Física 1 (Química)
Laboratorio

Verano 2021

JTP Nicolás Torasso

Ay 1ra Magalí Xaubet

Ay 2da Adán Garros

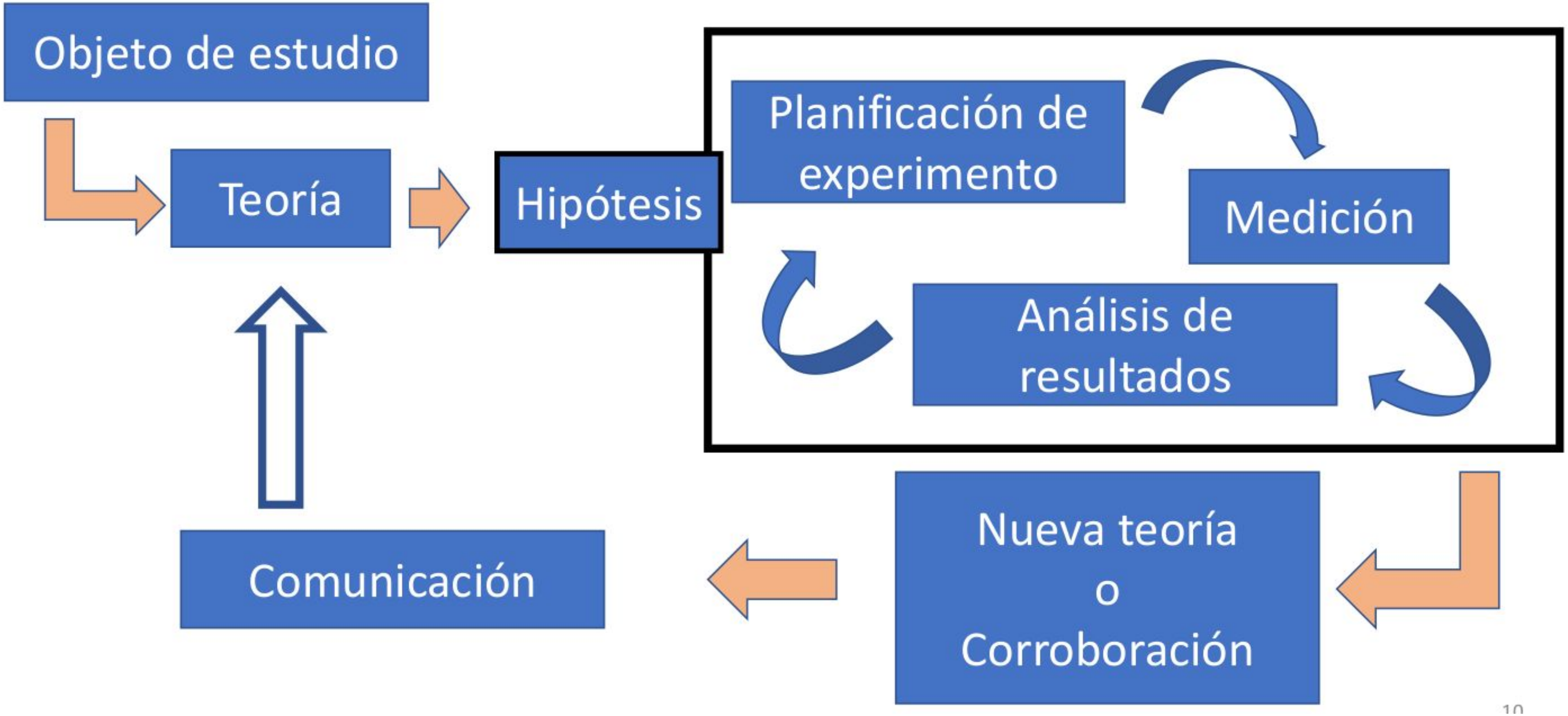
Lunes y Miércoles 14:30-18:30 hs

Prof. Gustavo Lozano

Hoy

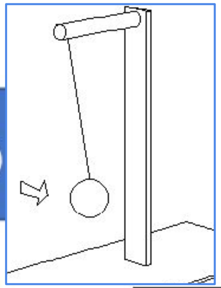
- Repaso breve.
- Testeo de un modelo mediante gráficos tipo “scatter”
- Testeo de modelo lineal o linealizado: Ajuste lineal - cuadrados mínimos (ponderados - no ponderados)
- Coeficientes r, R^2: estimadores de validez del modelo
- Cuadrados mínimos vs cálculo punto a punto
- Planificación del Experimento 4: determinación de g

El trabajo experimental: repaso



El trabajo experimental: repaso

Objeto de estudio



Teoría

Péndulo simple

Hipótesis

Masa puntual

Amortiguamiento despreciable

Planificación de experimento

Medición

Directa: **T**, L, m

Indirecta: T(L)

Estimación errores

Análisis de resultados

Estadística sobre T

Visualización: histograma

Nueva teoría o Corroboración

?

