

Hoja de Cálculo. Práctica

Coeficiente de correlación y recta de regresión

Introducción.

En este caso se va a crear un libro, para realizar el estudio de una distribución bidimensional; además de calcular los principales parámetros de cada variable (media, desviación típica y varianza) y de ambas variables conjuntamente (covarianza), permitirá averiguar el coeficiente de correlación lineal de Pearson y la ecuación de la recta de regresión.

Descripción.

La utilización del libro consistirá en introducir los datos en un rango de entrada que, estará sombreado de color azul claro y automáticamente se obtendrán todos los resultados mediante las funciones introducidas en las celdas correspondientes.

	A	B	C	D	E	F
1	Cálculo de la Correlación y Regresión					
2						
3		x_i	y_i	x^2	y^2	$x_i y_i$
4		30	450	900	202500	13500
5		43	500	1849	250000	21500
6		45	458	2025	209764	20610
7		56	700	3136	490000	39200
8		76	988	5776	976144	75088
9		67	498	4489	248004	33366
10		45	487	2025	237169	21915
11		46	576	2116	331776	26496
12		67	876	4489	767376	58692
13		34	434	1156	188356	14756
14	Sumas:	509	5967	27961	3901089	325123

En primer lugar, se introducirán los datos (x_i, y_i) , y se harán los cálculos previos para la tabla $(x^2, y^2, x_i y_i)$, así como sus sumatorios.

Las principales formulas introducidas, son las correspondientes a la fila 14, con todas las autosumas, así como las celdas D4, E4 y F4 como puede verse en la siguiente figura.

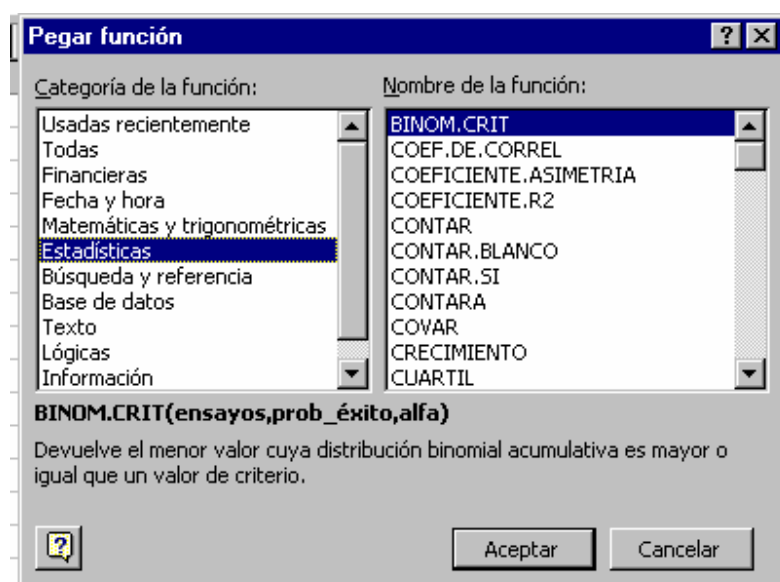
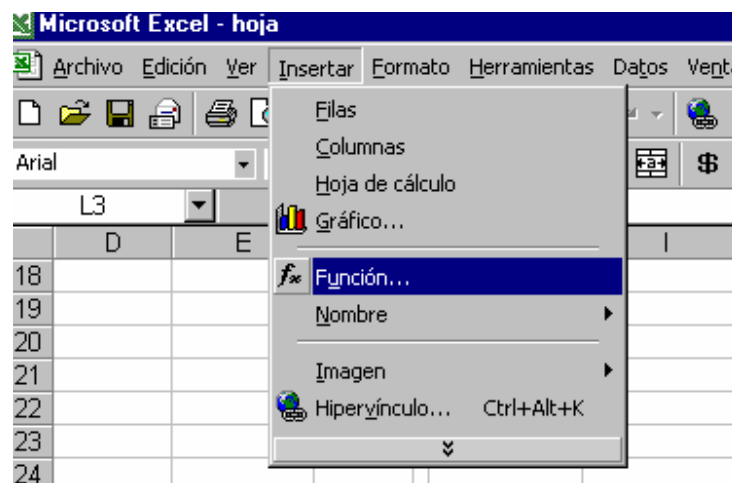
y_i	x^2	y^2	$x_i y_i$
0	=B4^2	=C4^2	=B4*C4
0	=B5^2	=C5^2	=B5*C5
8	=B6^2	=C6^2	=B6*C6
0	=B7^2	=C7^2	=B7*C7
8	=B8^2	=C8^2	=B8*C8
8	=B9^2	=C9^2	=B9*C9
7	=B10^2	=C10^2	=B10*C10
6	=B11^2	=C11^2	=B11*C11
6	=B12^2	=C12^2	=B12*C12
4	=B13^2	=C13^2	=B13*C13
SUMA(C4:C13)=SUMA(D4:D13) =SUMA(E4:E13) =SUMA(F4:F13)			

