FONEMATO 1

Se toma una muestra de 150 pinos piñoneros, observando en cada árbol su altura "X" (en centímetros) y el número "Y" de nidos de abejaruco que lo pueblan.

X / Y	1	2	3	4	5
(50;150]	2	4	6	10	8
(150;250]	1	2	3	5	4
(250;350]	3	6	9	15	12
(350;450]	4	8	12	20	16

1) Determine la media y la varianza de "X" y de las distribuciones X / Y.
2) Determine la media y la varianza de "Y" y de las distribuciones Y / X.
3) Determine la covarianza.
4) Analice la independencia de las variables.

SOLUCIÓN

En cada intervalo de "X" consideramos la marca de clase:

X / Y	1	2	3	4	5	$n_{i\bullet}$
100	2	4	6	10	8	30
200	1	2	3	5	4	15
300	3	6	9	15	12	45
400	4	8	12	20	16	60
$n_{j\bullet}$	10	20	30	40	50	N = 150

1) La distribución marginal de "X" es

x_i	100	200	300	400
n_i	30	15	45	60

; así:

$$\bar{x} = \frac{100.30 + 200.15 + 300.45 + 400.60}{150} = \frac{43500}{150} = 290 \text{ cm}$$

$$S_x^2 = \frac{100^2.30 + 200^2.15 + 300^2.45 + 400^2.60}{150} - (290)^2 = 12900$$

- La distribución X / Y = 1 es

x_i	100	200	300	400
n_i	2	1	3	4

: así:

$$\bar{x}_{y=1} = \frac{100.2 + 200.1 + 300.3 + 400.4}{2+1+3+4} = \frac{2900}{10} = 290 \text{ cm}$$

$$S_{x/y=1}^2 = \frac{100^2.2 + 200^2.1 + 300^2.3 + 400^2.4}{2+1+3+4} - 290^2 = 12900$$

- La distribución X / Y = 2 es

x_i	100	200	300	400
n_i	4	2	6	8

; asÍ

$$\bar{x}_{y=2} = \frac{100.4 + 200.2 + 300.6 + 400.8}{4+2+6+8} = \frac{5800}{20} = 290 \text{ cm}$$

$$S_{x/y=2}^2 = \frac{100^2.4 + 200^2.2 + 300^2.6 + 400^2.8}{4+2+6+8} - 290^2 = 12900$$

- La distribución $X / Y = 3$ es

x_i	100	200	300	400
n_i	6	3	9	12

; así:

$$\bar{x}_{y=3} = \frac{100.6 + 200.3 + 300.9 + 400.12}{6+3+9+12} = \frac{8700}{30} = 290 \text{ cm}$$

$$S^2_{x/y=3} = \frac{100^2.6 + 200^2.3 + 300^2.9 + 400^2.12}{6+3+9+12} - 290^2 = 12900$$

- La distribución $X / Y = 4$ es

x_i	100	200	300	400
n_i	10	5	15	20

; así:

$$\bar{x}_{y=4} = \frac{100.10 + 200.5 + 300.15 + 400.20}{5+10+15+20} = \frac{14500}{50} = 290 \text{ cm}$$

$$S^2_{x/y=4} = \frac{100^2.10 + 200^2.5 + 300^2.15 + 400^2.20}{5+10+15+20} - 290^2 = 12900$$

- La distribución $X / Y = 5$ es

x_i	100	200	300	400
n_i	8	4	12	16

; así:

$$\bar{x}_{y=5} = \frac{100.8 + 200.4 + 300.12 + 400.16}{8+4+12+16} = \frac{11600}{40} = 290 \text{ cm}$$

$$S^2_{x/y=5} = \frac{100^2.8 + 200^2.4 + 300^2.12 + 400^2.16}{8+4+12+16} - 290^2 = 12900$$

2) La distribución marginal de "Y" es

y_j	1	2	3	4	5
n_j	10	20	30	40	50

; así:

$$\bar{y} = \frac{1.10 + 2.20 + 3.30 + 4.40 + 5.50}{150} = \frac{540}{150} = 3'6 \text{ nidos}$$

$$S^2_y = \frac{1^2.10 + 2^2.20 + 3^2.30 + 4^2.40 + 5^2.50}{150} - 3'6^2 = 1'44$$

- La distribución de $Y / X = 100$ es

y_j	1	2	3	4	5
n_j	2	4	6	10	8

; así:

$$\bar{y}_{x=100} = \frac{1.2 + 2.4 + 3.6 + 4.10 + 5.8}{2+4+6+10+8} = \frac{108}{30} = 3'6 \text{ nidos}$$

$$S^2_{y/x=100} = \frac{1^2.2 + 2^2.4 + 3^2.6 + 4^2.10 + 5^2.8}{2+4+6+10+8} - 3'6^2 = 1'44$$

- La distribución de $Y / X = 200$ es

y_j	1	2	3	4	5
n_j	1	2	3	5	4

; así:

$$\bar{y}_{x=200} = \frac{1.1 + 2.2 + 3.3 + 4.5 + 5.4}{1+2+5+3+4} = \frac{54}{15} = 3'6 \text{ nidos}$$

$$S^2_{y/x=200} = \frac{1^2.1 + 2^2.2 + 3^2.3 + 4^2.5 + 5^2.4}{1+2+5+3+4} - 3'6^2 = 1'44$$

