

Práctica 5.2

Construcción de un modelo para el estudio de la Regresión

Para estudiar la regresión entre dos variables estadísticas deberás construir un modelo adecuado con Hoja de Cálculo. Puede ser similar al siguiente:

	A	B	C	D	E	F	G
1							
2							
3							
4							
5							
6							
7	Datos x	Datos y	Pronóstico	Error			
8	0	2	1,88	-0,12			
9	0	1	1,88	0,88			
10	1	3	3,39	0,39			
11	1	4	3,39	-0,61			
12	1	3	3,39	0,39			
13	1	2	3,39	1,39			
14	1	2	3,39	1,39			
15	2	4	4,91	0,91			
16	2	5	4,91	-0,09			
17	2	7	4,91	-2,09			
18	2	8	4,91	-3,09			
19	2	6	4,91	-1,09			
20	3	5	6,42	1,42			
21	4	8	7,93	-0,07			
22	4	10	7,93	-2,07			

(Instrucciones en la parte inferior del modelo)

Estudio de la Correlación y la Regresión	
Coeficiente de correlación $R=0,82$	
Pendiente $B=1,51$	
Coordenada en el origen $A=1,88$	
Error típico $E=1,55$	

Escribe todos los títulos, rótulos y colores a tu gusto

Reserva dos columnas con unas 15 ó 20 celdas cada una, para albergar los datos X e Y respectivamente. Rellénalas, tal como hiciste al principio de la sesión, con los datos de la tabla de horas de estudio.

Selecciona todos los datos de la X con el ratón y pide **Insertar > Nombre > Definir** y les das el nombre de **X**. Haz lo mismo para dar el nombre de **Y** a la segunda columna.

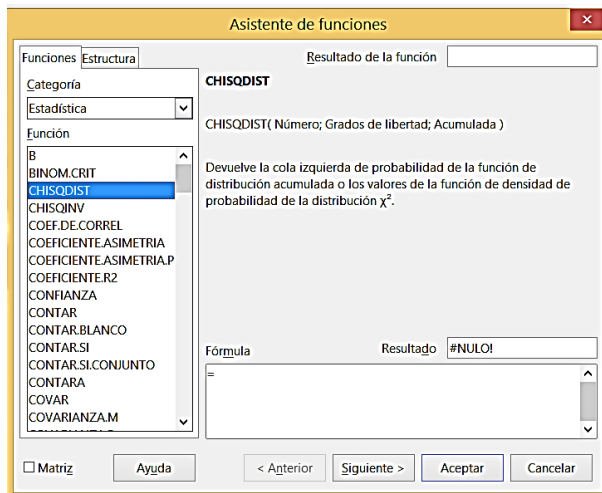
Coeficiente de correlación

Recuerda que es un número entre -1 y 1 que representa el grado de asociación entre X e Y. Si introducimos la recta de regresión, también se puede interpretar como el grado de ajuste de los datos a dicha recta. Para su cálculo debes usar la fórmula

=COEF.DE.CORREL(Y;X) Comprueba que te vuelve a resultar 0,82...

Esta fórmula y las siguientes las puedes insertar con el **asistente de funciones** eligiendo el tipo **funciones estadísticas**.





Sitúa el cálculo de R en la celda que desees, o siguiendo la estructura propuesta. Recuerda dejar dos columnas libres para pronósticos y errores.

Línea de regresión

Para hallar los coeficientes A y B de la recta de regresión puedes usar:

Para B, que es la pendiente: **=PENDIENTE(Y;X)**

Para A, que es la intersección con el eje Y: **=INTERSECCION.EJE(Y;X)**

Como ves, las fórmulas son muy directas y no necesitan cálculos complejos, que quedan a cargo del programa.

El valor de B es 1,513 y el de A 1,881, luego la ecuación de la recta de regresión será **Y' = 1,881+1,513X**, que es el mismo resultado que obtuvimos en la práctica 1.

Pronósticos

Los pronósticos se pueden calcular mediante la fórmula $Y'=A+BX$, pero LibreOffice dispone de una función especial, por lo que podemos rellenar la columna directamente.

La fórmula que debes usar es:

=PRONÓSTICO(celda de X que tiene enfrente;Y;X)

La celda de X escríbela como es usual en la Hoja: B7 ó C4, etc. Rellena así la primera celda de la columna que has reservado para pronósticos y copia luego hacia abajo. Comprueba que te resultan estos pronósticos:

Datos x	Datos y	Pronóstico	Error
0	2	1,88	-0,12
0	1	1,88	0,88
1	3	3,39	0,39
1	4	3,39	-0,61
1	3	3,39	0,39
1	2	3,39	1,39
1	2	3,39	1,39
2	4	4,91	0,91
2	5	4,91	-0,09
2	7	4,91	-2,09
2	8	4,91	-3,09
2	6	4,91	-1,09
3	5	6,42	1,42
4	8	7,93	-0,07
4	10	7,93	-2,07
5	7	9,45	2,45

Errores

Error típico: Es el error que se espera en los pronósticos. Lo tienes en la parte derecha del modelo. Su fórmula es:

=ERROR.TIPICO.XY(Y;X)

El valor que te da, 1,55, es el error que se espera cometer en los pronósticos. Su significado preciso es un poco más complicado.

Columna de errores: La consigues restando el pronóstico y el verdadero valor de Y. Comprueba los resultados.

Con esto tienes un modelo apropiado para estudiar la correlación y la regresión, pero no estaría completo sin un gráfico de dispersión XY.