Una vez obtenida la tabla, a partir de la columna G, colocaremos todos los parámetros que pretendemos calcular, tanto unidimensionales como bidimensionales.

	Variable x	Variable y
Media	50.9	613
Desv. Típ.	14.32794472	184.5481238
Varianza	205.29	34058.01
Covarianza	2140.27	
coef.corr =	0.809422052	
Pendiente=	10.42559306	
Ord. Origen=	66.03731307	
Ec. recta=	y = 10.4255930634712 x + (66.0373130693166)	

En la hoja de cálculo encontramos también funciones específicas para determinados parámetros (media, desviación típica, varianza, covarianza, pendiente, ordenada en el

origen, etc.), únicamente tenemos que desplegar el menú INSERTAR y escoger FUNCION. Aparecerán distintas opciones para insertar funciones, y escogeremos la que deseemos utilizar en cada momento, siguiendo las distintas ventanas que irán apareciendo sucesivamente para insertar los datos correctamente. Previamente, habremos situado el cursor en la celda donde deseemos insertar la función.



Debemos insertar todas las funciones deseadas y repetir el proceso por lo tanto para cada una de ellas, el resultado final debe ser:

	Variable x	Variable y
Media	=PROMEDIO(B4:B13)	=PROMEDIO(C5:C13)
Desv. Típ.	=DESVESTP(B4:B13)	=DESVESTP(C4:C13)
Varianza	=VARP(B4:B13)	=VARP(C4:C13)
Covarianza	=COVAR(B4:B13,C4:C13)	
coef.corr =	=COEF.DE.CORREL(B4:B13,C4:C13)	
Pendiente=	=ESTIMACION.LINEAL(C4:C13,B4:B13)	
Ord. Origen=	=INTERSECCION.EJE(C4:C13,B4:B13)	
Ec. recta=	=CONCATENAR(X1,X4,Y2,Y4,X6,X5,X3)	

Mediante la opción concatenar, conseguimos unir el contenido de varias celdas, consideradas como texto. Para ello se escogen unas celdas auxiliares separadas de los cálculos principales, donde insertaremos los caracteres que queramos que se unan para dar como resultado la expresión final de la recta de regresión.

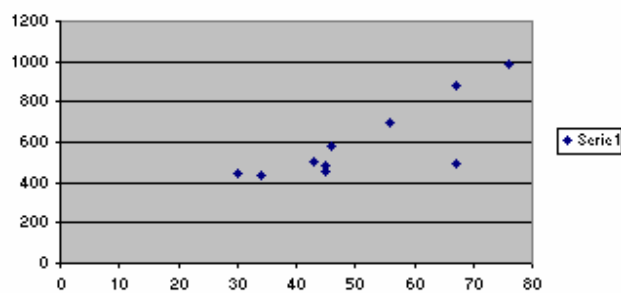
	X	Y
y =		
		x
)		
=I11		+
=I12		
(

En las celdas X4 y X5, se ponen la pendiente y la ordenada en el origen, calculadas en las celdas I11 e I12.