- La distribución de $Y / X = 300$ es

y_j	1	2	3	4	5
n_j	3	6	9	15	12

; así:

$$\bar{y}_{x=300} = \frac{1.3+2.6+3.9+4.15+5.12}{3+6+9+15+12} = \frac{162}{45} = 3'6 \text{ nidos}$$

$$S^2_{y/x=300} = \frac{1^2.3+2^2.6+3^2.9+4^2.15+5^2.12}{3+6+9+15+12} - 3'6^2 = 1'44$$

- La distribución de $Y / X = 400$ es

y_j	1	2	3	4	5
n_j	4	8	12	20	16

; así:

$$\bar{y}_{x=400} = \frac{1.3+2.6+3.9+4.15+5.12}{4+8+12+20+16} = \frac{180}{50} = 3'6 \text{ nidos}$$

$$S^2_{y/x=400} = \frac{1^2.3+2^2.6+3^2.9+4^2.15+5^2.12}{4+8+12+20+16} - 3'6^2 = 1'44$$

3) Covarianza S_{XY}:

$$\boxed{\left.\begin{array}{l}\bar{x} = 290 \\ \bar{y} = 3'6\end{array}\right\} \Rightarrow \bar{x}.\bar{y} = 1044}$$

$$S_{XY} = a_{11} - \bar{x}.\bar{y} = a_{11} - 1044 = 1044 - 1044 = 0$$

X / Y	1	2	3	4	5
100	2	4	6	10	8
200	1	2	3	5	4
300	3	6	9	15	12
400	4	8	12	20	16

$\Rightarrow$

$$\Rightarrow a_{11} = \frac{1}{150}.\sum_{i=1}^{4}\sum_{j=1}^{5} x_i.y_j.n_{ij} = \frac{156600}{150} = 1044$$

$$\sum_{i=1}^{4}\sum_{j=1}^{5} x_i.y_j.n_{ij} = 100.1.2+100.2.4+100.3.6+100.4.10+100.5.8+$$
$$+200.1.1+200.2.2+200.3.3+200.4.5+200.5.4+$$
$$+300.1.3+300.2.6+300.3.9+300.4.15+300.5.12+$$
$$+400.1.4+400.2.8+400.3.12+400.4.20+400.5.16 = 156600$$

4) Las variables "X" e "Y" son independientes, pues para cualesquiera valores de $i = 1, 2, 3, 4$ y de $j = 1, 2, 3, 4, 5$ sucede que:

$$n_{ij} = \frac{n_{i\bullet}.n_{\bullet j}}{N}$$

FONEMATO 2

En 1000 operaciones de venta, un concesionario de Seat observa los siguientes datos relativos al color del coche (blanco, azul, verde, negro) y a la forma de pago (contado/financiado). Analice la independencia de las estas variables.

	Blanco	Azul	Verde	Negro
Contado	180	240	144	36
Financiado	120	160	96	24

SOLUCIÓN

Calculamos las distribuciones marginales de frecuencias absolutas:

	Blanco	Azul	Azul	Negro	$n_{i\bullet}$
Contado	180	240	144	36	600
Financiado	120	160	96	24	400
$n_{\bullet j}$	300	400	240	60	1000

Las variables "Color" y "Forma de Pago" son independientes, pues la tabla de correlación es tan mágica que:

$$n_{ij} = \frac{n_{i\bullet}.n_{\bullet j}}{N}$$

FONEMATO 3

En 1000 operaciones de venta, un concesionario de Renault observa los siguientes datos relativos al color del coche (blanco, azul, verde, negro) y a la forma de pago (contado/financiado). Analice la independencia de las estas variables.

	Blanco	Azul	Verde	Negro
Contado	181	240	144	36
Financiado	119	160	96	24

SOLUCIÓN

Calculamos las distribuciones marginales de frecuencias absolutas:

	Blanco	Azul	Azul	Negro	$n_{i\bullet}$
Contado	181	240	144	36	600
Financiado	119	160	96	24	400
$n_{\bullet j}$	300	400	240	60	1000

Las variables "Color" y "Forma de Pago" no son independientes, pues la tabla de correlación no sucede que $n_{ij} = \frac{n_{i\bullet}.n_{\bullet j}}{N}$. Por ejemplo:

$$n_{11} = 181 \neq \frac{600.300}{1000} = 180$$

FONEMATO 4

Sea la tabla:

X / Y	−2	1	3
1	2	1	2
2	4	2	4
3	6	3	6

1) Determine la distribución marginal de "Y".
2) Determine la distribución de frecuencias relativas de "X" si Y = 3.
3) Analice la independencia de las variables.

SOLUCIÓN

1)

X / Y	−2	1	3	$n_{i\bullet}$
1	2	1	2	5
2	4	2	4	10
3	6	3	6	15
$n_{\bullet j}$	12	6	12	N = 30

$\Rightarrow$

y_j	$n_{\bullet j}$
−2	12
1	6
3	12
	30

2)

X / Y	−2	1	3
1	2	1	$\boxed{2}$
2	4	2	$\boxed{4}$
3	6	3	$\boxed{6}$
$n_{\bullet j}$	12	6	12

$\Rightarrow$

X / Y = 3	Frecuencia absoluta	Frecuencia relativa
−2	2	2/12
1	4	4/12
3	6	6/12
	12	1

3) Las variables "X" e "Y" son independientes, pues la tabla de correlación es tan mágica que:

$$n_{ij} = \frac{n_{i\bullet} . n_{\bullet j}}{N}$$

Como "X" e "Y" son independientes, la covarianza es 0, lo mismo que el coeficiente de correlación.

