Introducción a la Estadística

Tema 3

Distribuciones bidimensionales

3.1 CONSIDERACIONES GENERALES

Si de una población o muestra de tamaño "N" nos interesa observar a la vez dos caracteres, a cada elemento de la población le corresponde un par de valores o modalidades, según que los caracteres observados sean variables o atributos.

Sólo estudiaremos el caso en que los dos caracteres son cuantitativos, llamando **variable bidimensional** a la correspondiente variable estadística.

Del número de veces que observemos el par $(x_i;y_j)$ diremos que es la **frecuencia absoluta bidimensional** de dicho par, y se denota n_{ij}.

La **frecuencia relativa bidimensional** de $(x_i;y_j)$ se denota f_{ij}, y es el cociente entre su frecuencia absoluta bidimensional n_{ij} y la suma "n" de las frecuencias absolutas bidimensionales:

$$f_{ij} = \frac{n_{ij}}{N} = \frac{n_{ij}}{\sum_i \sum_j n_{ij}}$$

Llamaremos distribución bidimensional al conjunto de los pares $(x_i;y_j)$ con sus respectivas frecuencias absolutas o relativas.

3.2 TABLA DE CORRELACIÓN

Es un cuadro de doble entrada mediante el que representaremos las distribuciones bidimensionales.

X / Y	y_1	::::	y_j	::::	y_m	Suma
x_1	n_{11}	::::	n_{ij}	::::	n_{1m}	$n_{1\bullet}$
:	:	::::	:	::::	:	
x_i	n_{i1}	::::	n_{ij}	::::	n_{im}	$n_{i\bullet}$
:	:	::::	:	::::	:	
x_k	n_{k1}	::::	n_{kj}	::::	n_{km}	$n_{k\bullet}$
Suma	$n_{\bullet 1}$	::::	$n_{\bullet j}$	::::	$n_{\bullet m}$	$\boxed{N}$

siendo

$$n_{i\bullet} = \sum_{j=1}^{m} n_{ij} \equiv \text{frecuencia absoluta de } x_i$$

$$n_{\bullet j} = \sum_{i=1}^{k} n_{ij} \equiv \text{frecuencia absoluta de } y_j$$

$$f_{i\bullet} = \sum_{j=1}^{m} f_{ij} = \sum_{j=1}^{m} \frac{n_{ij}}{N} = \frac{1}{N}.\sum_{j=1}^{m} n_{ij} = \frac{n_{i\bullet}}{n} \equiv \text{frecuencia relativa de } x_i$$

$$f_{\bullet j} = \sum_{i=1}^{k} f_{ij} = \sum_{j=1}^{k} \frac{n_{ij}}{N} = \frac{1}{N}.\sum_{i=1}^{k} n_{ij} = \frac{n_{\bullet j}}{n} \equiv \text{frecuencia relativa de } y_j$$

$$\sum_{i=1}^{k} \sum_{j=1}^{m} f_{ij} = \sum_{i=1}^{k} f_{i\bullet} = \sum_{j=1}^{m} f_{\bullet j} = 1$$

Exactamente lo mismo si "X" e "Y" se expresan mediante intervalos de clase.

Por ejemplo: considera que al lanzar dos dados 24 veces se obtienen los siguientes resultados:

X	1	2	2	3	5	4	1	3	3	4	1	2	5	4	3	4	4	5	3	1	6	5	4	6
Y	2	3	1	4	3	2	6	4	1	6	6	5	1	2	5	1	1	2	6	6	2	1	2	5

Así, la correspondiente **tabla de correlación** es la siguiente:

X / Y	1	2	3	4	5	6	$n_{i\bullet}$
1	0	1	0	0	0	3	4
2	1	0	1	0	1	0	3
3	1	0	0	2	0	0	3
4	2	3	0	0	1	1	7
5	2	1	1	0	0	1	5
6	0	1	0	0	1	0	2
$n_{\bullet j}$	6	6	2	2	3	5	N = 24

3.3 DISTRIBUCIONES MARGINALES

En una distribución bidimensional (X;Y), la **distribución marginal de la variable "X"** es la distribución unidimensional de frecuencias definida por los valores que toma "X" y las frecuencias de los mismos .... y la **distribución marginal de la variable "Y"** es la distribución unidimensional de frecuencias definida por los valores que toma "Y" y las frecuencias de los mismos.