Comunicación

Informe

Review de pares

10

Análisis y testeo de un modelo

- Para testear un modelo podemos calcular los valores que predice y comparar con el valor que medimos en nuestro experimento.

- Otra alternativa es graficar pares de variables en ejes cartesianos y comparar con el tipo de dependencia (función matemática) que indica el modelo.

- Una visualización apropiada nos permite:
  - analizar los resultados del experimento y extraer información novedosa
  - simplificar la comunicación de los resultados a la comunidad

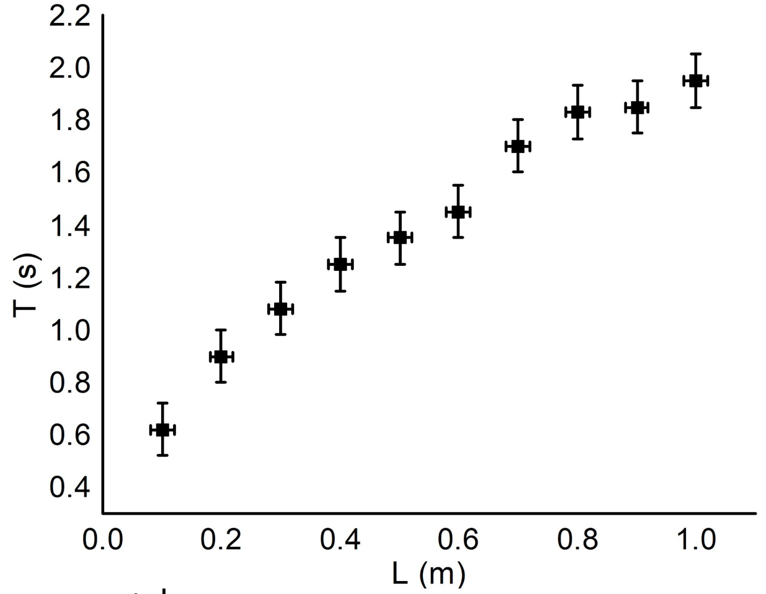
Análisis y testeo de un modelo

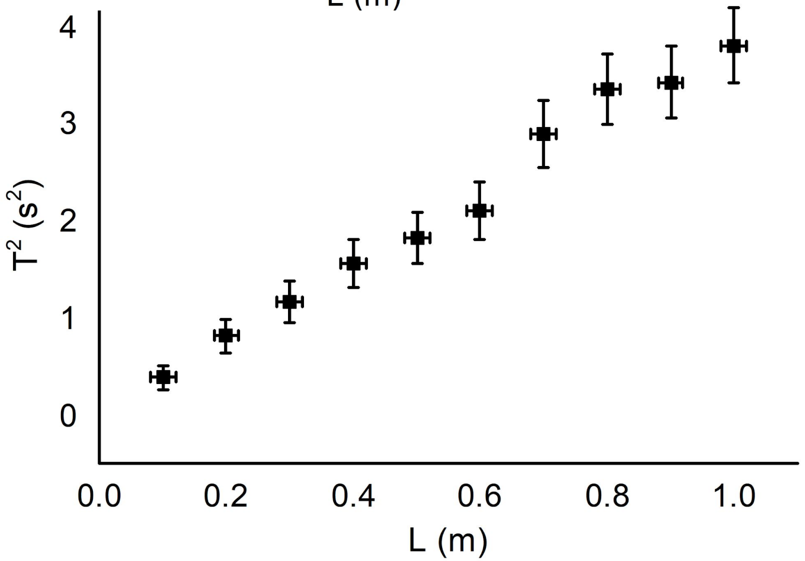
- A la hora de testear un modelo:
  - liberarnos de preconceptos. ej: el modelo es correcto. el modelo es incorrecto.
  - la naturaleza se manifestó frente nuestro: ser objetivos y reportar con fidelidad, sin descartar observaciones o alterar resultados