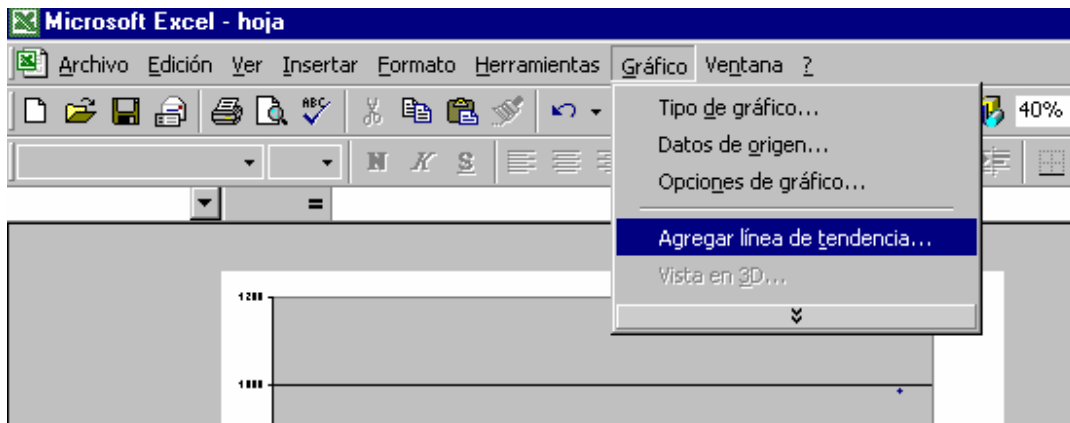
Una vez que se han presentado y calculado todos los parámetros en la hoja de cálculo se le da nombre y pasamos a representar la gráfica correspondiente de la nube de puntos y la recta de regresión calculada, para ello, seleccionaremos los datos con el rango B4:C13 y a continuación escogeremos la opción asistente para gráficos de la barra de herramientas, y el diagrama XY



quedará en este momento el gráfico como muestra la siguiente figura:



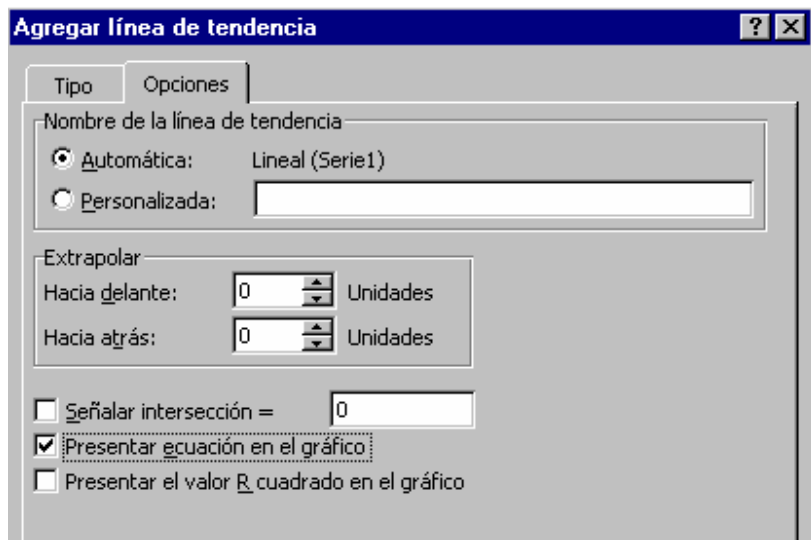
para que aparezca conjuntamente la recta de regresión y la ecuación de dicha recta, debemos seleccionar dentro del menú gráfico, AGREGAR LINEAS DE TENDENCIA.



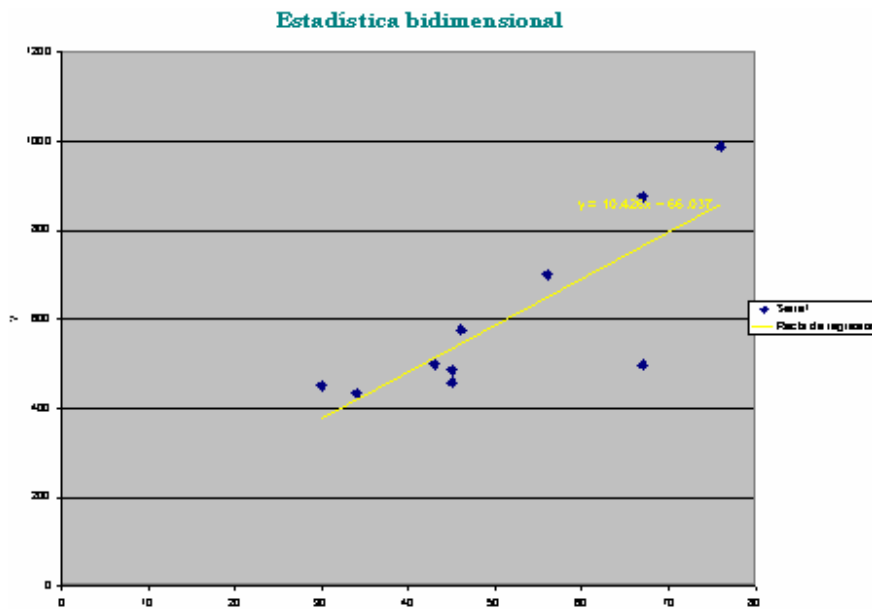
Aparecerá una nueva ventana donde podremos escoger el tipo de línea de tendencia que deseamos, en este caso LINEAL:



y dentro de la ficha de opciones, seleccionaremos, PRESENTAR ECUACIÓN EN EL GRAFICO.



