Ajuste de una recta por mínimos cuadrados

- Los datos y su interpretación
- Los parámetros que mejor ajustan.
- Estimación de la incertidumbre de los parámetros.
- Coeficiente de correlación lineal.
- Presentación de los resultados. Ejemplo.

Los datos y su interpretación

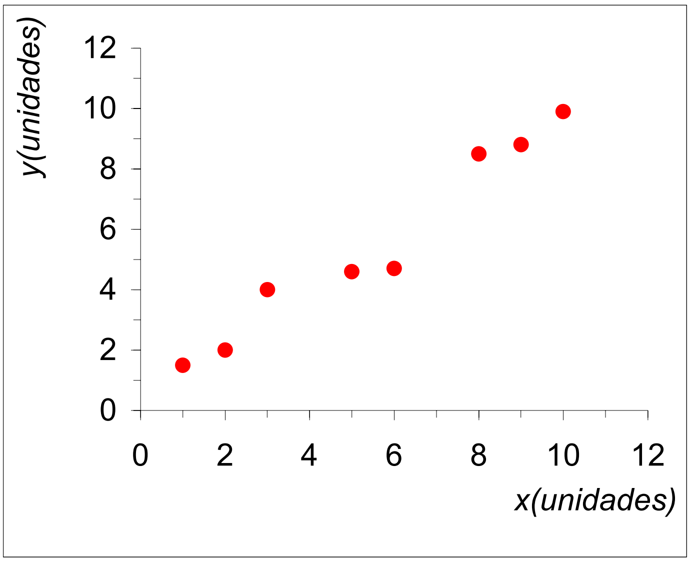
Razones teóricas: $y = mx + n$
N pares de medidas $(x_1, y_1); (x_2, y_2); \cdots; (x_N, y_N)$

☞ Antes de tomar las medidas:

- ✗ El intervalo elegido para la variable independiente, ¿abarca todo el rango de interés?
- ✗ ¿Están los puntos uniformemente distribuidos en este intervalo?

☞ Ordenación y representación gráfica de los datos

x_i	y_i
1	1.5
2	2.0
3	4.0
5	4.6
6	4.7
8	8.5
9	8.8
10	9.9

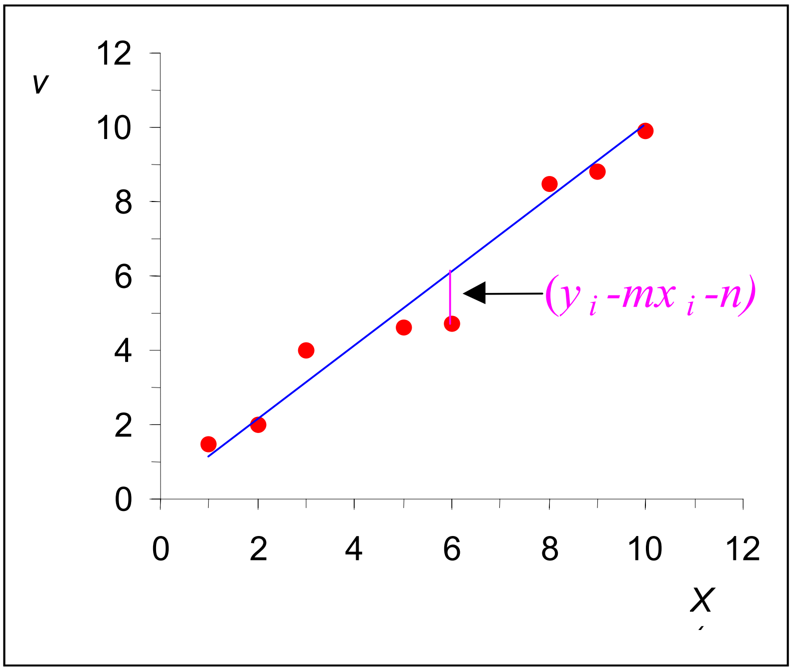


- ✗ ¿Se comportan los pares de medidas visualmente según una línea recta?
- ✗ ¿Hay algún punto que presente un comportamiento anómalo?