Testeo mediante gráficos tipo scatter

- Podemos graficar dos parámetros físicos medidos y visualizar su tipo de relación
- Podemos graficar los parámetros en forma **linealizada**. Para ello necesitamos conocer de antemano el modelo a testear. Ej:

$$T = 2\pi\sqrt{\frac{L}{g}}$$



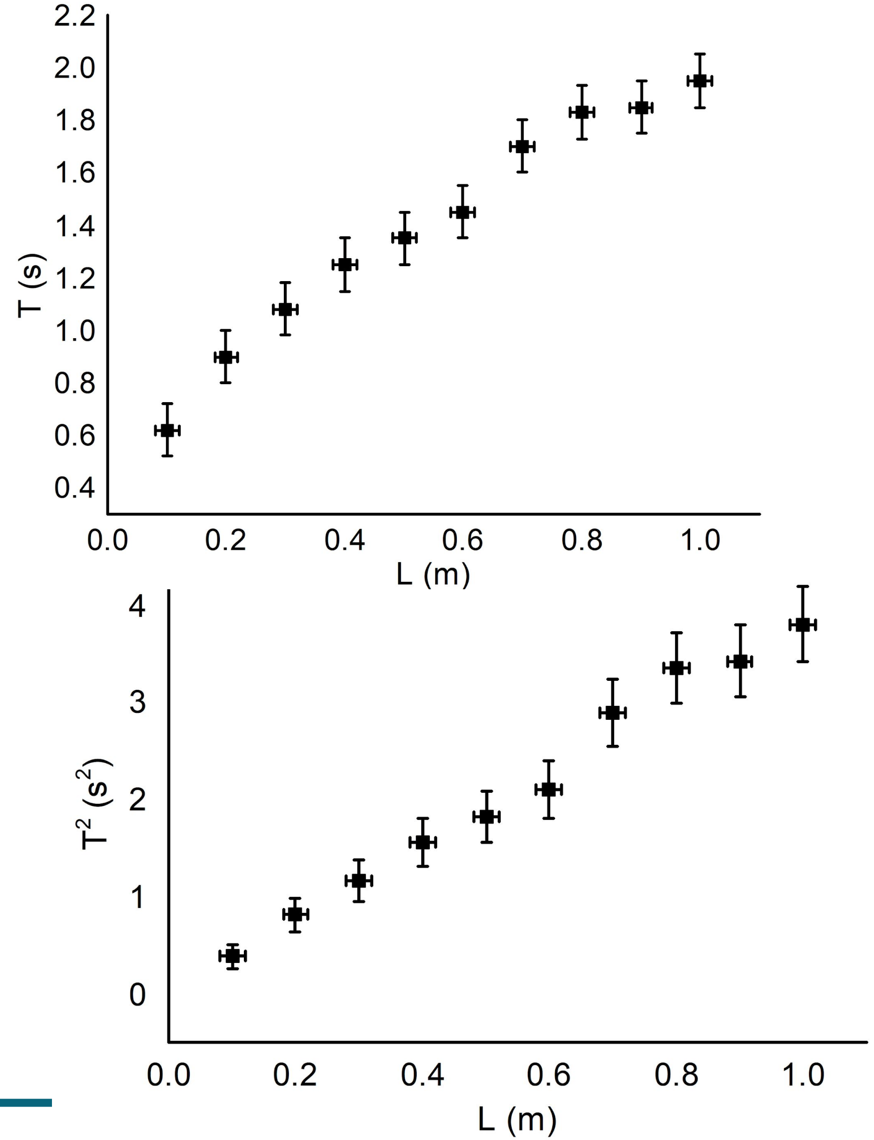


Testeo mediante gráficos tipo scatter

- Podemos graficar dos parámetros físicos medidos y visualizar su tipo de relación
- Podemos graficar los parámetros en forma **linealizada**. Para ello necesitamos conocer de antemano el modelo a testear. Ej:

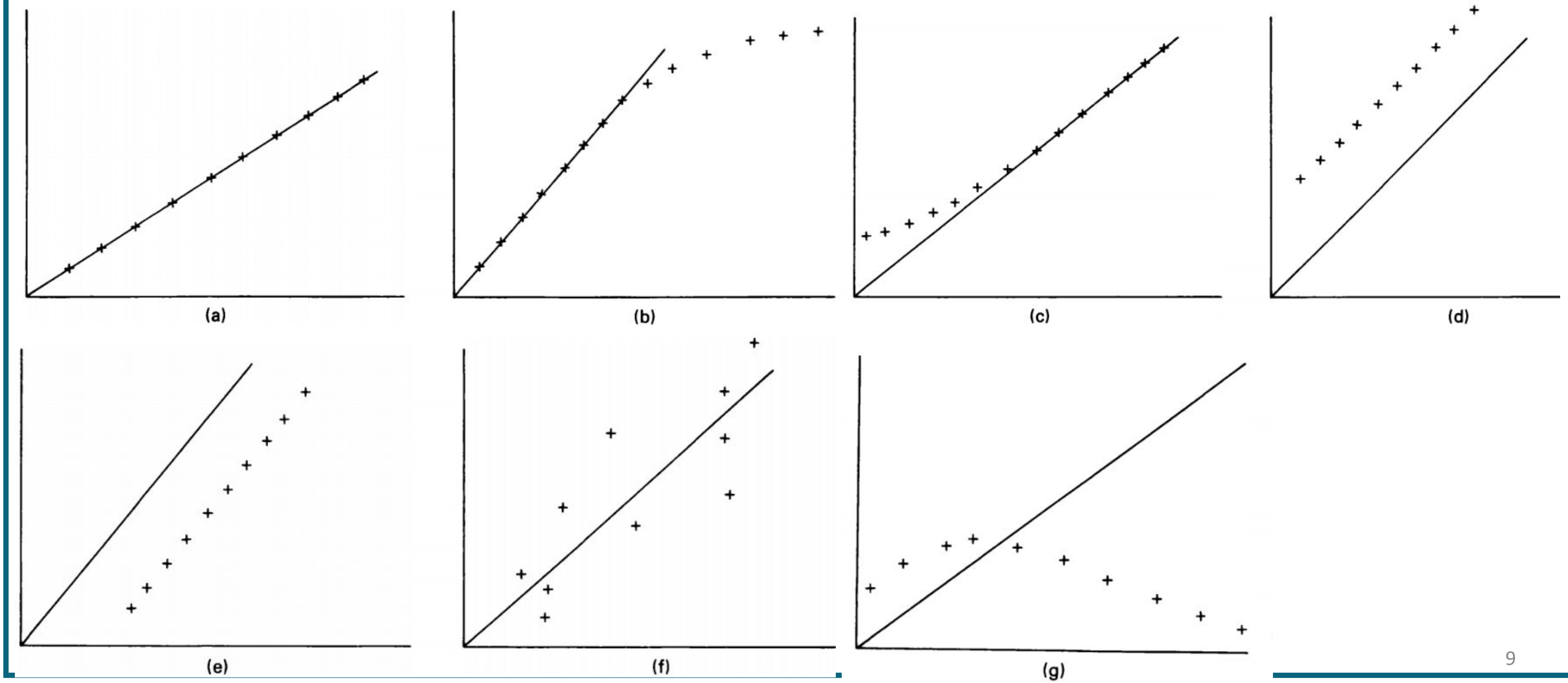
$$T = 2\pi\sqrt{\frac{L}{g}}$$

Opciones: T vs. √L
T^2 vs. L



Testeo de relación lineal

entre dos parámetros o dos parámetros linealizados. Puede pasar:



Regresión lineal - cuadrados mínimos

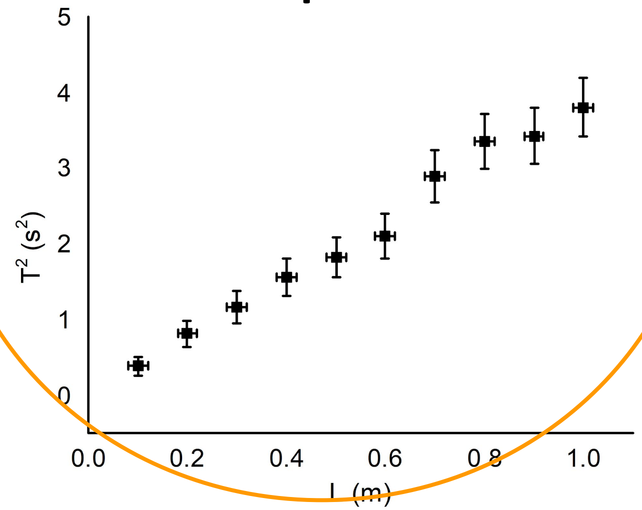
- Método matemático para obtener la recta que "mejor ajusta" la nube de puntos. f(x)= **a** x + **b**
- Obtenemos como resultado los parámetros **a** y **b** de la "recta que mejor ajusta"(y sus errores)
- Optativo: calcular estimadores de la validez del ajuste: r, R^2, χ^2
- Nota: el método de cuadrados mínimos no-lineal sirve para testear cualquier otro tipo de relaciones funcionales: cuadráticas, polinomiales, gaussianas, etc etc. $f(x) = ax^2 + bx + c$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Regresión lineal - cuadrados mínimos

