

Relación nº 2: Estadística Descriptiva Bidimensional

1.-Se han estudiado los pesos X (en Kg.) y las tallas Y (en cm.) de un grupo de individuos, obteniendo la siguiente información:

X \ Y	160	162	164	166	168	170
48	3	2	2	1	0	0
51	2	3	4	2	2	1
54	1	3	6	8	5	1
57	0	0	1	2	8	3
60	0	0	0	2	4	4

- Calcular el peso medio y la talla media.
- Calcular el porcentaje de individuos que pesan menos de 55 Kg. y miden más de 165 cm.
- Dentro del conjunto de los que miden más de 165 cm. ¿cuál es el porcentaje de los que pesan más de 52 Kg.?
- ¿Cuál es la altura más frecuente entre los individuos cuyo peso oscila entre 51 y 57 Kg.?
- ¿Qué peso medio es más representativo, el de los individuos que miden 164 cm., o el de los que miden 168 cm.?

2.-Se han lanzado dos dados varias veces. Designamos con X el resultado del primer dado, y con Y el resultado del segundo. La información obtenida se recoge en la siguiente tabla:

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

- Construir la tabla bidimensional de correlación.
- Calcular las puntuaciones medias obtenidas con cada dado.
- ¿Qué puntuaciones son más homogéneas, las obtenidas con el primer dado o con el segundo?
- Calcular la puntuación más frecuente obtenida con el segundo dado, si con el primero se obtuvo un 3.
- Calcular la puntuación máxima del 50% de las puntuaciones más bajas obtenidas con el primer dado, si con el segundo se obtuvo un 2 o un 5.

3.-Se han medido dos caracteres simultáneamente sobre cada uno de los miembros de un colectivo, obteniéndose los datos de la tabla:

X \ Y	90-100	100-120	120-140
10-15	6	3	1
15-20	5	10	2
20-25	4	1	7
25-30	2	2	4

- Calcular la media de Y, si X toma valores superiores a 20.
- Calcular la media de X, si Y toma valores entre 100 y 120.
- Calcular las medias y las varianzas marginales.
- En el conjunto de los individuos que presentan valores de Y inferiores a 120, calcular el valor de X mínimo del 30% de las observaciones más altas.
- En el conjunto de los individuos que presentan valores de X entre 15 y 25, calcular el porcentaje de éstos que presentan valores superiores a 117.

4.-Con los datos de la siguiente tabla:

Y \ X	1	2	4
5	1	0	2
10	2	1	0
15	0	1	3

- Obtener las medias marginales.
- Obtener las varianzas marginales.
- Obtener la covarianza. Interpretar su valor.

5.-Se considera la siguiente distribución bidimensional:

Y \ X	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

- Calcular las medias marginales.
- ¿Qué distribución está más agrupada en torno a su media?
- Calcular las medias condicionadas. ¿Qué se puede comentar acerca de los resultados obtenidos?
- Obtener la covarianza y comentar su valor.
- ¿Sería adecuado estimar linealmente los valores de Y en función de los de X? ¿Son X e Y independientes?

6.-Para una compañía aérea se han obtenido datos relativos a 24 trayectos realizados por el avión DC-9, obteniéndose los siguientes datos para las variables Y= consumo total de combustible, en miles de litros, correspondiente a un vuelo de duración X (expresado en horas):

$$\sum x_i = 31,47 \quad \sum y_i = 219,719 \quad \sum x_i y_i = 349,486 \quad \sum x_i^2 = 51,075 \quad \sum y_i^2 = 2396,504$$

- Ajustar un modelo lineal que permita la estimación de Y conociendo X.
- Obtener el coeficiente de correlación lineal.
- ¿Qué consumo total estimaría para un programa de vuelos compuesto de 100 vuelos de media hora, 200 de una hora y 100 de dos horas? ¿Es fiable esta estimación?

7.-Las notas en Física (X) y Matemáticas (Y) obtenidos por 10 alumnos elegidos al azar han sido las siguientes:

X	9	7	3	6	7	5	10	8	2	5
Y	8	5	4	2	9	6	10	9	1	5

- Obtener las rectas de regresión de Y/X y de X/Y.
- Obtener el coeficiente de correlación lineal. Interpretación.
- Representar gráficamente las rectas de regresión. Interpretación del gráfico.
- Para un alumno que haya obtenido un 4 en Física, ¿qué nota le estimaría para Matemáticas? Comentar la fiabilidad de la estimación.
- Para un alumno que haya obtenido un 7 en Matemáticas, ¿qué nota le estimaría para Física? Comentar la fiabilidad de la estimación.