Una vez creado el gráfico, podemos cambiar fácilmente las opciones de color de la recta y de todos los textos, fondos, títulos, colores, etc. de la gráfica final, pudiendo presentar un aspecto finalmente similar a este:



Ejercicios

1º) Haz el ejercicio propuesto en la descripción.

2º) Modifica el aspecto de la hoja anterior para que puedas introducir 15 pares de datos, en lugar de los 10 que presentaba. (Utiliza la opción, INSERTAR FILA del menú INSERTAR), y modifica mediante CORTAR y PEGAR, el aspecto de la hoja a partir de la columna H.

3º) Una vez que tengas hecho el ejercicio 2, aprovéchalo para calcular y dibujar la recta de regresión de los siguientes datos:

X	Y
34	356
56	578
56	589
67	654
89	897
34	356
54	532
13	147
45	468
67	684
23	222
78	801
52	498
98	1004
46	502

Guía para los ejercicios

1º) Sigue los pasos de la descripción.

2º) Mediante insertar fila, repetir el proceso hasta obtener las filas requeridas. Situarse previamente en una celda donde el aspecto de la hoja de cálculo no se vea muy afectado. Las formulas introducidas, conservarán todos los aspectos relativos de las nuevas celdas introducidas, por lo que no habrá que modificarlas.

	A	B	C	D	E	F	G	H		
1		Cálculo de la Correlación y Regresión								
2										
3		x_i	y_i	x^2	y^2	$x_i y_i$			Varia	
4		30	450	900	202500	13500				
5		43	500	1849	250000	21500	Media			
6		45	458	2025	209764	20610	Desv. Típ.	14.32		
7		56	700	3136	490000	39200	Varianza			
8										
9										
10										
11										
12										
13		76	988	5776	976144	75088				
14		67	498	4489	248004	33366	Covarianza			
15		45	487	2025	237169	21915	coef.corr =	0.809		

Una vez insertadas las filas, vemos que a partir de la columna H, ha quedado modificada la hoja de calculo. Debemos volver a dejar su aspecto inicial:

	A	B	C	D	E	F	G	H	I	
1		Cálculo de la Correlación y Regresión								
2										
3		x_i	y_i	x^2	y^2	$x_i y_i$		Variable x		
4		30	450	900	202500	13500				
5		43	500	1849	250000	21500	Media		50.9	
6		45	458	2025	209764	20610	Desv. Típ.		14.32794472	
7		56	700	3136	490000	39200	Varianza		205.29	
8										
9							Covarianza		2140.27	
10							coef. corr =		0.809422052	
11							Pendiente=		#¡VALOR!	
12							Ord. Origen=		66.03731307	
13		76	988	5776	976144	75088				
14		67	498	4489	248004	33366	Ec. recta=		#¡VALOR!	
15		45	487	2025	237169	21915				

	Variable x	
Media	=PROMEDIO(B4:B18)	=PROMEDIO(B4:B18)
Desv. Típ.	=DESVESTP(B4:B18)	=DESVESTP(B4:B18)
Varianza	=VARP(B4:B18)	=VARP(B4:B18)
Covarianza	=COVAR(B4:B18,C4:C18)	=COVAR(B4:B18,C4:C18)
coef. corr =	=COEF.DE.CORREL(B4:B18,C4:C18)	=COEF.DE.CORREL(B4:B18,C4:C18)
Pendiente=	=ESTIMACION.LINEAL(C4:C18,B4:B18)	=ESTIMACION.LINEAL(C4:C18,B4:B18)
Ord. Origen=	=INTERSECCION.EJE(C4:C18,B4:B18)	=INTERSECCION.EJE(C4:C18,B4:B18)
Ec. recta=	=CONCATENAR(X1,X4,Y2,Y4,X6,X8)	

Los rangos han cambiado automáticamente

Con lo cual esta nueva hoja ya esta lista para que le introduzcan los datos correspondientes.

3º) Introduce los datos y represéntalos, tal como se explicó en la descripción.