

Formulario

Elementos de la distribución de frecuencias de una variable cualitativa

Modalidades (a_i)	n_i	N_i	f_i	F_i
a_1	n_1	$N_1 = n_1$	$f_1 = n_1/N$	$F_1 = f_1 = N_1/N$
a_2	n_2	$N_2 = n_1 + n_2 = N_1 + n_2$	$f_2 = n_2/N$	$F_2 = f_1 + f_2 = F_1 + f_2 = N_2/N$
\vdots	\vdots	\vdots	\vdots	\vdots
a_i	n_i	$N_i = n_1 + \dots + n_i = N_{i-1} + n_i$	$f_i = n_i/N$	$F_i = f_1 + \dots + f_i = F_{i-1} + f_i = N_i/N$
\vdots	\vdots	\vdots	\vdots	\vdots
a_k	n_k	$N_k = n_1 + \dots + n_k = N_{k-1} + n_k = N$	$f_k = n_k/N$	$F_k = f_1 + \dots + f_k = F_{k-1} + f_k = N_k/N = 1$
	$\sum_{i=1}^k n_i = N$		$\sum_{i=1}^k f_i = 1$	

Elementos de la distribución de frecuencias de una variable cuantitativa (no agrupada en intervalos)

Valores (x_i)	n_i	N_i	f_i	F_i
x_1	n_1	$N_1 = n_1$	$f_1 = n_1/N$	$F_1 = f_1 = N_1/N$
x_2	n_2	$N_2 = n_1 + n_2 = N_1 + n_2$	$f_2 = n_2/N$	$F_2 = f_1 + f_2 = F_1 + f_2 = N_2/N$
\vdots	\vdots	\vdots	\vdots	\vdots
x_i	n_i	$N_i = n_1 + \dots + n_i = N_{i-1} + n_i$	$f_i = n_i/N$	$F_i = f_1 + \dots + f_i = F_{i-1} + f_i = N_i/N$
\vdots	\vdots	\vdots	\vdots	\vdots
x_k	n_k	$N_k = n_1 + \dots + n_k = N_{k-1} + n_k = N$	$f_k = n_k/N$	$F_k = f_1 + \dots + f_k = F_{k-1} + f_k = N_k/N = 1$
	$\sum_{i=1}^k n_i = N$		$\sum_{i=1}^k f_i = 1$	

Elementos de la distribución de frecuencias de una variable cuantitativa (agrupada en intervalos)

Intervalos ($L_{i-1} - L_i]$)	n_i	N_i	f_i	F_i	Marcas de clase (x_i)	c_i	d_i
$L_0 - L_1$	n_1	N_1	f_1	F_1	$x_1 = (L_0 + L_1)/2$	$c_1 = L_1 - L_0$	$d_1 = n_1/c_1$
$L_1 - L_2$	n_2	N_2	f_2	F_2	$x_2 = (L_1 + L_2)/2$	$c_2 = L_2 - L_1$	$d_2 = n_2/c_2$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$L_{i-1} - L_i$	n_i	N_i	f_i	F_i	$x_i = (L_{i-1} + L_i)/2$	$c_i = L_i - L_{i-1}$	$d_i = n_i/c_i$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$L_{k-1} - L_k$	n_k	$N_k = N$	f_k	$F_k = 1$	$x_k = (L_{k-1} + L_k)/2$	$c_k = L_k - L_{k-1}$	$d_k = n_k/c_k$
	$\sum_{i=1}^k n_i = N$		$\sum_{i=1}^k f_i = 1$				

Tabla de correlación

$x_i \setminus y_j$	y_1	...	y_j	...	y_h	$n_{i.}$
x_1	n_{11}	...	n_{1j}	...	n_{1h}	$n_{1.} = n_{11} + \dots + n_{1h}$
x_2	n_{21}	...	n_{2j}	...	n_{2h}	$n_{2.} = n_{21} + \dots + n_{2h}$
\vdots	\vdots		\vdots		\vdots	\vdots
x_i	n_{i1}	...	n_{ij}	...	n_{ih}	$n_{i.} = n_{i1} + \dots + n_{ih}$
\vdots	\vdots		\vdots		\vdots	\vdots
x_k	n_{k1}	...	n_{kj}	...	n_{kh}	$n_{k.} = n_{k1} + \dots + n_{kh}$
$n_{.j}$	$n_{.1} = n_{11} + \dots + n_{k1}$...	$n_{.j} = n_{1j} + \dots + n_{kj}$...	$n_{.h} = n_{1h} + \dots + n_{kh}$	$\sum_{i=1}^k \sum_{j=1}^h n_{ij} = N$

