probabilidad\ del\ evento\ B\ dado\ que\ ocurrió\ A$$

Por lo tanto $P(B/A) = \frac{P(A \cap B)}{P(A)}$

Por tanto:

P(B/A) = P(B) y *P(A/B) = P(A)*

Entonces A y B son eventos *independientes*.

$$M \cap N = \emptyset$$

Si dos eventos no pueden ocurrir simultáneamente se dice que son "Mutuamente Excluyentes"

Teorema de probabilidad total

$$P(A) = \sum_{i=1}^{k} P(B_i \cap A) = \sum_{i=1}^{k} [P(B_i).P(A/B_i)]$$

B_i: Partición del Espacio Muestral $\Rightarrow \sum_{i=1}^{k} P(B_i) = 1$

Regla de Bayes (Probabilidad de las causas)

Si los eventos $B_1, B_2, \ldots, B_k$ constituyen una partición del Espacio Muestral S y $P(B_i) \neq 0$ para $i = 1,2,\ldots,k$ entonces cualquier evento B_r en S tal que $P(A) \neq 0$ puede expresarse como $P(B_r/A) = P(B_r \cap A) / P(A)$ $[\ Probabilidad\ Condicional]$.

Como $P(A/B_r) = P(B_r \cap A) / P(B_r) \Rightarrow P(B_r \cap A) = P(B_r).P(A/B_r)$

$$P(B_r/A) = \frac{P(B_r \cap A)}{\sum_{i=1}^{k} P(B_i \cap A)} = \frac{P(B_r).P(A/B_r)}{\sum_{i=1}^{k} [P(B_i).P(A/B_i)]}$$

Datos sin agrupar	***Datos agrupados*** $NIC = 5log_{10}n \quad o \quad NIC = \sqrt{n}, \qquad 5 \leq NIC \leq 15$ *Ancho del IC:* $A = \frac{R}{NIC}$
Medidas de posición o tendencia central	
Media: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ *Se puede obtener con Alcula.	$\bar{X} = \sum_{i=1}^{k}(X_{PM_i} . fr_i)$
Mediana/cuartiles: n impar: $Q_k = X_{\frac{k}{4}(n+1)}$ con k=1,2,3. n par: $Q_k = \frac{1}{2}\left[X_{\left(\frac{k}{4}n\right)} + X_{\left(\frac{k}{4}n+1\right)}\right]$ con k=1,2,3	Se determina a cuál intervalo de clase pertenece el cuartil. Es el primero que tenga Fr $\geq \frac{k}{4}$. $Q_k = L_i + \frac{\frac{k}{4} - Fr_{(i-1)}}{Fr_i - Fr_{(i-1)}} . A$ con k=1,2,3
Moda: Valor de mayor frecuencia.	Se determina a cuál intervalo de clase pertenece la moda. Es aquél con mayor fr. $Mo = L_i + \frac{A}{\frac{f_i - f_{(i+1)}}{f_i - f_{(i-1)}} + 1}$
Percentiles: $P_k = X_{\left(\frac{k}{100}n\right)}$ con k=1,…,99	Se determina a cuál intervalo de clase pertenece el percentil. Es el primero que tenga Fr $\geq \frac{k}{100}$. $P_k = L_i + \frac{\frac{k}{100} - Fr_{(i-1)}}{Fr_i - Fr_{(i-1)}} . A$ con k=1,…,99
Medidas de dispersión	
Varianza: $S^2 = \frac{n\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2}{n\,(n-1)}$ **Desvío estándar:** $S = \sqrt{S^2}$ *Se puede obtener con Alcula.	$S^2 = \frac{n\sum_{i=1}^{k} X_{PM_i}^2 . fa_i - \left[\sum_{i=1}^{k}(X_{PM_i} . fa_i)\right]^2}{n\,(n-1)}$ $S = \sqrt{S^2}$
Rango: $R = X_{máx} - X_{mín}$	$R = X_{máx} - X_{mín}$
Rango intercuartil: $IQ = Q_3 - Q_1$	$IQ = Q_3 - Q_1$
Coeficiente de variación: $CV = S/\bar{X}$	$CV = S/\bar{X}$
Medidas de forma	
Asimetría: $SK = \frac{3(\bar{X} - \tilde{X})}{S}$	$SK = \frac{3(\bar{X} - \tilde{X})}{S}$
Curtosis: $Cu = \frac{1}{n}\frac{\sum_{i=1}^{n}(x_i - \bar{X})^4}{S^4} - 3$	

Teorema de Chebyshev: en relación a un conjunto de datos cualquiera (poblacional o muestral) y una constante k > 1 cuando menos (1 - 1/k²) de los datos debe estar dentro de k desvíos estándar a uno y otro lado de la media para que la dispersión se considere pequeña.

Ejemplo: si elegimos $k = 2$ entonces $1 - \frac{1}{k^2} = \frac{3}{4} = 0{,}75$. El 75% de los datos debe estar a $\bar{X} + 2S$ y $\bar{X} - 2S$ para que la desviación se considere pequeña.