X	x_1	x_2	::::	x_k
$n_{i\bullet}$	$n_{1\bullet}$	$n_{2\bullet}$	::::	$n_{k\bullet}$

;

Y	y_1	y_2	::::	y_m
$n_{\bullet j}$	$n_{\bullet 1}$	$n_{\bullet 1}$	::::	$n_{\bullet m}$

Por ejemplo, la distribución de frecuencias marginales de "X" en el caso de los dos dados, es:

X / Y	1	2	3	4	5	6	$n_{i\bullet}$
1	0	1	0	0	0	3	4
2	1	0	1	0	1	0	3
3	1	0	0	2	0	0	3
4	2	3	0	0	1	1	7
5	2	1	1	0	0	1	5
6	0	1	0	0	1	0	2
							N = 24

$\Rightarrow$

X	1	2	3	4	5	6
$n_{i\bullet}$	4	3	3	7	5	2

• La media de "X" es:

$$\bar{x} = \frac{1.4 + 2.3 + 3.3 + 4.7 + 5.5 + 6.2}{24} = \frac{84}{24} = 3'5$$

• El momento de segundo orden de "X" respecto al origen es:

$$\frac{1^2.4 + 2^2.3 + 3^2.3 + 4^2.7 + 5^2.5 + 6^2.2}{24} = \frac{352}{24} = 14'66$$

• La varianza de "X" es $14'66 - 3'5^2 = 2'41$.

Por ejemplo, la distribución de frecuencias marginales de "Y" en el caso de los dos dados, es:

Y	1	2	3	4	5	6
$n_{\bullet j}$	6	6	2	2	3	5

• La media de "Y" es:

$$\bar{y} = \frac{1.6 + 2.6 + 3.2 + 4.2 + 5.3 + 6.5}{24} = \frac{77}{24} = 3'21$$

• El momento de segundo orden de "Y" respecto al origen es:

$$\frac{1^2.6 + 2^2.6 + 3^2.2 + 4^2.2 + 5^2.3 + 6^2.5}{24} = \frac{335}{24} = 13'95$$

• La varianza de "Y" es $13'95 - 3'21^2 = 3'64$.