Los parámetros que mejor ajustan

☞ ¿Cuál es la recta que mejor se ajusta a las N medidas?

$$\chi^2(n,m) = \sum_{i=1}^{N} (y_i - mx_i - n)^2$$



$$m = \frac{NS_{xy} - S_x S_y}{NS_{xx} - S_x S_x}$$

$$n = \frac{S_{xx} S_y - S_x S_{xy}}{NS_{xx} - S_x S_x}$$

☞ ¿Qué valores de m y n hacen mínimo χ^2?

$$\frac{\partial \chi^2}{\partial m} = 0 \to 0 = \sum_{i=1}^{N} -2\left(y_i - mx_i - n\right)x_i = -2\sum_{i=1}^{N}\left(y_i x_i - mx_i^2 - nx_i\right)$$

$$\frac{\partial \chi^2}{\partial n} = 0 \to 0 = \sum_{i=1}^{N} -2\left(y_i - mx_i - n\right)$$

Definiendo

$$S_x = \sum_{i=1}^{N} x_i \quad S_y = \sum_{i=1}^{N} y_i \quad S_{xx} = \sum_{i=1}^{N} x_i^2 \quad S_{xy} = \sum_{i=1}^{N} x_i y_i$$

Estimación de la incertidumbre de los parámetros

☞ ¿Cuál es el mejor estimador de las incertidumbres de m y de n?

Suponemos que:

- Solo los valores y_i tienen error: δy_i
- Los errores en y son todos iguales: $\delta y_i = \delta y = \sigma_y$ y se estima a partir de la varianza de los datos:

$$\sigma_y^2 = \frac{1}{N-2}\sum_{i=1}^{N}\left(y_i - mx_i - n\right)^2 = \frac{\chi^2(n,m)}{N-2}$$

Aplicando propagación de errores:

$$\sigma_m^2 = \sum_{j=1}^{N}\left(\frac{\partial m}{\partial y_j}\sigma_y\right)^2 \quad ; \quad \sigma_n^2 = \sum_{j=1}^{N}\left(\frac{\partial n}{\partial y_j}\sigma_y\right)^2$$

y operando se obtiene:

$$\sigma_n^2 = \frac{S_{xx}}{NS_{xx} - S_x S_x}\frac{\chi^2(n,m)}{N-2}$$

$$\sigma_m^2 = \frac{N}{NS_{xx} - S_x S_x}\frac{\chi^2(n,m)}{N-2}$$

Coeficiente de correlación lineal

¿Cómo podemos saber cuán bueno es el comportamiento lineal de los N pares de datos medidos?

☞ Los errores en las medidas σ_{y_i} son conocidos:

- ¿La recta pasa por casi todos las barras de error de los puntos?
- Test de χ^2.

☞ Los errores en las medidas σ_{y_i} son desconocidos:

- A partir de la dispersión de los datos.
- Coeficiente de correlación lineal: r
- Mide el grado de correlación lineal entre x e y.
- $|r| \leq 1$
    - $|r| = 1$ Correlación total.
    - $r = 0$ No hay correlación.

$$r = \frac{NS_{xy} - S_x S_y}{\sqrt{NS_{xx} - S_x S_x}\sqrt{NS_{yy} - S_y S_y}} \qquad \text{siendo} \qquad S_{yy} = \sum_{i=1}^{N} y_i^2$$

Presentación de los resultados

☞ Ejemplo

Tabla de datos y cálculos

i	x_i	y_i	$x_i\,y_i$	x_i^2	y_i^2	$(n+mx_i-y_i)^2$
1	1	1.5	1.5	1.0	2.25	0.042
2	2	2.0	4.0	4.0	4.00	0.052
3	3	4.0	12.0	9.0	16.00	0.699
4	5	4.6	23.0	25.0	21.16	0.187
5	6	4.7	28.2	36.0	22.09	1.606
6	8	8.5	68.0	64.0	72.25	0.440
7	9	8.8	79.2	81.0	77.44	0.000
8	10	9.9	99.0	100.0	98.01	0.037

