

Estadística aplicada

Práctica 6: DIAGRAMAS BIVARIANTES Y MEDIDAS DE RELACIÓN

Para realizar el estudio de una distribución bidimensional, además de calcular de calcular los principales parámetros de cada variable (media, desviación típica y varianza), calcularemos la covarianza que nos permitirá averiguar el coeficiente de Pearson. También, en la siguiente práctica hallaremos la ecuación de la recta de regresión y realizaremos estimaciones para una de las variables conociendo el valor de la otra.

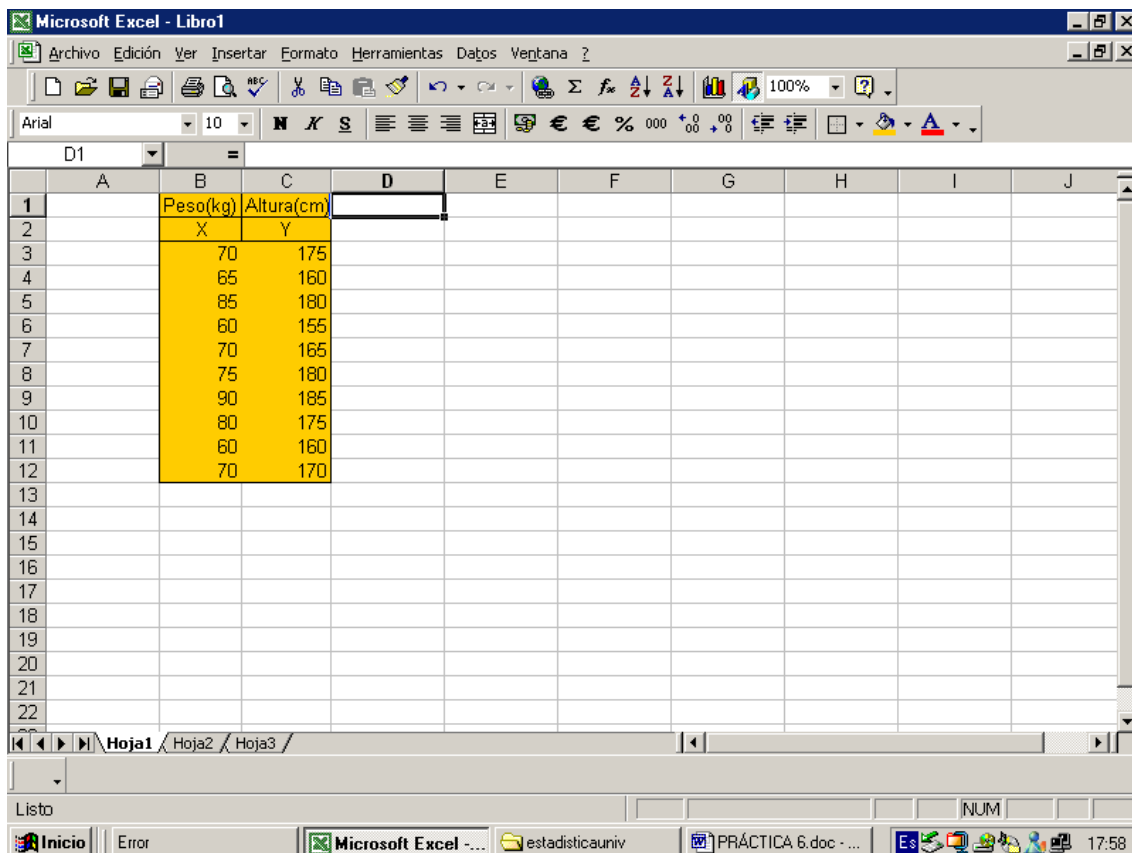
Ejemplo 1: Dada la siguiente distribución:

Peso(Kg.)	70	65	85	60	70	75	90	80	60	70
Altura(cm.)	175	160	180	155	165	180	185	175	160	170

- Representa la nube de puntos.
- Calcula las desviaciones típicas marginales, la covarianza y el coeficiente de correlación.

DIAGRAMAS BIVARIANTES

- Abre **Microsoft Excel** y en libro nuevo en Hoja 1 copia los datos iniciales en el rango B3:C12 :



- Selecciona el **Asistente de gráficos**. Luego en tipo de gráfico **Dispersión (XY)**.

3. En Nombre: **Nube de puntos**. En el eje X: **Peso** y en eje Y: **Altura**.

MEDIDAS DE RELACIÓN

Ahora para cada una de las variables X e Y calcularemos la media, la desviación típica y la varianza.