Cuartiles, Deciles y Percentiles: Dividen los datos en cuatro, diez o cien partes iguales.
Los datos deben estar ordenados.

$Q_1 = X_{\frac{n+1}{4}}$; $Q_2 = X_{\frac{n+1}{2}} = \tilde{X}$; $Q_3 = X_{\frac{3.(n+1)}{4}}$ con n impar

$Q_1 = \frac{1}{2}\left(X_{\frac{n}{4}} + X_{\frac{n}{4}+1}\right)$; $Q_2 = \frac{1}{2}\left(X_{\frac{n}{2}} + X_{\frac{n}{2}+1}\right)$; $Q_3 = \frac{1}{2}\left(X_{\frac{3.n}{2}} + X_{\frac{3}{4}.(n+1)}\right)$ con n par

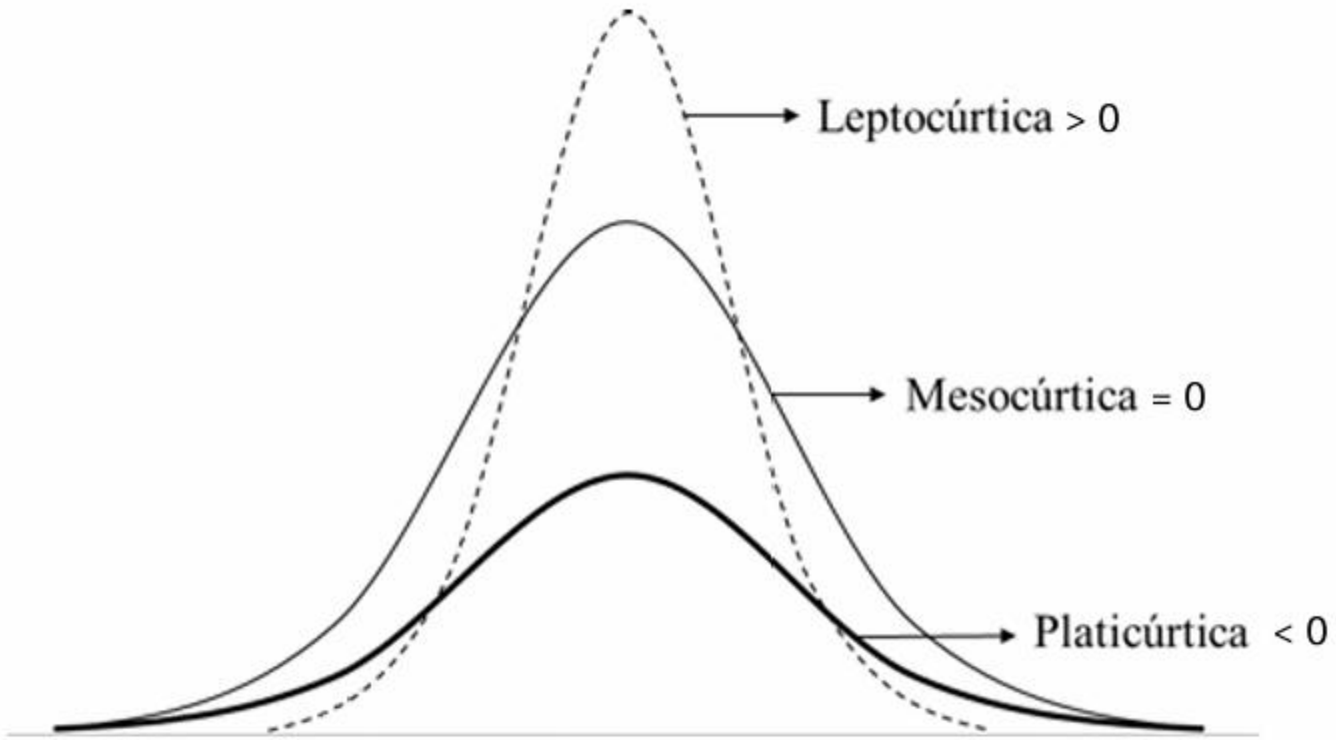
Una regla para determinar si un dato es **anómalo** (outlier) es:

- Si un dato es < Q1 – 1.5(Q3-Q1)
- Si un dato es > Q3 + 1.5(Q3-Q1)

Curtosis: Miden la mayor o menor concentración de datos alrededor de la media.

El grado de Curtosis es:

$$Cu = \frac{1}{n}\frac{\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^4}{S^4} - 3$$



Teorema de probabilidad total

$$P(A) = \sum_{i=1}^{k} P(B_i \cap A) \ = \sum_{i=1}^{k} [P(B_i).P(A/B_i)]$$

B_i: Partición del Espacio Muestral $\Rightarrow \sum_{i=1}^{k} P(B_i) \ = \ 1$

Si quisiéramos encontrar la probabilidad del evento A en términos de los eventos B_i donde todos los B_i son mutuamente excluyentes.