Tabla de contingencia

$a_i \setminus b_j$	b_1	...	b_j	...	b_h	$n_{i.}$
a_1	n_{11}	...	n_{1j}	...	n_{1h}	$n_{1.} = n_{11} + \dots + n_{1h}$
a_2	n_{21}	...	n_{2j}	...	n_{2h}	$n_{2.} = n_{21} + \dots + n_{2h}$
\vdots	\vdots		\vdots		\vdots	\vdots
a_i	n_{i1}	...	n_{ij}	...	n_{ih}	$n_{i.} = n_{i1} + \dots + n_{ih}$
\vdots	\vdots		\vdots		\vdots	\vdots
a_k	n_{k1}	...	n_{kj}	...	n_{kh}	$n_{k.} = n_{k1} + \dots + n_{kh}$
$n_{.j}$	$n_{.1} = n_{11} + \dots + n_{k1}$...	$n_{.j} = n_{1j} + \dots + n_{kj}$...	$n_{.h} = n_{1h} + \dots + n_{kh}$	$\sum_{i=1}^k \sum_{j=1}^h n_{ij} = N$

Bloque I. Análisis de una variable					
Media aritmética	$\bar{x} = \frac{1}{N} \sum_{i=1}^k x_i n_i$			Rango	$R = \max\{x_1, \dots, x_k\} - \min\{x_1, \dots, x_k\}$
				Recorrido intercuartílico	$R_I = Q_3 - Q_1$
Media aritmética ponderada	$\bar{x}_w = \frac{w_1 x_1 + w_2 x_2 + \dots + w_k x_k}{w}$			Varianza	$S^2 = \frac{1}{N} \sum_{i=1}^k (x_i - \bar{x})^2 n_i = \frac{1}{N} \sum_{i=1}^k x_i^2 n_i - \bar{x}^2$
Media de la composición de poblaciones	$\bar{x}_p = \frac{N_1 \bar{x}_1 + N_2 \bar{x}_2 + \dots + N_k \bar{x}_k}{N}$			Desviación típica	$S = +\sqrt{S^2}$
Media geométrica	$\bar{x}_g = \sqrt[N]{x_1^{n_1} \dots x_k^{n_k}} = (x_1^{n_1} \dots x_k^{n_k})^{1/N}$			Coefficiente de apertura	$A = \max\{x_1, \dots, x_k\} / \min\{x_1, \dots, x_k\}$
Media armónica	$\bar{x}_a = \frac{N}{\frac{n_1}{x_1} + \frac{n_2}{x_2} + \dots + \frac{n_k}{x_k}}$			Recorrido relativo	$R_R = [\max\{x_1, \dots, x_k\} - \min\{x_1, \dots, x_k\}] / \bar{x}$
				Recorrido semi-intercuartílico	$R_S = (Q_3 - Q_1) / (Q_1 + Q_3)$
Mediana	No agrupada en intervalos	$N_i = N/2$	$Me = (x_i + x_{i+1}) / 2$	Coefficiente de variación de Pearson	$CV = \frac{S}{ \bar{x} }$
		$N_i > N/2$	$Me = x_i$	Variable tipificada	$Z = \frac{X - \bar{x}}{S}$
	Agrupada en intervalos	$N_i = N/2$	$Me = L_i$	Coefficiente de asimetría de Fisher	$g_1 = \frac{m_3}{S^3} = \frac{\frac{1}{N} \sum_{i=1}^k (x_i - \bar{x})^3 n_i}{(S^2)^{3/2}}$
		$N_i > N/2$	$Me = L_{i-1} + \frac{N/2 - N_{i-1}}{n_i} \cdot c_i$	Coefficiente de asimetría de Yule-Bowley	$AB = \frac{Q_1 + Q_3 - 2Q_2}{Q_3 - Q_1}$
Moda	Agrupada en intervalos	$Mo = L_{i-1} + \frac{n_{i+1}}{n_{i-1} + n_{i+1}} \cdot c_i$		Coefficiente de curtosis	$g_2 = \frac{m_4}{S^4} - 3 = \frac{\frac{1}{N} \sum_{i=1}^k (x_i - \bar{x})^4 n_i}{(S^2)^2} - 3$
		$Mo = L_{i-1} + \frac{d_{i+1}}{d_{i-1} + d_{i+1}} \cdot c_i$			
Cuantiles <small>Cuartiles (Q_r): $k = 4, r = 1, 2, 3$ Deciles (D_r): $k = 10, r = 1, \dots, 9$ Percentiles (P_r): $k = 100, r = 1, \dots, 99$</small>	No agrupada en intervalos	$N_i = (r/k)N$	$C_{r/k} = \frac{x_i + x_{i+1}}{2}$	Índice de Gini	$IG = \frac{\sum_{i=1}^{k-1} (p_i - q_i)}{\sum_{i=1}^{k-1} p_i}$
		$N_i > (r/k)N$	$C_{r/k} = x_i$		
	Agrupada en intervalos	$N_i = (r/k)N$	$C_{r/k} = L_i$		
		$N_i > (r/k)N$	$C_{r/k} = L_{i-1} + \frac{(r/k) \cdot N - N_{i-1}}{n_i} \cdot c_i$		
		$p_i = \frac{N_i}{N} \cdot 100$	$q_i = \frac{u_i}{u_k} \cdot 100$		