3.4 DISTRIBUCIONES CONDICIONADAS

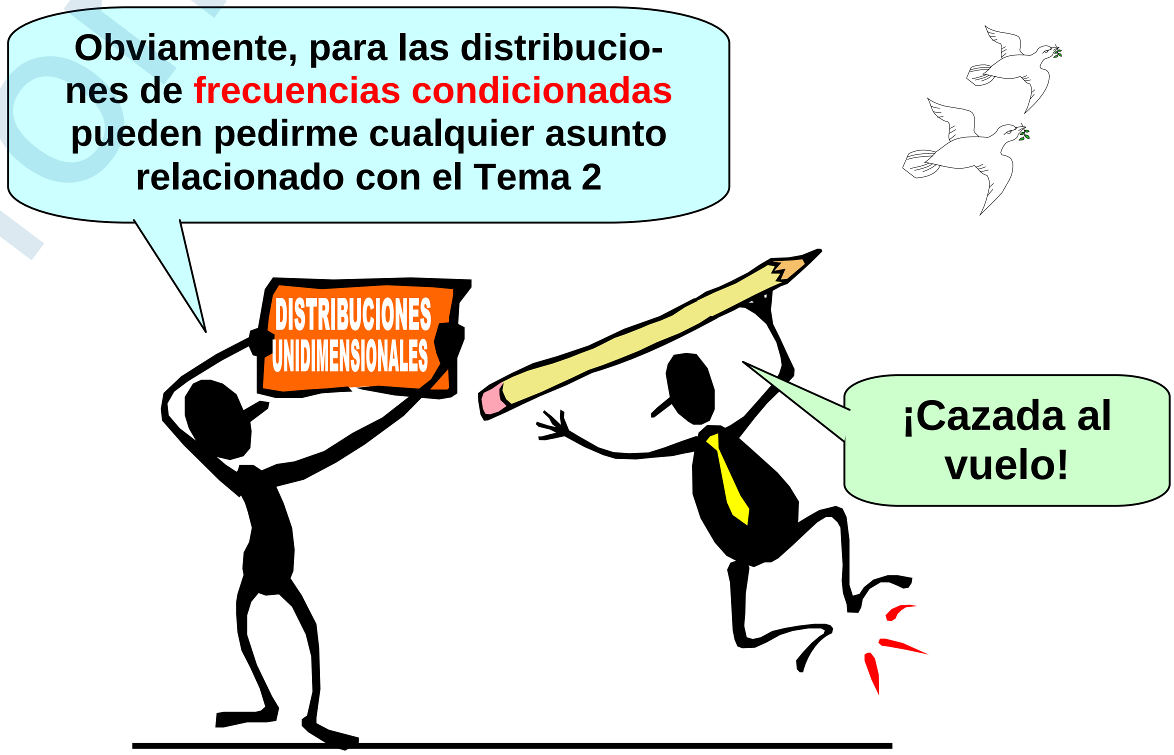
- La frecuencia de y_j condicionada a que $X = x_i$ se denota $f(y_j / x_i)$, y es la frecuencia relativa con que se observa y_j en el conjunto de observaciones en que $X = x_i$; o sea: $f(y_j / x_i) = \dfrac{n_{ij}}{n_{i\bullet}} = \dfrac{f_{ij}}{f_{i\bullet}}$. La distribución de frecuencias de "Y" condicionada a que $X = x_i$ es la distribución unidimensional que forman los valores $y_1, \ldots, y_m$ que toma "Y" con sus correspondientes frecuencias condicionadas $f(y_1 / x_i), \ldots, f(y_m / x_i)$.

- La frecuencia de x_i condicionada a que $Y = y_j$ se denota $f(x_i / y_j)$, y es la frecuencia relativa con que se observa x_i en el conjunto de observaciones en que $Y = y_j$; o sea: $f(x_i / y_j) = \dfrac{n_{ij}}{n_{\bullet j}} = \dfrac{f_{ij}}{f_{\bullet j}}$. La distribución de frecuencias de la variable "X" condicionada a que $Y = y_j$ es la distribución unidimensional que forman los valores $x_1, \ldots, x_k$ que toma "X" con sus correspondientes frecuencias condicionadas $f(x_1 / y_j), \ldots, f(x_k / y_j)$.

- Se verifica que

$$\sum_{i=1}^{k} f(x_i / y_j) = \frac{\sum_{i=1}^{k} f_{ij}}{f_{\bullet j}} = \frac{f_{\bullet j}}{f_{\bullet j}} = 1$$

$$\sum_{j=1}^{m} f(y_j / x_i) = \frac{\sum_{j=1}^{m} f_{ij}}{f_{i\bullet}} = \frac{f_{i\bullet}}{f_{i\bullet}} = 1$$



Por ejemplo, en el caso de los dos dados, determinemos la distribución de frecuencias de la variable "Y" condicionada a que $X = 5$:

X / Y	1	2	3	4	5	6	$n_{i\bullet}$
1	0	1	0	0	0	3	4
2	1	0	1	0	1	0	3
3	1	0	0	2	0	0	3
4	2	3	0	0	1	1	7
5	$\boxed{2}$	$\boxed{1}$	$\boxed{1}$	$\boxed{0}$	$\boxed{0}$	$\boxed{1}$	$\boxed{5}$
6	0	1	0	0	1	0	2
$n_{\bullet j}$	6	6	2	2	3	5	$N = 24$

$\Rightarrow$

$Y / X = 5$	1	2	3	4	5	6
$f(y_j / x_i = 5)$	2/5	1/5	1/5	0	0	1/5

- La media de $Y / X = 5$ es:

$$\bar{y}_{x_i=5} = \sum_{j=1}^{6} y_j . f(y_j / x_i = 5) = 1.\frac{2}{5} + 2.\frac{1}{5} + 3.\frac{1}{5} + 4.0 + 5.0 + 6.\frac{1}{5} = \frac{13}{5} = 2'6$$

- El momento de segundo orden de $Y / X = 5$ respecto al origen es:

$$1^2.\frac{2}{5} + 2^2.\frac{1}{5} + 3^2.\frac{1}{5} + 4^2.0 + 5^2.0 + 6^2.\frac{1}{5} = \frac{50}{5} = 10$$

- La varianza de $Y / X = 5$ es $10 - 2'6^2 = 3'24$.