1. Copiamos la columna con los valores de la variable X en la columna E (Seleccionar y luego Pegar). Añadimos dos columnas más. La columna F para las frecuencias absolutas (iguales a 1 en este caso para todos los valores de X).
2. En la G calcularemos los cuadrados de los valores de X. Para ello en la celda G3 introduciremos la fórmula $=F3^2$ y utilizando el Control de relleno la copiaremos hasta G12.

En la celda E13 utilizando **Autosuma** calcularemos la suma de los valores de X.

En la celda F13 utilizando **Autosuma** calcularemos la frecuencia total de X.

En la celda G13 utilizando **Autosuma** calcularemos la suma de los cuadrados de valores de X.

3. Copiamos la columna con los valores de la variable Y en la columna I (Seleccionar y luego Pegar). Añadimos dos columnas más. La columna J para las frecuencias absolutas (iguales a 1 en este caso para todos los valores de Y). En la K calcularemos los cuadrados de los valores de Y. Para ello en la celda K3 introduciremos la fórmula $=I3^2$ y utilizando el Control de relleno la copiaremos hasta I12.

En la celda I13 utilizando **Autosuma** calcularemos la suma de los valores de Y.

En la celda J13 utilizando **Autosuma** calcularemos la frecuencia total de Y.

En la celda K13 utilizando **Autosuma** calcularemos la suma de los cuadrados de valores de Y.

4. Media, varianza y desviación típica de X.

En la celda C16 introduciremos la fórmula $=E13/F13$ y calcularemos la **media**.

En la celda C17 introduciremos la fórmula $=1/F13*G13-C16^2$ y calcularemos la **varianza**.

En la celda C18 introduciremos la fórmula $=RAIZ(C17)$ y calcularemos la **desviación típica**.

5. Media, varianza y desviación típica de Y.

En la celda C20 introduciremos la fórmula $=I13/J13$ y calcularemos la **media**.

En la celda C21 introduciremos la fórmula $=1/J13*K13-C20^2$ y calcularemos la **varianza**.

En la celda C22 introduciremos la fórmula $=RAIZ(C21)$ y calcularemos la **desviación típica**.

6. Covarianza y coeficiente de correlación.

En la celda O3 introduciremos la fórmula $=M3*N3$ y utilizando el Control de relleno la copiaremos hasta O12.

En la celda O13 utilizando **Autosuma** calcularemos la suma de los productos X.Y.

En la celda C24 introduciremos la fórmula $=1/10*O13-C16*C20$ y calcularemos la covarianza.

En la celda C25 introduciremos la fórmula $=C24/(C18*C22)$ y calcularemos el **coeficiente de correlación**.

Funciones de Excel

Ahora calcularemos la **covarianza** y el **coeficiente de correlación** utilizando las funciones de Excel:

En la celda J24 introduciremos =COVAR(B3:B12;C3:C12).

En la celda J24 introduciremos =COEF.DE.CORREL(B3:B12;C3:C12).

Vemos que se obtienen los mismos resultados.

The screenshot shows an Excel spreadsheet with the following data and calculations:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1		Peso(kg)	Altura(cm)		Peso(kg)	n_i	x_i^2		Altura(cm)	n_j	y_j^2		Peso(kg)	Altura(cm)	$x_i y_j$			
2		X	Y		X				Y				X	Y				
3		70	175		70	1	4900		175	1	30625		70	175	12250			
4		65	160		65	1	4225		160	1	25600		65	160	10400			
5		85	180		85	1	7225		180	1	32400		85	180	15300			
6		60	155		60	1	3600		155	1	24025		60	155	9300			
7		70	165		70	1	4900		165	1	27225		70	165	11550			
8		75	180		75	1	5625		180	1	32400		75	180	13500			
9		90	185		90	1	8100		185	1	34225		90	185	16650			
10		80	175		80	1	6400		175	1	30625		80	175	14000			
11		60	160		60	1	3600		160	1	25600		60	160	9600			
12		70	170		70	1	4900		170	1	28900		70	170	11900			
13					725	10	53475		1705	10	291625				124450			
14																		
15																		
16																		
17		Variable X	Media															
18			Varianza															
19			Desv.típica															
20			72,5															
21		Variable X	Media															
22			Varianza															
23			Desv.típica															
24			91,25															
25			9,552487															
26																		
27																		
28																		
29																		
30																		
31																		
32																		
33																		
34																		
35																		
36																		
37																		
38																		
39																		
40																		
41																		
42																		
43																		
44																		
45																		
46																		
47																		
48																		
49																		
50																		

Summary of statistical results from the spreadsheet:

Variable	Media	Varianza	Desv.típica
Variable X	72,5	91,25	9,552487
Variable Y	170,5	92,25	9,604696

Covariance and Correlation Coefficient:

Metric	Value
Covarianza	83,75
Coeficiente de correlación	0,91282