Bloque II. Análisis conjunto de dos variables							
Independencia estadística	$n_{ij} = \frac{n_i \cdot n_j}{N} \quad \forall i, j$	Covarianza	$S_{XY} = \frac{1}{N} \sum_{i=1}^k \sum_{j=1}^h (x_i - \bar{x})(y_j - \bar{y})n_{ij} = \frac{1}{N} \sum_{i=1}^k x_i \sum_{j=1}^h y_j n_{ij} - \bar{x}\bar{y}$				
Coefficiente de correlación lineal	$r_{XY} = \frac{S_{XY}}{S_X \cdot S_Y}$	Rectas de regresión lineal	$y - \bar{y} = \frac{S_{XY}}{S_X^2} (x - \bar{x})$		$x - \bar{x} = \frac{S_{XY}}{S_Y^2} (y - \bar{y})$		
Coefficiente de determinación	$R_{XY}^2 = r_{XY}^2$						
ATRIBUTOS							
Coefficiente básico de dependencia	$D = \frac{n_{11}n_{22} - n_{12}n_{21}}{N} = n_{11} - \frac{n_{1.} \times n_{.1}}{N}$	Coefficiente de asociación Q de Yule	$Q = \frac{n_{11}n_{22} - n_{12}n_{21}}{n_{11}n_{22} + n_{12}n_{21}} = \frac{N \cdot D}{n_{11}n_{22} + n_{12}n_{21}}$				
Estadístico chi-cuadrado	$\chi^2 = \sum_{i=1}^k \sum_{j=1}^h \frac{(n_{ij} - e_{ij})^2}{e_{ij}}$	Coefficiente de contingencia	$C = \sqrt{\frac{\chi^2}{N + \chi^2}}$	Coefficiente V de Cramer	$V = \sqrt{\frac{\chi^2}{N[\text{mín}(k, h) - 1]}}$	Coefficiente T de Tschuprow	$T = \frac{\chi^2}{N\sqrt{(k-1)(h-1)}}$
Coefficiente de correlación por rangos de Spearman				$\rho = 1 - \frac{6}{N(N^2 - 1)} \sum_{i=1}^n d_i^2$			
Bloque III. Análisis de las variables a lo largo del tiempo							
Índices simples	$I_{t/0} = \frac{x_t}{x_0} \cdot 100$	Tasa de variación absoluta	$\nabla x_t = x_t - x_{t-1}$				
Índices en cadena	$I_{t/t-1} = \frac{x_t}{x_{t-1}} \cdot 100$	Tasa de variación relativa	$\dot{x}_t = \left(\frac{x_t - x_{t-1}}{x_{t-1}} \right) \cdot 100 = \left(\frac{x_t}{x_{t-1}} - 1 \right) \cdot 100$				
Cambio de base	$I_{t/t'} = \frac{I_{t/t''}}{I_{t'/t''}}$	Tasa media acumulativa de un periodo	$r_{t_1/t_2} = \left[\left(\frac{x_{t_2}}{x_{t_1}} \right)^{\frac{1}{t_2 - t_1}} - 1 \right] \cdot 100$				
Deflactación	Magnitud del año t a precios constantes (base 0) = $\frac{\text{Magnitud del año } t \text{ a precios corrientes}}{I_{t0}}$						