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

$$= \underbrace{\left(\frac{1}{N} . \sum_{i=1}^{k} \sum_{j=1}^{m} x_i . y_j . n_{ij} \right)}_{a_{11}} - \underbrace{\left(\frac{1}{N} . \sum_{i=1}^{k} x_i . n_{i\bullet} \right)}_{\bar{x}} . \underbrace{\left(\frac{1}{N} . \sum_{j=1}^{m} y_j . n_{\bullet j} \right)}_{\bar{y}} = 0$$

FONEMATO 5

Siendo "X" la renta familiar e "Y" el gasto en alimentación, sea la tabla siguiente tabla de correlación:

X / Y	3	4	5	8
5	2	0	0	0
7	2	1	0	0
12	0	0	2	0
20	0	0	1	1
40	0	0	0	1

1) Determínese el gasto que realizan al menos la mitad de las familias.
2) Determínese la distribución de frecuencias relativas de "X" si $Y < 5$.
3) Determínese el momento a_{12}.

SOLUCIÓN

1) Piden la mediana de la distribución marginal de "Y".

X / Y	3	4	5	8	$n_{i\bullet}$
5	2	0	0	0	2
7	2	1	0	0	5
12	0	0	2	0	2
20	0	0	1	1	2
40	0	0	0	1	1
$n_{\bullet j}$	4	1	3	2	$N = 10$

$\Rightarrow$

y_j	$n_{\bullet j}$	$N_{\bullet j}$
3	4	4
4	1	5
5	3	8
8	2	10
	10	

La mediana de "Y" es "4", pues la frecuencia relativa acumulada correspondiente a "4" es la primera mayor o igual a $N/2 = 5$.

2)

X / Y	3	4	5	8
5	$\boxed{2}$	$\boxed{0}$	0	0
7	$\boxed{2}$	$\boxed{1}$	0	0
12	$\boxed{0}$	$\boxed{0}$	2	0
20	$\boxed{0}$	$\boxed{0}$	1	1
40	$\boxed{0}$	$\boxed{0}$	0	1

$\Rightarrow$

$X / Y < 5$	Frecuencia absoluta	Frecuencia relativa
5	2	2/5
7	3	3/5
12	0	0
20	0	0
40	0	0
	5	1

3) Es:

$$a_{12} = \frac{1}{N}.\sum \sum x_i . y_j^2 . n_{ij} =$$

$$= \frac{5.3^2.2 + 7.3^2.2 + 7.4^2.1 + 12.5^2.2 + 20.5^2.1 + 20.8^2.1 + 40.8^2.1}{10} = 491'8$$

FONEMATO 6

Siendo "X" la renta familiar e "Y" el gasto en alimentación, sea la tabla siguiente tabla de correlación:

X / Y	30	35	40	50	60
50	6	5	2	0	0
60	10	7	3	1	0
70	25	20	13	6	1
80	23	19	15	8	3
90	10	10	7	5	1

1) Determínese el gasto medio de las familias con renta 80.
2) Determínese la varianza de los ingresos de las familias que gastan 50.

SOLUCIÓN

1)

X / Y	30	35	40	50	60
50	6	5	2	0	0
60	10	7	3	1	0
70	25	20	13	6	1
80	23	19	15	8	3
90	10	10	7	5	1

$\Rightarrow$

y_j	Frecuencia absoluta
30	23
35	19
40	15
50	8
60	3
	68

$\Rightarrow$

$$\Rightarrow \bar{y}_{x=80} = \frac{30.23 + 35.19 + 40.15 + 50.8 + 60.3}{68} = 37'27$$

2)

X / Y	30	35	40	50	60
50	6	5	2	0	0
60	10	7	3	1	0
70	25	20	13	6	1
80	23	19	15	8	3
90	10	10	7	5	1

$\Rightarrow$

x_j	Frecuencia absoluta
50	0
60	1
70	6
80	8
90	5
	20

$\Rightarrow$

$$\bar{x}_{y=50} = \frac{50.0 + 60.1 + 70.6 + 80.8 + 90.5}{20} = \frac{1570}{20}$$

$$S^2_{x/y=50} = \frac{50^2.0 + 60^2.1 + 70^2.6 + 80^2.8 + 90^2.5}{20} - \left(\frac{1570}{20}\right)^2 = 72'75$$

FONEMATO 7

Sea la tabla:

X / Y	−1	0	1
0	3	4	a
1	1	b	3

1) Calcule "a" y "b" si la distribución marginal de frecuencias relativas de "Y" es:

y_j	−1	0	1
$f_{\bullet j}$	0'25	0'3125	0'4375

2) Si $a = 2$ y $b = 3$, calcule la covarianza y la media de "X" condicionada a $Y \leq 0$.