Por ejemplo, en el caso de los dos dados, determinemos la distribución de frecuencias de la variable "X" condicionada a que $Y = 2$:

X / Y	1	2	3	4	5	6	$n_{i\bullet}$
1	0	$\boxed{1}$	0	0	0	3	4
2	1	$\boxed{0}$	1	0	1	0	3
3	1	$\boxed{0}$	0	2	0	0	3
4	2	$\boxed{3}$	0	0	1	1	7
5	2	$\boxed{1}$	1	0	0	1	5
6	0	$\boxed{1}$	0	0	1	0	2
$n_{\bullet j}$	6	$\boxed{6}$	2	2	3	5	$N = 24$

$\Rightarrow$

$X / Y = 2$	1	2	3	4	5	6
$f(x_i / y_j = 2)$	1/6	0	0	3/6	1/6	1/6

- La media de $X / Y = 2$ es:

$$\bar{x}_{y_j=5} = \sum_{i=1}^{6} x_i . f(x_i / y_j = 2) = 1.\frac{1}{6} + 2.0 + 3.0 + 4.\frac{3}{6} + 5.\frac{1}{8} + 6.\frac{1}{6} = \frac{24}{6} = 4$$

- El momento de segundo orden de $X / Y = 2$ respecto al origen es:

$$1^2.\frac{1}{6} + 2^2.0 + 3^2.0 + 4^2.\frac{3}{6} + 5^2.\frac{1}{8} + 6^2.\frac{1}{6} = \frac{109}{6} = 18'16$$

- La varianza de $X / Y = 2$ es $18'16 - 4^2 = 2'16$

3.5 MOMENTOS BIDIMENSIONALES

Momentos respecto al origen

El momento de orden (r;s) respecto al origen se denota a_{rs}, siendo:

$$a_{rs} = \sum_{i=1}^{k} \sum_{j=1}^{m} x_i^r . y_j^s . f_{ij}$$

Casos famosos:

$$a_{10} = \sum_{i=1}^{k} \sum_{j=1}^{m} x_i^1 . y_j^0 . f_{ij} = \sum_{i=1}^{k} \sum_{j=1}^{m} x_i . f_{ij} = \sum_{i=1}^{k} x_i . \left(\sum_{j=1}^{m} f_{ij} \right) = \sum_{i=1}^{k} x_i . f_{i\bullet} = \overline{x}$$

$$a_{01} = \sum_{i=1}^{k} \sum_{j=1}^{m} x_i^0 . y_j^1 . f_{ij} = \sum_{i=1}^{k} \sum_{j=1}^{m} y_j . f_{ij} = \sum_{j=1}^{m} y_j . \left(\sum_{i=1}^{k} f_{ij} \right) = \sum_{j=1}^{m} y_j . f_{\bullet j} = \overline{y}$$

De igual modo, $a_{20} = \sum_{i=1}^{k} \sum_{j=1}^{m} x_i^2 . y_j^0 . f_{ij}$ es el momento de orden 2 que respecto al origen tiene la distribución marginal de "X", y $a_{02} = \sum_{i=1}^{k} \sum_{j=1}^{m} x_i^0 . y_j^2 . f_{ij}$ es el momento de orden 2 que respecto al origen tiene la distribución marginal de "Y".

