

Distribuciones bidimensionales

Los N elementos de la población son clasificados según dos caracteres X e Y , cuyas modalidades las notamos respectivamente por x_i e y_j donde i varía desde 1 hasta p y j varía desde 1 hasta q .

n_{ij} → **Frecuencia absoluta**: número de individuos de la población que presentan la modalidad x_i de X y la modalidad y_j de Y .

Tabla de **frecuencias absolutas** de una variable bidimensional

(X/Y)	y_1	y_2	y_3	...	y_q	<i>suma</i>
x_1	n_{11}	n_{12}	n_{13}	...	n_{1q}	$n_{1\cdot}$
x_2	n_{21}	n_{22}	n_{23}	...	n_{2q}	$n_{2\cdot}$
x_3	n_{31}	n_{32}	n_{33}	...	n_{3q}	$n_{3\cdot}$
...
x_p	n_{p1}	n_{p2}	n_{p3}	...	n_{pq}	$n_{p\cdot}$
suma	$n_{\cdot 1}$	$n_{\cdot 2}$	$n_{\cdot 3}$...	$n_{\cdot q}$	$n_{\cdot \cdot} = N$

f_{ij} → **Frecuencia relativa**: proporción de individuos de la población que presentan la modalidad x_i de X y la modalidad y_j de Y . Se obtiene dividiendo la frecuencia absoluta entre el número de elementos de la población (N).

Tabla de **frecuencias relativas** de una variable bidimensional

(X/Y)	y_1	y_2	y_3	...	y_q	<i>suma</i>
x_1	f_{11}	f_{12}	f_{13}	...	f_{1q}	$f_{1\cdot}$
x_2	f_{21}	f_{22}	f_{23}	...	f_{2q}	$f_{2\cdot}$
x_3	f_{31}	f_{32}	f_{33}	...	f_{3q}	$f_{3\cdot}$
...
x_p	f_{p1}	f_{p2}	f_{p3}	...	f_{pq}	$f_{p\cdot}$
suma	$f_{\cdot 1}$	$f_{\cdot 2}$	$f_{\cdot 3}$...	$f_{\cdot q}$	$f_{\cdot \cdot} = 1$

En la tabla el primer subíndice indica la fila y el segundo subíndice indica la columna.

COVARIANZA

$$\sigma_{XY} = Cov(X, Y) = \frac{1}{N} \sum_{j=1}^q \sum_{i=1}^p (x_i - \bar{x})(y_j - \bar{y}) n_{ij} = \frac{1}{N} \sum_{j=1}^q \sum_{i=1}^p x_i y_j n_{ij} - \bar{x} \bar{y}$$

DISTRIBUCIONES MARGINALES

Estas distribuciones nos indican cómo se distribuye una variable independientemente de los valores que tome la otra.

Distribución marginal de X :

X	$n_{i\cdot}$
x_1	$n_{1\cdot}$
x_2	$n_{2\cdot}$
x_3	$n_{3\cdot}$
...	...
x_p	$n_{p\cdot}$
suma	$n_{\cdot\cdot}=N$

Distribución marginal de Y :

Y	$n_{\cdot j}$
y_1	$n_{\cdot 1}$
y_2	$n_{\cdot 2}$
y_3	$n_{\cdot 3}$
...	...
y_q	$n_{\cdot q}$
suma	$n_{\cdot\cdot}=N$

Existe una única distribución marginal de X y una única distribución marginal de Y .

DISTRIBUCIONES CONDICIONADAS

Son distribuciones unidimensionales obtenidas a partir de las bidimensionales, manteniendo fijo un valor en una de las variables y considerando los valores que toma la otra con sus respectivas frecuencias

Distribución de X condicionada al valor y_j de Y :

$X Y=y_j$	n_{ij}
x_1	n_{1j}
x_2	n_{2j}
x_3	n_{3j}
...	...
x_p	n_{pj}
suma	$n_{\cdot j}$

Distribución de Y condicionada al valor x_i de X :

$Y X=x_i$	n_{ji}
y_1	n_{i1}
y_2	n_{i2}
y_3	n_{i3}
...	...
y_q	n_{iq}
suma	$n_{i\cdot}$

- Las variables X e Y se dicen **dependientes** estadísticamente cuando la variación de una influye en la distribución de la otra.

- Cuando las distribuciones condicionadas a cualquier valor de la otra variable sean iguales, diremos que las variables X e Y son estadísticamente **independientes**. Las variables son independientes si y solamente si ocurre que $f_{ij} = f_{i\cdot}f_{\cdot j}$ para cualquier pareja (i,j) .

- Cuando las variables son independientes la covarianza vale 0 (siempre).

- Cuando la covarianza vale 0 las variables son **INCORRELADAS**.