SOLUCIÓN

1)

X / Y	−1	0	1	
0	3	4	a	
1	1	b	3	
$n_{\bullet j}$	4	4 + b	a + 3	N = 11 + a + b

$$\Rightarrow \left\{ \begin{array}{l} \dfrac{4}{11+a+b} = 0'25 \\ \dfrac{4+b}{11+a+b} = 0'3125 \\ \dfrac{a+3}{11+a+b} = 0'4375 \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} a = 4 \\ b = 1 \end{array} \right.$$

2) Si $a = 2$ y $b = 3$, es:

$$S_{XY} = a_{11} - \bar{x}.\bar{y} = \frac{2}{16} - \frac{6}{16}.\frac{2}{16} = 0'07812$$

X / Y	−1	0	1	$n_{i\bullet}$
0	3	4	3	10
1	1	2	3	6
$n_{\bullet j}$	4	6	6	N = 16

$$\bar{x} = 0.\frac{9}{16} + 1.\frac{6}{16} = \frac{6}{16} \quad ; \quad \bar{y} = (-1).\frac{4}{16} + 0.\frac{6}{16} + 1.\frac{6}{16} = \frac{2}{16}$$

$$a_{11} = \frac{1}{16}.\sum_{i=1}^{2}\sum_{j=1}^{3} x_i.y_j.n_{ij} = \frac{2}{16}$$

$$\sum_{i=1}^{2}\sum_{j=1}^{3} x_i.y_j.n_{ij} = 0.(-1).3 + 0.0.4 + 0.1.3 + 1.(-1).1 + 1.0.0 + 1.1.3 = 2$$

• Se tiene que:

X / Y	−1	0	1
0	$\boxed{3}$	$\boxed{4}$	3
1	$\boxed{1}$	$\boxed{2}$	3

$\Rightarrow$

X / Y ≤ 0	Frecuencia absoluta
0	7
1	3

$\Rightarrow$

$$\Rightarrow \bar{x}_{y \leq 0} = 0.\frac{7}{10} + 1.\frac{3}{10} = 0'3$$

FONEMATO 8

Siendo $a > 0$, sea la tabla:

X / Y	0	1
0	a	5
1	3	2

Calcule la covarianza y determine "a" para que "X" e "Y" sean independientes.

SOLUCIÓN

$$S_{XY} = a_{11} - \bar{x}.\bar{y} = \frac{2}{a+10} - \frac{5}{a+10}.\frac{7}{a+10}$$

X / Y	0	1	$n_{i\bullet}$
0	a	5	a + 5
1	3	2	5
$n_{\bullet j}$	a + 3	7	a + 10

$$\bar{x} = 0.\frac{a+5}{a+10} + 1.\frac{5}{a+10} = \frac{5}{a+10}$$

$$\bar{y} = 0.\frac{a+3}{a+10} + 1.\frac{7}{a+10} = \frac{7}{a+10}$$

$$a_{11} = \frac{1}{a+10}.\sum_{i=1}^{2}\sum_{j=1}^{2} x_i.y_j.n_{ij} = \frac{2}{a+10}$$

$$\boxed{\sum_{i=1}^{2}\sum_{j=1}^{2} x_i.y_j.n_{ij} = 0.0.a + 0.1.5 + 1.0.3 + 1.1.2 = 2}$$

Las variables "X" e "Y" son independientes si la tabla es tan mágica que:

$$n_{ij} = \frac{n_{i\bullet}.n_{\bullet j}}{N} \Rightarrow$$

$$\Rightarrow n_{ij}.N = n_{i\bullet}.n_{\bullet j} \Rightarrow \begin{cases} a.(a+10) = (a+5).(a+3) \Rightarrow a = 15/2 \\ 5.(a+10) = 7.(a+5) \Rightarrow a = 15/2 \\ 3.(a+10) = 5.(a+3) \Rightarrow a = 15/2 \\ 2.(a+10) = 35 \Rightarrow a = 15/2 \end{cases}$$

FONEMATO 9

Demuéstrese que si "X" e "Y" son incorreladas son independientes.

X / Y	0	1
0	a	b
1	c	d

SOLUCIÓN

Por comodidad, sin perder generalidad, consideramos que las frecuencias dadas son relativas (por tanto, es $a + b + c + d = 1$), pues si fueran absolutas, dividiéndolas $a + b + c + d$ obtendríamos la siguiente tabla de frecuencias relativas:

X / Y	0	1
0	a'	b'
1	c'	d'

Es:

$$S_{XY} = a_{11} - \bar{x}.\bar{y} = d - (b+d).(c+d)$$

X / Y	0	1	$f_{\bullet j}$
0	a	b	a + b
1	c	d	c + d
$f_{i\bullet}$	a + c	b + d	

$$\Rightarrow \begin{cases} \bar{x} = 0.(a+b) + 1.(c+d) = c+d \\ \bar{y} = 0.(a+c) + 1.(b+d) = b+d \end{cases}$$

$$a_{11} = \sum_{i=0}^{1} \sum_{j=0}^{1} z_i.x_j.f_{ij} = 0.0.a + 0.1.c + 1.0.b + 1.1.d = d$$

Por tanto, "X" e "Y" son incorreladas (tienen covarianza nula) sólo si

$$a+b+c+d = 1 \quad \text{(I)}$$
$$d - (b+d).(c+d) = 0 \quad \text{(II)}$$

Las variables "X" e "Y" son independientes si $f_{ij} = f_{i\bullet}.f_{\bullet j}$ Comprobemos que ocurre tal cosa si se satisfacen (I) y (II):