- Método matemático para obtener la recta que “mejor ajusta” la nube de puntos. f(x)= **a** x + **b**

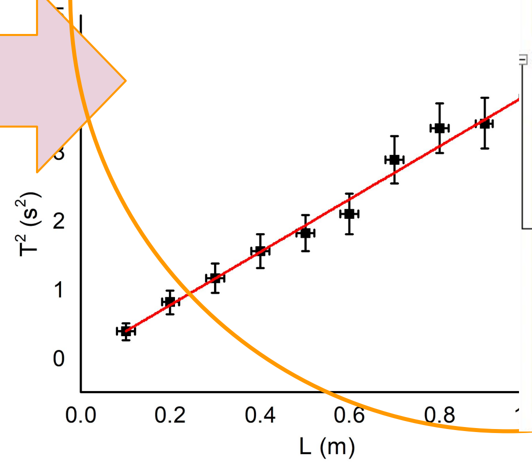
Input



Cuadrados mínimos

método matemático

Output



Parameters

		Value	Standard Error
T^2	Intercept	0.00197	0.05303
	Slope	3.85968	0.12218

Statistics

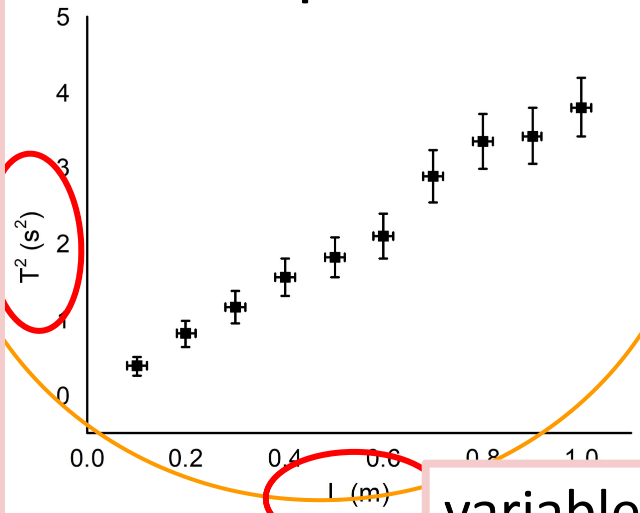
	T^2
Number of Points	10
Degrees of Freedom	8
Reduced Chi-Sqr	0.20082
Residual Sum of Squares	1.60659
Pearson's r	0.99602
R Value	0.99602
R-Square(COD)	0.99205
Adj. R-Square	0.99105

Regresión lineal - cuadrados mínimos

- Método matemático para obtener la recta que “mejor ajusta” la nube de puntos. f(x)= **a** x + **b**