Gráfico de la recta de regresión

Selecciona con el ratón las dos columnas: X, Y con sus rótulos

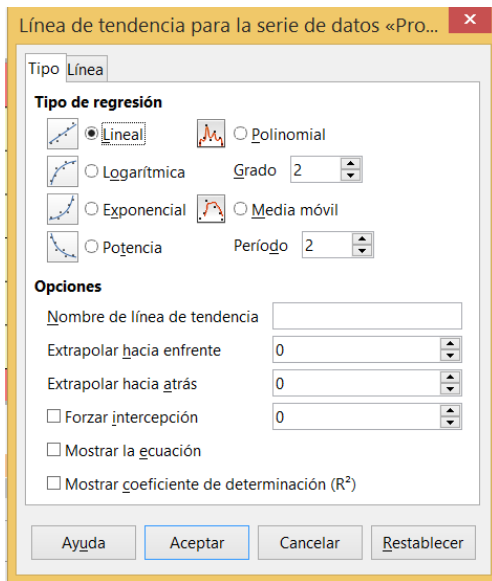
Construye un gráfico **XY Dispersión**. Ya sabes cómo



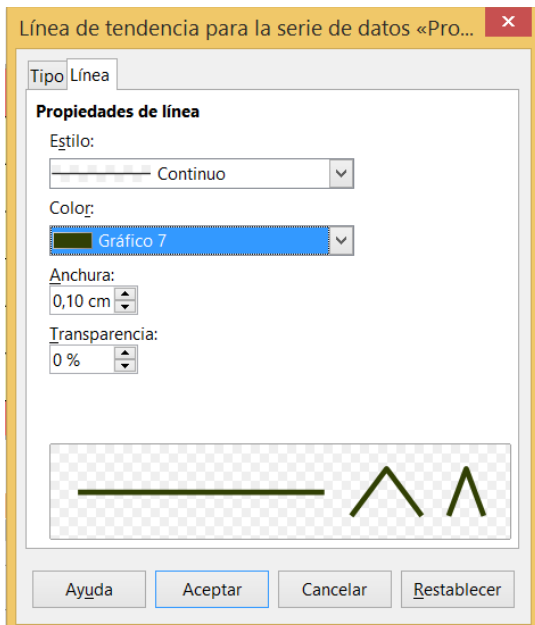
Para incluir la recta de regresión de manera automática puedes proceder como sigue:

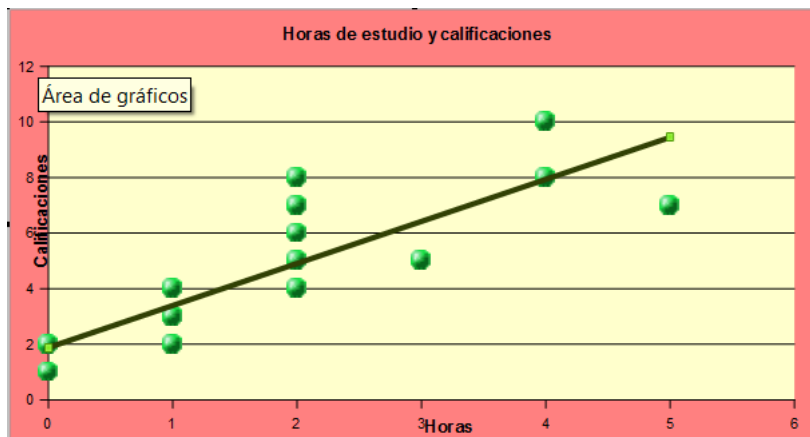
- Una vez construido el gráfico a tu gusto, haz doble clic sobre él, señala un punto de la nube con el ratón y vuelve a hacer clic hasta que veas seleccionados todos los puntos de esa nube.

- Pulsa entonces con el botón derecho y elige **Insertar línea de tendencia**. En ella activa el *añadir como curva de regresión la línea de regresión lineal*.



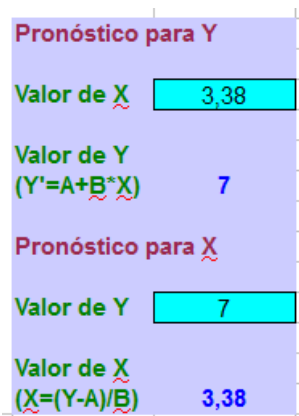
Acepta y verás dibujada la recta de regresión deseada. Si no la ves a tu gusto, vuelve a repetir el proceso, pero elige la pestaña línea para cambiar sus características.





Pronósticos para otros valores

La recta de regresión es útil para hacer pronósticos sobre valores que no figuran en la tabla. Vamos a continuación a resolver la cuestión previa de la sesión. ¿Cuánto tengo que estudiar para sacar notable? Según la tabla de pronósticos, entre 3 y 4 horas. Si quisiéramos concretar más (si es que eso es razonable en este caso) deberíamos usar la fórmula de la recta de regresión, ampliando el modelo creado con estas otras celdas:



La simple inspección de la figura te dará idea de cómo organizar los pronósticos. El valor de Y lo consigues con la fórmula de regresión

$$Y' = A + BX$$

y el de X, despejando: $X = (Y' - A) / B$

En la figura tienes la solución de la cuestión previa: **El notable supone estudiar 3,38 horas.**

Si tienes dificultades en algún detalle de la confección de este modelo, puedes consultar **bidim0.ods** en la carpeta de modelos de esta sesión.