- $$f_{0\bullet}.f_{\bullet 0} = (a+b).(a+c) = (1-(c+d)).(1-(b+d)) =$$
  $$\boxed{a+b+c+d = 1 \Rightarrow \begin{cases} a+b = 1-(c+d) \\ a+c = 1-(b+d) \end{cases}}$$
  $$= 1-(b+d)-(c+d)+(c+d).(b+d) =$$
  $$= 1-c-d-b-d+(b+d).(c+d) = a = f_{00}$$
  $$\boxed{(I) \Rightarrow 1-c-d-b = a \;\; ; \;\; (II) \Rightarrow -d+(b+d).(c+d) = 0}$$
- $$f_{1\bullet}.f_{\bullet 0} = (c+d).(a+c) = (c+d).(1-(b+d)) =$$
  $$\boxed{a+b+c+d = 1 \Rightarrow a+c = 1-(b+d)}$$
  $$= c+d-(c+d).(b+d) = c+d-d = c = f_{10}$$
  $$\boxed{(II) \Rightarrow (b+d).(c+d) = d}$$
- $$f_{0\bullet}.f_{\bullet 1} = (a+b).(b+d) = (1-(c+d)).(b+d) =$$
  $$\boxed{a+b+c+d = 1 \Rightarrow a+b = 1-(c+d)}$$
  $$= (b+d)-(c+d).(b+d) = b+d-d = b = f_{01}$$
  $$\boxed{(II) \Rightarrow (b+d).(c+d) = d}$$
- $$f_{1\bullet}.f_{\bullet 1} = (c+d).(b+d) = d = f_{11}$$
  $$\boxed{(II) \Rightarrow (b+d).(c+d) = d}$$

FONEMATO 10

De la tabla de correlación entre "X" e "Y" se obtiene:

$$\sum_i x_i.n_{i\bullet} = 500 \ ; \ \sum_j y_j.n_{\bullet j} = 1000 \ ; \ \sum_i x_i^2.n_{i\bullet} = 3000 \ ; \ \sum_j y_j^2.n_{\bullet j} = 20000$$

$$\sum_i \sum_j x_i.y_j.n_{ij} = 6000 \ ; \ N = 100$$

Determínese la covarianza y el coefciente de correlación entre las variables "U" y "V" tales que $2.X = 3.U + 4$ y $2.Y = 2.V + 3$.

SOLUCIÓN

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

$$S_{UV} = \frac{S_{XY}}{B_1.B_2} \ ; \ r_{UV} = \frac{|B_1|.|B_2|}{B_1.B_2}.r_{XY}$$

En nuestro caso es $B_1 = 3/2$ y $B_2 = 1$:

$$\left.\begin{array}{l} 2.X = 3.U + 4 \\ 2.Y = 2.V + 3 \end{array}\right\} \Rightarrow \left\{\begin{array}{l} U = \dfrac{2.X - 4}{3} = \dfrac{X - 2}{3/2} \Rightarrow B_1 = 3/2 \\ V = \dfrac{2.Y - 3}{2} = \dfrac{Y - (3/2)}{1} \Rightarrow B_2 = 1 \end{array}\right.$$

Es:

$$S_{UV} = \frac{S_{XY}}{B_1.B_2} = \frac{S_{XY}}{3/2} = \frac{m_{11} - \overline{x}.\overline{y}}{3/2} = \frac{60 - 5.10}{3/2} = \frac{20}{3}$$

$$\boxed{\begin{array}{c} m_{11} = \dfrac{1}{N}.\displaystyle\sum_i \sum_j x_i.y_j.n_{ij} = \dfrac{6000}{100} = 60 \\ \overline{x} = \dfrac{1}{N}.\displaystyle\sum_i x_i.n_{i\bullet} = \dfrac{500}{100} = 5 \ ; \ \overline{y} = \dfrac{1}{N}.\displaystyle\sum_j y_j.n_{\bullet j} = \dfrac{1000}{100} = 10 \end{array}}$$

Es:

$$\boxed{\frac{|B_1|.|B_2|}{B_1.B_2} = 1, \text{ pues } B_1 = 3/2 \text{ y } B_2 = 1 \text{ tienen igual signo}}$$

$$r_{UV} = \frac{|B_1|.|B_2|}{B_1.B_2}.r_{XY} = r_{XY} = \frac{S_{XY}}{S_X.S_Y} = \frac{60 - 5.10}{\sqrt{5}.10} = \frac{1}{\sqrt{5}}$$

$$\boxed{\begin{array}{l} S_X = +\sqrt{\left(\dfrac{1}{N}.\displaystyle\sum_i x_i^2.n_{i\bullet}\right) - (\overline{x})^2} = \sqrt{\dfrac{3000}{100} - 5^2} = \sqrt{5} \\ S_Y = +\sqrt{\left(\dfrac{1}{N}.\displaystyle\sum_j y_j^2.n_{\bullet j}\right) - (\overline{y})^2} = \sqrt{\dfrac{20000}{100} - 10^2} = 10 \end{array}}$$

FONEMATO 11

De la tabla de correlación entre "X" e "Y" se conocen las frecuencias absolutas de la distribución marginal de "X" y las frecuencias relativas de las distribuciones condicionadas de "Y" para cada valor de "X":

X	1	2	3
Frec. abs.	10	20	15

;

Y / X = 1	2	4	6
Frec. rel.	0'2	0'5	0'3

Y / X = 2	2	4	6
Frec. rel.	0'45	03	0'25

;

Y / X = 3	2	4	6
Frec. rel.	0'4	0'6	0

1) Determine la tabla de correlacion de frecuencias absolutas y la marginal de "Y"
2) Calcule la media y la varianza de $X / Y = 2$.
3) Si $Z = 2.X - 1$, calcule la media y la varianza de $Z / Y = 2$.