Momentos respecto a las medias

El momento de orden (r;s) respecto a las medias se denota m_{rs}, siendo:

$$m_{rs} = \sum_{i=1}^{k} \sum_{j=1}^{m} (x_i - \overline{x})^r . (y_j - \overline{y})^s . f_{ij}$$

Casos famosos:

$$m_{10} = \sum_{i=1}^{k} \sum_{j=1}^{m} (x_i - \overline{x})^1 . f_{ij} = \sum_{i=1}^{k} (x_i - \overline{x}) . \left(\sum_{j=1}^{m} f_{ij} \right) = \sum_{i=1}^{k} (x_i - \overline{x}) . f_{i\bullet} = 0$$

$$m_{01} = \sum_{i=1}^{k} \sum_{j=1}^{m} (y_j - \overline{y})^1 . f_{ij} = \sum_{j=1}^{m} (y_j - \overline{y}) . \left(\sum_{i=1}^{k} f_{ij} \right) = \sum_{j=1}^{m} (y_j - \overline{y}) . f_{\bullet j} = 0$$

$$m_{20} = \sum_{i=1}^{k} \sum_{j=1}^{m} (x_i - \overline{x})^2 . f_{ij} = \sum_{i=1}^{k} (x_i - \overline{x})^2 . \left(\sum_{j=1}^{m} f_{ij} \right) = \sum_{i=1}^{k} (x_i - \overline{x})^2 . f_{i\bullet} = S_X^2$$

$$m_{02} = \sum_{i=1}^{k} \sum_{j=1}^{m} (y_j - \overline{y})^2 . f_{ij} = \sum_{j=1}^{m} (y_j - \overline{y})^2 . \left(\sum_{i=1}^{k} f_{ij} \right) = \sum_{j=1}^{m} (y_j - \overline{y})^2 . f_{\bullet j} = S_Y^2$$

Del momento m_{11} se dice que es la **covarianza,** y se denota S_{XY}:

$$S_{XY} \equiv m_{11} = \sum_{i=1}^{k} \sum_{j=1}^{m} (x_i - \overline{x}) . (y_j - \overline{y}) . f_{ij}$$

El signo de la covarianza S_{XY} indica el sentido de variación conjunta de las variables. Si $S_{XY} > 0$, las dos variables varían en el mismo sentido (en promedio) .... y si $S_{XY} < 0$, las dos variables varían en sentido contrario(en promedio)

Es:

$$S_{XY} \equiv m_{11} = \frac{1}{N}.\sum_{i=1}^{k}\sum_{j=1}^{m}(x_i - \overline{x}).(y_j - \overline{y}).n_{ij} =$$

$$= \frac{1}{N}.\sum_{i=1}^{k}\sum_{j=1}^{m}(x_i.y_j - x_i.\overline{y} - \overline{x}.y_j + \overline{x}.\overline{y}).n_{ij} =$$

$$= \left(\frac{1}{N}.\sum_{i=1}^{k}\sum_{j=1}^{m}x_i.y_j.n_{ij}\right) - \overline{y}.\left(\frac{1}{N}.\sum_{i=1}^{k}x_i.\left(\sum_{j=1}^{m}n_{ij}\right)\right) -$$

$$-\overline{x}.\left(\frac{1}{N}.\sum_{j=1}^{m}y_j.\left(\sum_{ij=1}^{k}n_{ij}\right)\right) + \overline{x}.\overline{y}.\sum_{i=1}^{k}\sum_{j=1}^{m}n_{ij} =$$

$$= \underbrace{\left(\frac{1}{N}.\sum_{i=1}^{k}\sum_{j=1}^{m}x_i.y_j.n_{ij}\right)}_{a_{11}} - \overline{y}.\underbrace{\left(\frac{1}{N}.\sum_{i=1}^{k}x_i.n_{i\bullet}\right)}_{\overline{x}} - \overline{x}.\underbrace{\left(\frac{1}{N}.\sum_{j=1}^{m}y_j.n_{\bullet j}\right)}_{\overline{y}} + \overline{x}.\overline{y} =$$

$$= a_{11} - \overline{x}.\overline{y} = a_{11} - a_{10}.a_{01}$$

$$\boxed{\overline{x} = a_{10} \;\; ; \;\; \overline{y} = a_{01}}$$

Por ejemplo, en el caso de los dos dados, sabiendo ya que $\overline{x} = 3'25$ e $\overline{y} = 3'21$, es:

$$S_{XY} = a_{11} - \overline{x}.\overline{y} = a_{11} - 3'5.3'21 = 10'33 - 3'5.3'21 = -0'905$$