Input

variable dependiente

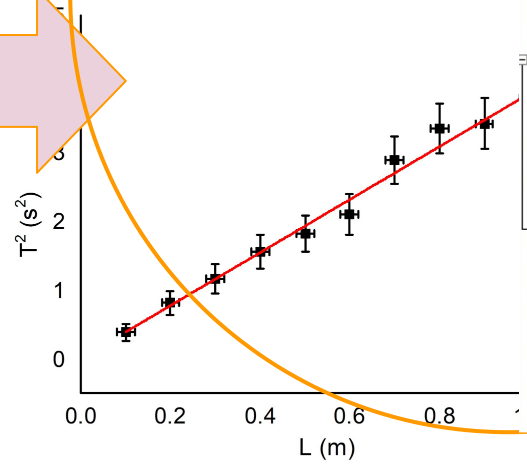


variable independiente

Cuadrados mínimos

método matemático

Output



Parameters

		Value	Standard Error
T^2	Intercept	0.00197	0.05303
	Slope	3.85968	0.12218

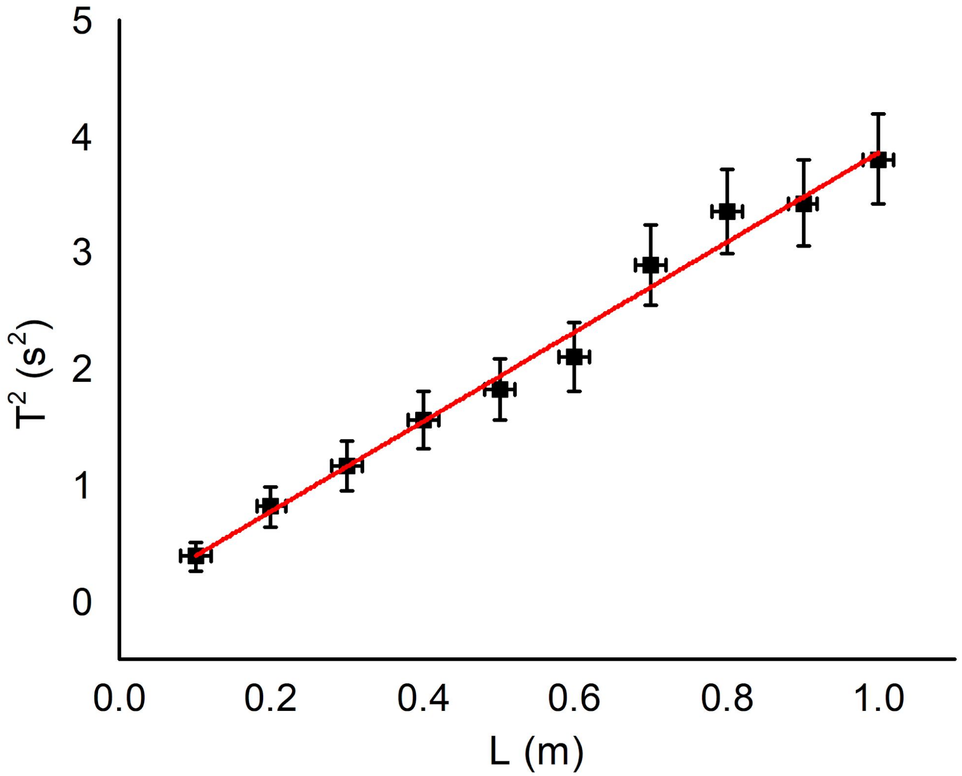
Statistics

	T^2
Number of Points	10
Degrees of Freedom	8
Reduced Chi-Sqr	0.20082
Residual Sum of Squares	1.60659
Pearson's r	0.99602
R Value	0.99602
R-Square(COD)	0.99205
Adj. R-Square	0.99105

universidad de buenos aires - exactas
departamento de física

Regresión lineal - cuadrados mínimos

- Método matemático para obtener la recta que "mejor ajusta" la nube de puntos. f(x)= **a** x + **b**



Parameters

		Value	Standard Error
T^2	Intercept	0.00197	0.05303
	Slope	3.85968	0.12218

a y **b**:
valor central y error (sin unidades, sin redondear)

Statistics

	T^2
Number of Points	10
Degrees of Freedom	8
Reduced Chi-Sqr	0.20082
Residual Sum of Squares	1.60659
Pearson's r	0.99602
R Value	0.99602
R-Square(COD)	0.99205
Adj. R-Square	0.99105