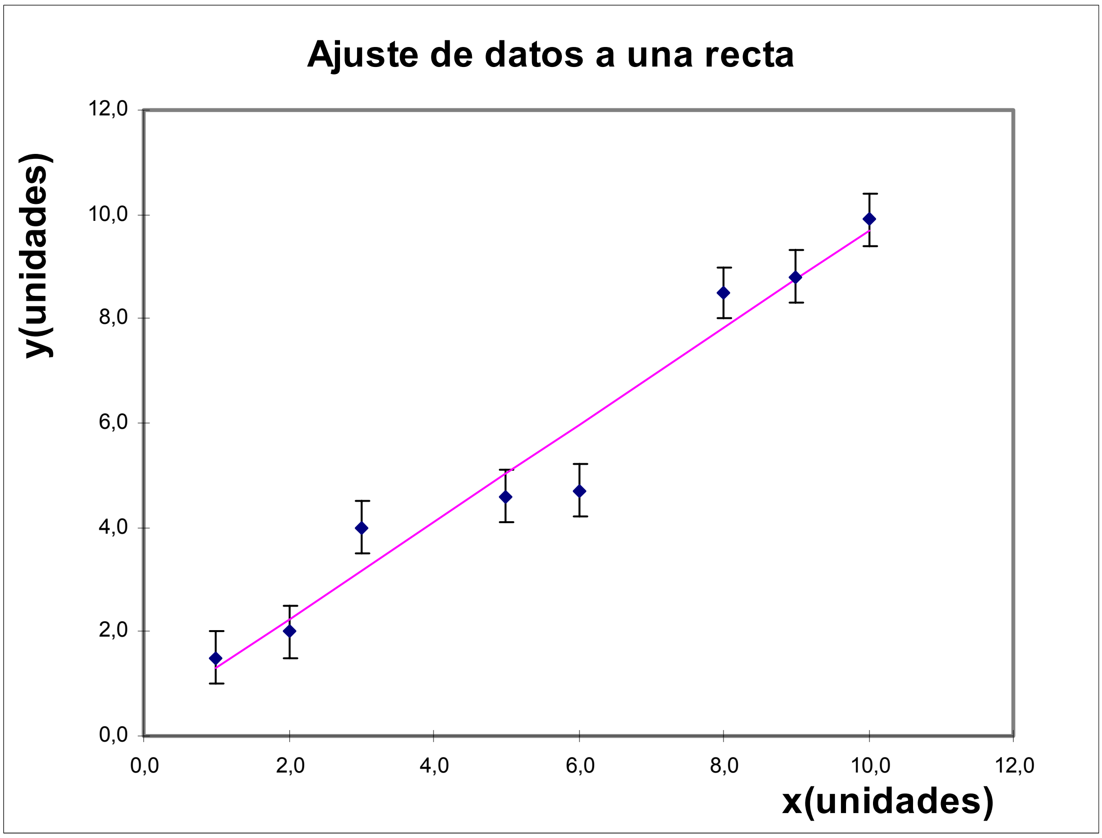
N=8	S_x=44	S_y=44	S_{xy}=314.9	S_{xx}=320	S_{yy}=313.2	χ^2=3.066

PARÁMETROS DEL AJUSTE :

$$m=\frac{NS_{xy}-S_xS_y}{NS_{xx}-S_xS_x}=0.935 \qquad \varepsilon(m)=\sqrt{\frac{N}{NS_{xx}-S_xS_x}\frac{\chi^2(n,m)}{N-2}}=0.081$$

$$n=\frac{S_{xx}S_y-S_xS_{xy}}{NS_{xx}-S_xS_x}=0.36 \qquad \varepsilon(n)=\sqrt{\frac{S_{xx}}{NS_{xx}-S_xS_x}\frac{\chi^2(n,m)}{N-2}}=0.512$$

$$r=\frac{NS_{xy}-S_xS_y}{\sqrt{NS_{xx}-S_xS_x}\sqrt{NS_{yy}-S_yS_y}}=0.978$$



$$\boxed{y = (0.94 \pm 0.08)x + (0.4 \pm 0.5)}$$