X / Y	1	2	3	4	5	6
1	0	1	0	0	0	3
2	1	0	1	0	1	0
3	1	0	0	2	0	0
4	2	3	0	0	1	1
5	2	1	1	0	0	1
6	0	1	0	0	1	0

$\Rightarrow$

$$\Rightarrow a_{11} = \frac{1}{n}.\sum_{i=1}^{6}\sum_{j=1}^{6}x_i.y_j.n_{ij} =$$

$$= \frac{1.2.1 + 1.6.3 + 2.1.1 + 2.3.1 + 2.5.1 + 3.1.1 + 3.4.2 + 4.1.2 + 4.2.3 + 4.5.1}{24} +$$

$$+ \frac{4.6.1 + 5.1.2 + 5.2.1 + 5.3.1 + 5.6.1 + 6.2.1 + 6.5.1}{24} =$$

$$= \frac{248}{24} = 10'33$$

La covarianza ante los cambios de origen y escala

Si $U = \dfrac{X - A_1}{B_1}$ y $V = \dfrac{Y - A_2}{B_2}$, es:

$$S_{XY} \equiv m_{11} = \sum_{i=1}^{k}\sum_{j=1}^{m}(x_i - \overline{x}).(y_j - \overline{y}).f_{ij} =$$

$$\boxed{\begin{array}{l} u_i = \dfrac{x_i - A_1}{B_1} \Rightarrow \begin{cases} x_i = B_1.u_i + A_1 \\ \overline{x} = B_1.\overline{u} + A_1 \end{cases} \\ v_j = \dfrac{y_j - A_2}{B_2} \Rightarrow \begin{cases} y_j = B_2.v_j + A_2 \\ \overline{y} = B_2.\overline{v} + A_2 \end{cases} \end{array}}$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{m}(B_1.u_i + A_1 - B_1.\overline{u} - A_1).\left(B_2.v_j + A_2 - B_2.\overline{v} - A_2\right).f_{ij} =$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{m}(B_1.u_i - B_1.\overline{u}_1).\left(B_2.v_j - B_2.\overline{v}\right).f_{ij} =$$

$$= B_1.B_2.\sum_{i=1}^{k}\sum_{j=1}^{m}(u_i - \overline{u}_1).\left(v_j - \overline{v}\right).f_{ij} =$$

$$= B_1.B_2.S_{UV}$$

Por tanto, eligiendo adecuadamente el nuevo origen $(A_1; A_2)$ y los factores de escala B_1 y B_2, podemos aligerar notablemente el cálculo de la covarianza.

El coeficiente de correlación

$$r_{XY} = \frac{S_{XY}}{S_X.S_Y}$$

Si $U = \dfrac{X - A_1}{B_1}$ y $V = \dfrac{Y - A_2}{B_2}$, es:

$$r_{UV} = \frac{S_{UV}}{S_U.S_V} = \frac{\dfrac{S_{XY}}{B_1.B_2}}{\dfrac{S_X}{|B_1|}.\dfrac{S_Y}{|B_2|}} = \frac{|B_1|.|B_2|}{B_1.B_2}.\frac{S_{XY}}{S_X.S_Y} = \frac{|B_1|.|B_2|}{B_1.B_2}.r_{XY}$$

$$\boxed{S_{UV} = \frac{S_{XY}}{B_1.B_2} \;;\; S_U = \frac{S_X}{|B_1|} \;;\; S_V = \frac{S_Y}{|B_2|}}$$

Si B_1 y B_2 tienen igual signo, es $\dfrac{|B_1|.|B_2|}{B_1.B_2} = 1$, por lo que $r_{UV} = r_{XY}$.

Si B_1 y B_2 no tienen igual signo, es $\dfrac{|B_1|.|B_2|}{B_1.B_2} = -1$, por lo que $r_{UV} = -r_{XY}$.