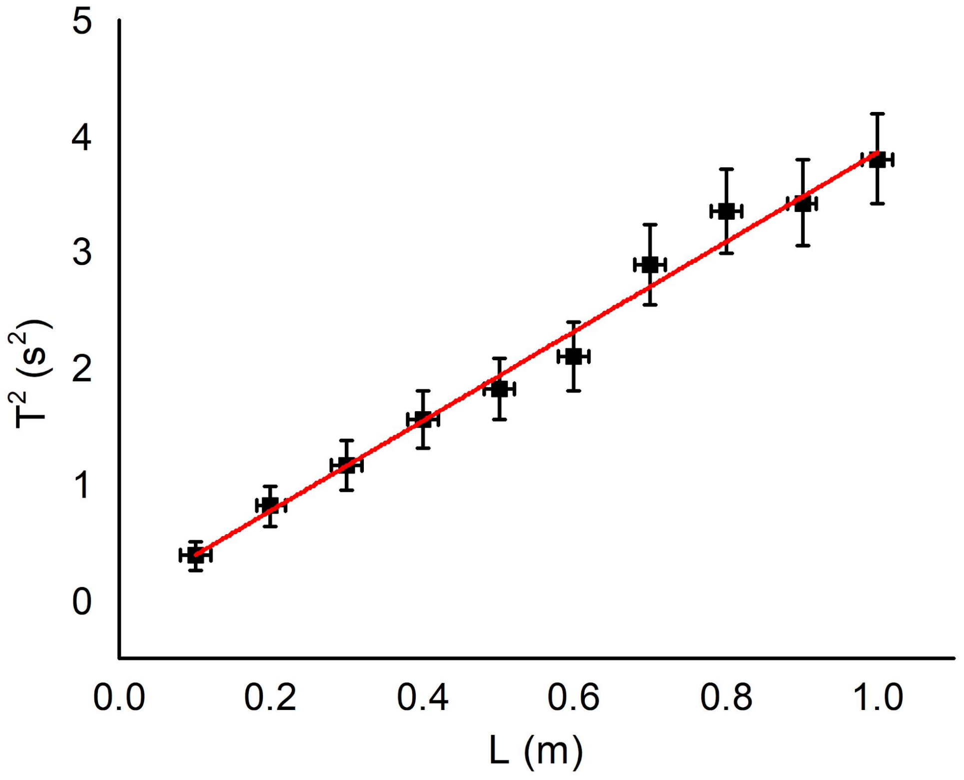
r, R^2, χ^2
calidad del ajuste

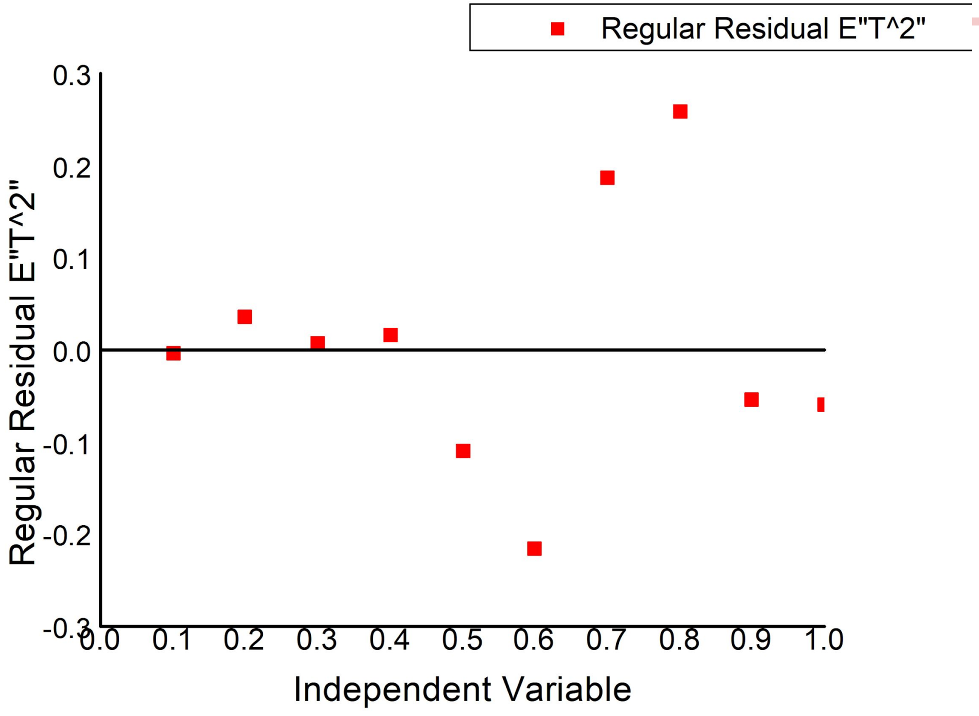
Regresión lineal - cuadrados mínimos

- Método matemático para obtener la recta que "mejor ajusta" la nube de puntos. f(x)= **a** x + **b**

más info sobre calidad del ajuste:

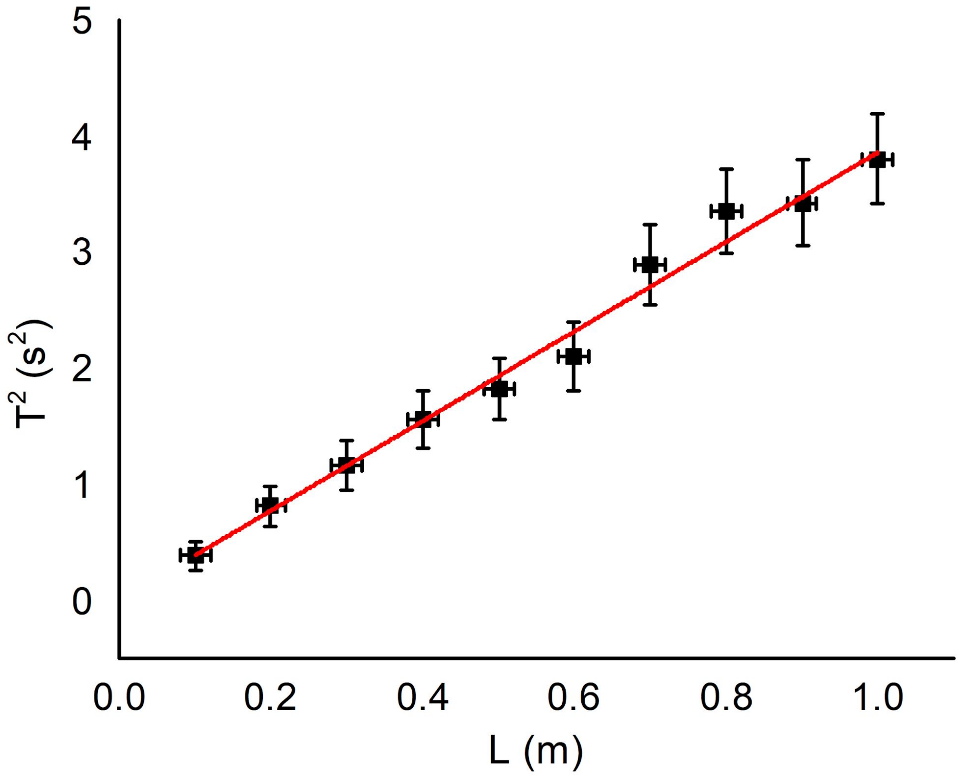
residuos: distancia entre cada punto y la recta