SOLUCIÓN

1) Tabla de correlacion de frecuencias absolutas:

X / Y	2	4	6
1	2	5	3
2	9	6	5
3	6	9	0

Y / X = 1	2	4	6
Frec. rel.	0'2	0'5	0'3

$\Rightarrow$

Y / X = 1	2	4	6
Frec. abs.	2	5	3

la frecuencia absoluta de $X = 1$ es 10

Y / X = 2	2	4	6
Frec. rel.	0'45	03	0'25

$\Rightarrow$

Y / X = 2	2	4	6
Frec. abs.	9	6	5

la frecuencia absoluta de $X = 2$ es 20

Y / X = 3	2	4	6
Frec. rel.	0'4	0'6	0

$\Rightarrow$

Y / X = 3	2	4	6
Frec. abs.	6	9	0

la frecuencia absoluta de $X = 3$ es 15

2)

X / Y	2	4	6
1	2	5	3
2	9	6	5
3	6	9	0

$\Rightarrow$

X / Y = 2	1	2	3
Frec. abs.	2	9	6

$\Rightarrow$

$$\Rightarrow \begin{cases} \text{Media} = \dfrac{1.2 + 2.9 + 3.6}{2 + 9 + 6} = \dfrac{38}{17} = 2'23 \\ \text{Varianza} = \dfrac{1^2.2 + 2^2.9 + 3^2.6}{2 + 9 + 6} - \left(\dfrac{38}{17}\right)^2 = 0'438 \end{cases}$$

3) Denotando "W" a la variable $X / Y = 2$, si $Z = 2.W - 1$, es:

$$\overline{z} = 2.\overline{w} - 1 = 2.2'23 - 1 = 3'46$$

$$S_Z^2 = 2^2.S_W^2 = 2^2.0'438 = 1'752$$

FONEMATO 12

Sea la tabla:

X / Y	1	3	5
0 − 4	3	1	2
4 − 6	0	2	1
6 − 10	1	2	3

Determine la distribución de frecuencias relativas de "X" si $Y \leq 3$, el momento a_{12} y el intervalo que incluye al 40% central de la distribución de "X".

SOLUCIÓN

• Frecuencia relativa de $X / Y \leq 3$:

X / Y	1	3	5
0 − 4	$\boxed{3}$	$\boxed{1}$	2
4 − 6	$\boxed{0}$	$\boxed{2}$	1
6 − 10	$\boxed{1}$	$\boxed{2}$	3

$\Rightarrow$

$X / Y \leq 3$	0 − 4	4 − 6	6 − 10
Frec. abs.	$3+1=4$	$0+2=2$	$1+2=3$

$\Rightarrow$

$\Rightarrow$

$X / Y \leq 3$	0 − 4	4 − 6	6 − 10
Frec. rel.	$\frac{4}{4+2+3}$	$\frac{2}{4+2+3}$	$\frac{3}{4+2+3}$

• Momento a_{12}:

$$a_{12} = \frac{1}{N}.\sum_{i}\sum_{j} x_i.y_j^2.n_{ij} =$$

$\boxed{x_i \equiv \text{marca de clase del i - ésimo intervalo de "X"}}$

$$= \frac{2.1^2.3+2.3^2.1+2.5^2.2+5.1^2.0+5.3^2.2+5.5^2.1+8.1^2.1+8.3^2.2+8.5^2.3}{15}$$

• La distribución de "X" es:

X	0 − 4	4 − 6	6 − 10
n_i	6	3	6
N_i	6	9	15

Los percentiles P_{30} y P_{70} son los extremos del intervalo que incluye al 40% central de la distribución de "X".

• El percentil P_{30} está en el intervalo $0-4$, que es el primero con frecuencia absoluta acumulada mayor o igual que $30.N/100 = 30.15/100 = 4'5$:

$$P_{30} = L_{i-1} + \frac{4'5 - N_{i-1}}{n_i}.(L_i - L_{i-1}) = 0 + \frac{4'5-0}{6}.(4-0) = 3$$

• El percentil P_{85} está en el intervalo $6-8$, que es el primero con frecuencia absoluta acumulada mayor o igual que $70.N/100 = 70.15/100 = 10'5$:

$$P_{70} = L_{i-1} + \frac{10'5 - N_{i-1}}{n_i}.(L_i - L_{i-1}) = 6 + \frac{10'5-9}{6}.(10-6) = 7$$

FONEMATO 13

La tabla recoge la información relativa a la plantilla de una empresa, en lo que se refiere a la edad de los trabajadores y a su antigüedad en la empresa.

EDAD "X"	ANTIGÜEDAD "Y"										
	1	2	5	6	7	8	10	12	14	15	18
18–30	2	3		8		3					
30–40			1	5	16		2				
40–50					6	13		2	2		
50–64									6	8	2

1) Determine los cuartiles primero y tercero de la distribución de antigüedades de los trabajadores con edad superior a 40.
2) Determine la media y la varianza de la distribución antigüedades de los trabajadores con edad entre 30 y 40.

SOLUCIÓN

1)

Antigüedad	7	8	10	12	14	15	18
Frec abs.	6	13	0	5	10	8	2
Frec. abs. acum.	6	19	19	24	34	42	44

Como $N/4 = 44/4 = 11$ y $3.N/4 = 3.44/4 = 33$, es $C_1 = 8$ y $C_3 = 14$.

2)

Antigüedad	5	6	7	10
Frec abs.	1	5	16	2

$\Rightarrow$

$$\Rightarrow \begin{cases} \text{Media} = \dfrac{5.1+6.5+7.16+10.2}{1+5+16+2} = \dfrac{167}{24} \\ \text{Varianza} = \dfrac{5^2.1+6^2.5+7^2.16+10^2.2}{1+5+16+2} - \left(\dfrac{167}{24}\right)^2 \end{cases}$$

FONEMATO 14

Sea la tabla:

X / Y	0 – 2	2 – 4	4 – 6
5 – 10	20	0	0
10 – 20	5	35	5
20 – 40	0	15	20

1) Determine las medias aritmética, geométrica y armónica de "Y", y S_Y^2.
2) Determine mediana y la moda de "X".
3) Determine m_{11}.