8.-En la estimación de los parámetros de un modelo de regresión lineal simple se han obtenido los siguientes valores:

Media marginal de X : 5

Media marginal de Y : 8

Varianza marginal de Y : 20

Covarianza : 15

Coeficiente de determinación: 0,9

Determinar:

- La varianza marginal de X.
- Las rectas de regresión de Y/X y de X/Y.
- El coeficiente de correlación lineal.

9.-En un mercado de una ciudad se han observado durante 6 días consecutivos las cantidades de naranjas vendidas C (en cientos de Kg.) y el precio P correspondiente por Kg.(en euros), obteniéndose los siguientes datos:

$$\sum c_i = 21 \quad \sum p_i = 840 \quad \sum c_i^2 = 91 \quad \sum p_i^2 = 120.200 \quad \sum c_i p_i = 2.730$$

- Obtener la recta de regresión que permita estimar la cantidad de naranjas vendidas.
- Hallar el coeficiente de correlación lineal entre C y P. Interpretación.
- Para un precio por Kg. de 1 €, ¿qué cantidad de naranjas se venderán? ¿Es una buena estimación?

10.-Una factoría de refrescos ha tomado al azar 10 semanas del año, observando la temperatura media correspondiente a cada una de ellas, y la cantidad de refrescos pedidos durante cada semana, obteniéndose los siguientes resultados:

Temperatura Media (°C)	10	28	12	31	30	19	24	5	9	15
Cantidades de refrescos	21	65	19	72	75	39	67	11	12	24

¿Puede la factoría planificar la cantidad de producción en función de la temperatura esperada para cada semana? ¿De qué forma? ¿Es adecuada esta planificación?

11.-Sea la distribución unidimensional

X_i	3	5	8	9
n_i	5	1	2	1

que es una marginal de una distribución bidimensional (X,Y), de la que se conoce:

$$\sum n_i y_i^2 = 3.240 \quad Y/X \equiv Y = 5X - 20$$

Determinar la recta de regresión de X/Y, y comentar la bondad del ajuste.

12.-La siguiente tabla muestra la edad X y la presión sanguínea Y de 10 mujeres:

Nombre	A	B	C	D	E	F	G	H	I	J
Edad	56	42	72	36	63	47	55	47	38	42
Presión Sanguínea	148	126	159	118	149	130	151	142	114	141

- Obtener la recta de regresión de Y/X. Representarla gráficamente.
- Obtener el coeficiente de correlación lineal.
- Dar una predicción lineal para la presión sanguínea de una mujer de 51 años. ¿Es fiable esta estimación?

13.-Dos variables tienen las siguientes rectas de regresión:

$$8.X + 2.Y = 1 \quad 16.X + 9.Y = 1$$

- Obtener las rectas de regresión de Y/X y de X/Y.
- Obtener el coeficiente de correlación lineal.
- Obtener las medias marginales.

14.-Cinco niños de 2, 4, 6, 7 y 8 años pesan respectivamente 15, 19, 25, 33 y 34 Kg, Hallar la recta de regresión que permita estimar el peso de un niño de 10 años y de otro de 5 años. ¿Qué fiabilidad tendrían las estimaciones anteriores?

15.-En una distribución bidimensional (X, Y) las rectas de regresión son

$$X/Y: X + Y = 1 \quad Y/X: X + 2Y = -1$$

- Para $x = 2$, ¿se puede esperar que y valga -1 ?

16.-Los siguientes datos muestran el crecimiento de la población de truchas en 7 años de funcionamiento de una piscifactoría:

X Años	1	2	3	5	6	7
Y Truchas (miles)	3	7	15	100	250	650

Para predecir el número de truchas ajustamos una recta. Efectuar predicciones para los 4 y los 10 años, y comentar la validez de las estimaciones.

17.-Dada la tabla

X	-4	-2	0	2	4
Y	0	3,4641	4	3,4641	0

¿Son X e Y variables incorreladas?

¿Son X e Y variables independientes?