Resumen de fórmulas de Tema 2

Datos sin agrupar	***Datos agrupados*** $NIC = log_{10}\, n \quad o \quad NIC = \sqrt{n}, \qquad 5 \leq NIC \leq 15$ *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.