SOLUCIÓN

1) La distribución marginal de "Y" es

Y	0 – 2	2 – 4	4 – 6
Frec abs.	25	50	25

Tomamos la marca de clase de cada intervalo como representativa él.

$$\bar{y} = \frac{1.25 + 3.50 + 5.25}{25 + 50 + 25} = 3 \; ; \; G = \sqrt[100]{1^{25}.3^{50}.5^{25}} = 2'59$$

$$H = \frac{25 + 50 + 25}{\frac{25}{1} + \frac{50}{3} + \frac{25}{5}} = 2'14 \; ; \; S_Y^2 = \frac{1^2.25 + 3^2.50 + 5^2.25}{25 + 50 + 25} - 3^2 = 2$$

2) La marginal de "X" es

X	5 – 10	10 – 20	20 – 40
n_i	20	45	35
N_i	20	65	100
d_i	4	4'5	1'75

- La mediana está en el intervalo 10 – 20, que es el primero con frecuencia absoluta acumulada mayor o igual que $N/2 = 100/2 = 50$.

$$Me = L_{i-1} + \frac{50 - N_{i-1}}{n_i}.(L_i - L_{i-1}) = 10 + \frac{50 - 20}{45}.(20 - 10) = 16'67$$

- La moda está en el intervalo 10 – 20, el de mayor densidad de frecuencia.

$$Mo = L_{i-1} + \frac{d_{i+1}}{d_{i-1} + d_{i+1}}.(L_i - L_{i-1}) = 10 + \frac{1'75}{1'75 + 4}.(20 - 10) = 13'04$$

3)

$$m_{11} \equiv S_{XY} = a_{11} - \bar{x}.\bar{y} = 65'25 - 18'75.3 = 9$$

$$\bar{y} = 3 \; ; \; \bar{x} = \frac{7'5.20 + 15.45 + 30.35}{20 + 45 + 35} = 18'75$$

$$a_{11} = \frac{1}{100}.\sum_i \sum_j x_i.y_j.n_{ij} = \frac{6525}{100} = 65'25$$

$$\sum_i \sum_j x_i.y_j.n_{ij} = 7'5.1.20 + 15.1.5 + 15.3.35 + $$

$$+15.5.5 + 30.3.15 + 30.5.20 = 6525$$

FONEMATO 15

En un tipo de empresas, siendo "X" el número de empleados e "Y" la producción, se ha obtenido la información que recoge la tabla:

Y / X	0 – 1	1 – 100	100 – 1000
1 – 10	10	0	0
10 – 50	2	3	1
50 – 100	0	4	3
> 100	0	1	1

1) Determine la distribución (frecuencias absolutas y relativas) de la producción en las empresas de no más de 100 trabajadores.
2) Determine el índice de Gini de "Y" y representa la curva de Lorenz.
3) Determine la varianza y el coeficiente de variación de "X".
4) Entre el 40% las empresas que más producen, ¿cuál es la producción mínima?

SOLUCIÓN

1)

Y / X	0 – 1	1 – 100	100 – 1000
1 – 10	$\boxed{10}$	$\boxed{0}$	$\boxed{0}$
10 – 50	$\boxed{2}$	$\boxed{3}$	$\boxed{1}$
50 – 100	$\boxed{0}$	$\boxed{4}$	$\boxed{3}$
> 100	0	1	1

$\Rightarrow$

$\Rightarrow$

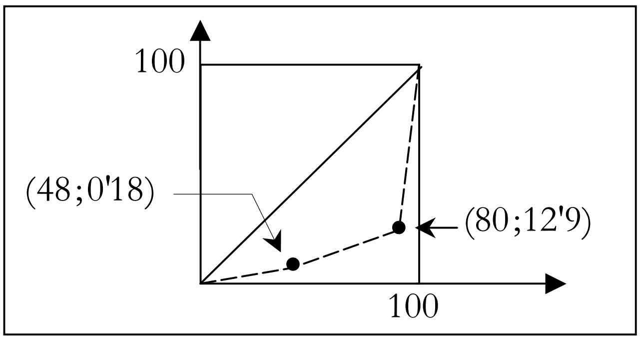
$X / Y \le 100$	0 – 1	1 – 100	100 – 1000
Frec. abs.	12	7	4
Frec. rel.	$\frac{12}{12+7+4}$	$\frac{7}{12+7+4}$	$\frac{4}{12+7+4}$

2)

x_i	n_i	N_i	$x_i.n_i$	u_i	$p_i = \frac{N_i}{N}.100$	$q_i = \frac{u_i}{u_k}.100$	$p_i - q_i$
0'5	12	12	6	6	48	0'18	47'82
50'5	8	20	404	410	80	12'9	67'1
550	5	25	2750	3160	×	×	×
					128		114'92

$\boxed{u_k}$

$$I_G = \frac{\sum_{i=1}^{k-1} (p_i - q_i)}{\sum_{i=1}^{k-1} p_i} = \frac{114'92}{128} = 0'89 \Rightarrow \text{alta concentración}$$



3) La distribución marginal de "X" es:

X	$0-1$	$1-100$	$100-1000$
Marca	0'5	50'5	550
Frec. abs.	12	8	5

$\Rightarrow$

$$\Rightarrow \left\{ \begin{array}{c} \bar{y} = \dfrac{0'5.12 + 50'5.8 + 550.5}{12+8+5} = 126'4 \\ S_Y^2 = \dfrac{0'5^2.12 + 50'5^2.8 + 550^2.5}{12+8+5} - 126'4^2 = 45339'24 \end{array} \right\} \Rightarrow$$

$$\Rightarrow C_V = \frac{\sqrt{45339'24}}{126'4}$$

4) El percentil 60 de "X" está en el intervalo $1-100$, que es el primero con frecuencia absoluta acumulada mayor o igual que $60.N/100 = 60.25/100 = 15$.

$$P_{60} = L_{i-1} + \frac{15 - N_{i-1}}{n_i}.(L_i - L_{i-1}) =$$

$$= 1 + \frac{15-12}{8}.(100-1) = 38'125$$

FONEMATO 16

En un tipo de empresas, siendo "X" el número de empleados e "Y" la producción, se ha obtenido la información que recoge la tabla:

Y / X	0 – 10	10 – 20	20 – 40	40 – 100
2 – 6	15	0	0	0
6 – 14	0	10	0	0
14 – 26	0	10	5	0
26 – 54	0	0	0	10

En la distribución de trabajadores, determine la moda y el intervalo que contiene el 40% central.

SOLUCIÓN

X	0 – 10	10 – 20	20 – 40	40 – 100
n_i	15	20	5	10
N_i	15	35	40	50
$d_i = \dfrac{n_i}{L_i - L_{i-1}}$	1'5	2	0'25	0'16

El intervalo de mayor densidad de frecuencia $d_i = n_i/(L_{i-1} - L_i)$, es el 2º:

$$Mo = L_{i-1} + \frac{d_{i+1}}{d_{i-1} + d_{i+1}}.(L_i - L_{i-1}) =$$

$$= 10 + \frac{0'25}{1'5 + 0'25}.(20 - 10)$$

El percentil 30 de "X" está en el intervalo $1 - 100$, que es el primero con frecuencia absoluta acumulada mayor o igual que $30.N/100 = 30.50/100 = 15$.

$$P_{30} = L_{i-1} + \frac{15 - N_{i-1}}{n_i}.(L_i - L_{i-1}) = 0 + \frac{15 - 0}{15}.(10 - 0) = 15$$

El percentil 70 de "X" está en el intervalo $1 - 100$, que es el primero con frecuencia absoluta acumulada mayor o igual que $70.N/100 = 70.50/100 = 35$.

$$P_{70} = L_{i-1} + \frac{35 - N_{i-1}}{n_i}.(L_i - L_{i-1}) = 10 + \frac{35 - 15}{20}.(20 - 10) = 20$$

FONEMATO 17

Se hace un estudio sobre 50 empresas invierten en publicidad por primera vez, observando la inversión "X" en publicidad y el porcentaje "Y" de variación de beneficio, con los siguientes resultados:

X / Y	0 – 20	20 – 50	50 – 100	100 – 150
20 – 40	15	2	0	0
40 – 100	0	1	10	0
100 – 150	0	0	3	8
150 – 200	0	0	1	10

1) Determine la mediana de "Y". ¿Qué inversión en publicidad es más frecuente?

2) Calcular la media y la varianza de la inversión en las empresas que obtuvieron variación de beneficios entre el 50 y el 100%. ¿Cuál es la inversión mínima del 20% de las empresas con mayor inversión en publicidad?

SOLUCIÓN

1) La mediana de "Y" está en el intervalo $50-100$, que es el primero con frecuencia absoluta acumulada mayor o igual que $N/2 = 50/2 = 25$.

$$Me = L_{i-1} + \frac{25 - N_{i-1}}{n_i}.(L_i - L_{i-1}) = 50 + \frac{25-18}{14}.(100-50) = 75$$

Y	0 – 20	20 – 50	50 – 100	100 – 150
Frec. abs.	15	3	14	18
Frec. abs. acum.	15	18	32	50

- La moda de "X" está en $20-40$, por ser el de mayor densidad de frecuencia.

$$Mo = L_{i-1} + \frac{d_{i+1}}{d_{i-1} + d_{i+1}}.(L_i - L_{i-1}) = 20 + \frac{0'18}{0 + 0'18}.(40-20) = 40$$

X	20 – 40	40 – 100	100 – 150	150 – 200
n_i	17	11	11	11
d_i	0'85	0'18	0'22	0'22

2)

	40 – 100	100 – 150	150 – 200
Feec. abs.	10	3	1

$\Rightarrow$

$$\Rightarrow \begin{cases} \text{Media} = \dfrac{70.10 + 125.3 + 175.1}{10+3+1} = \dfrac{1250}{14} \\ \text{Varianza} = \dfrac{70^2.10 + 125^2.3 + 175^2.1}{10+3+1} - \left(\dfrac{1250}{14}\right)^2 \end{cases}$$

- El percentil 80 de "X" está en el intervalo $150-200$, que es el primero con frecuencia absoluta acumulada mayor o igual que $80.N/100 = 80.50/100 = 40$.

$$P_{80} = L_{i-1} + \frac{40 - N_{i-1}}{n_i}.(L_i - L_{i-1}) = 150 + \frac{40-39}{11}.(200-150)$$

X	20 – 40	40 – 100	100 – 150	150 – 200
n_i	17	11	11	11
N_i	17	28	39	50