3.6 INDEPENDENCIA ESTADÍSTICA

Se dice que la variable "Y" se distribuye independientemente de la variable "X" si sea cual sea el valor elegido para "j", la frecuencia de y_j condicionada a que $X = x_i$ es la misma para todo valor de "i"; o sea:

$$f(y_j / x_1) = f(y_j / x_2) = ... = f(y_j / x_k), \ \ \forall j = 1, 2, ... ,m$$

Propiedades que se derivan de la independencia

1) Si todas las frecuencias condicionadas de y_j son iguales, coinciden con la frecuencia marginal de y_j. En efecto:

$$\boxed{\text{si } f(y_j / x_i) \text{ es constante}}$$

$$f_{\bullet j} = \sum_{i=1}^{k} f_{ij} = \sum_{i=1}^{k} f(y_j / x_i).f_{i\bullet} = f(y_j / x_i).\sum_{i=1}^{k} f_{i\bullet} = f(y_j / x_i)$$

$$\boxed{f(y_j / x_i) = \frac{f_{ij}}{f_{i\bullet}} \Rightarrow f_{ij} = f(y_j / x_i).f_{i\bullet}} \quad \boxed{\sum_{i=1}^{k} f_{i\bullet} = 1}$$

Obvio: siendo siempre $f_{ij} = f(y_j / x_i).f_{i\bullet}$, si cuando las frecuencias condicionadas de y_j son iguales sucede que $f_{\bullet j} = f(y_j / x_i)$, también sucede que $f_{ij} = f_{\bullet j}.f_{i\bullet}$; o sea, la frecuencia bidimensional relativa es el producto de las correspondientes frecuencias marginales relativas.

Obvio: si cuando "Y" es independiente de "X" sucede que $f_{ij} = f_{i\bullet}.f_{\bullet j}$; se tiene que

$$f_{ij} = f_{i\bullet}.f_{\bullet j} \Rightarrow \frac{n_{ij}}{N} = \left(\frac{n_{i\bullet}}{N}\right).\left(\frac{n_{\bullet j}}{N}\right) \Rightarrow n_{ij} = \frac{n_{i\bullet}.n_{\bullet j}}{N}$$

$$\boxed{f_{i\bullet} = \frac{n_{i\bullet}}{N} \ ; \ f_{\bullet j} = \frac{n_{\bullet j}}{N}}$$

2) Si la variable "Y" se distribuye independientemente de la variable "X", la variable "X" se distribuye independientemente de "Y", pues por ser "Y" independiente de "X", es $f_{ij} = f_{i\bullet}.f_{\bullet j}$, y como siempre es $f_{ij} = f(x_i / y_j).f_{\bullet j}$, resulta que $f(x_i / y_j) = f_{i\bullet}$; es decir, sea cual sea el valor elegido para "i", la frecuencia de x_i condicionada a que $Y = y_j$ es la misma para todo valor de "j".

3) Si "Y" ("X") es independiente de "X" ("Y"), las medias de "Y" (de "X") condicionadas a los diversos valores de "X" (de "Y") coinciden entre sí .... y también coinciden con la media marginal de "Y" (de "X").

4) Todos los momentos respecto al origen o a la media de "Y" ("X") condicionadas a los diversos valores de "X" (de "Y") coinciden entre sí .... y también coinciden con los momentos marginales de "Y" (de "X").

5) La covarianza es 0:

$$S_{XY} = \sum_{i=1}^{k} \sum_{j=1}^{m} (x_i - \bar{x}).(y_j - \bar{y}).f_{ij} =$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{m} (x_i - \bar{x}).(y_j - \bar{y}).f_{i\bullet}.f_{\bullet j} =$$

$$= \underbrace{\left(\sum_{i=1}^{k} (x_i - \bar{x}).f_{i\bullet} \right)}_{0} . \underbrace{\left(\sum_{j=1}^{m} (y_j - \bar{y}).f_{\bullet j} \right)}_{0} = 0$$

Naturalmente, si la covarianza es 0, el coeficiente de correlación también.

Ojo!: la covarianza puede ser 0 sin que "X" e "Y" sean independientes.

El que "X" e "Y" sean independientes significa que cada una va a su bola; es decir, los valores de una no condicionan en modo alguno los valores de la otra. Por ejemplo, "X" expresa el peso de los estudiantes e "Y" la nota que obtienen en el examen de Matemáticas.