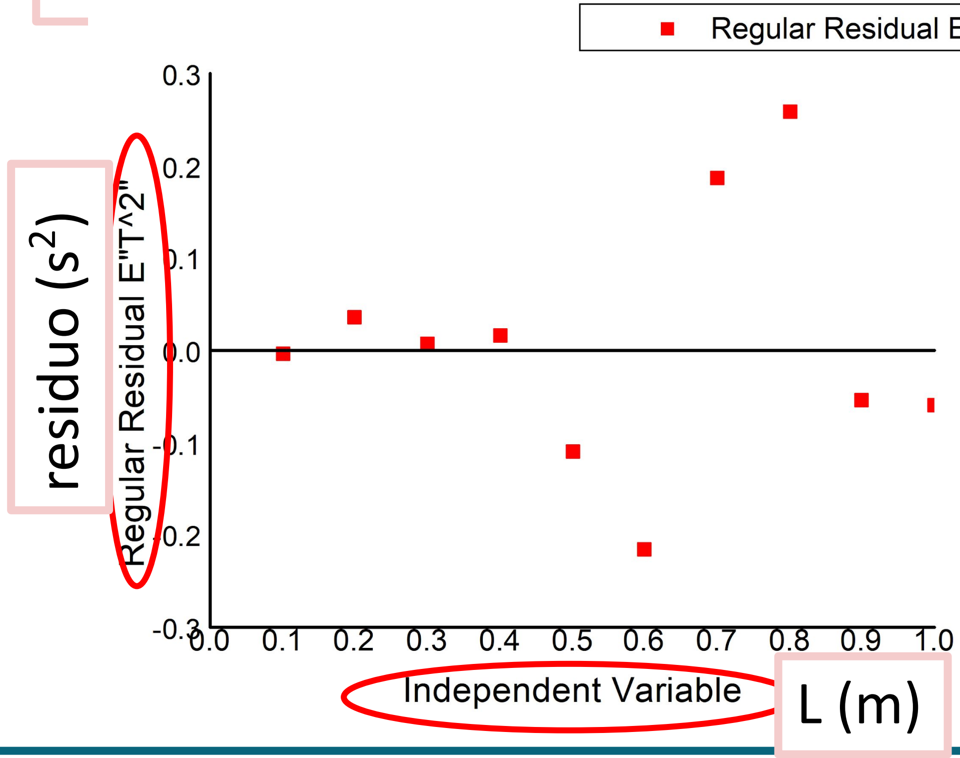
Regresión lineal - cuadrados mínimos

- Método matemático para obtener la recta que “mejor ajusta” la nube de puntos. f(x)= **a** x + **b**



más info sobre calidad del ajuste:

residuos: distancia entre cada punto y la recta



Ajuste por cuadrados mínimos

- El método de ajuste por cuadrados mínimos **busca los parámetros** del modelo propuesto que **minimizan la sumatoria de residuos**.

- Cuando la regresión es **lineal**, existe un cálculo análitico para calcular los coeficientes **a** y **b** y sus errores. (ver siguientes diapos)

- Si la regresión es no-lineal (cuadrática, gaussiana, etc) se usa un algoritmo iterativo (=método numérico) que repite unas operaciones hasta obtener los parámetros que minimizan la suma de residuos.

Hipótesis del modelo de regresión lineal simple

universidad de buenos aires - exactas
departamento de Física

*Diapos adaptadas de presentaciones de:
Dra. Silvia Goyanes
Dra. Adriana Márquez
Lic. Santiago Estévez Areco

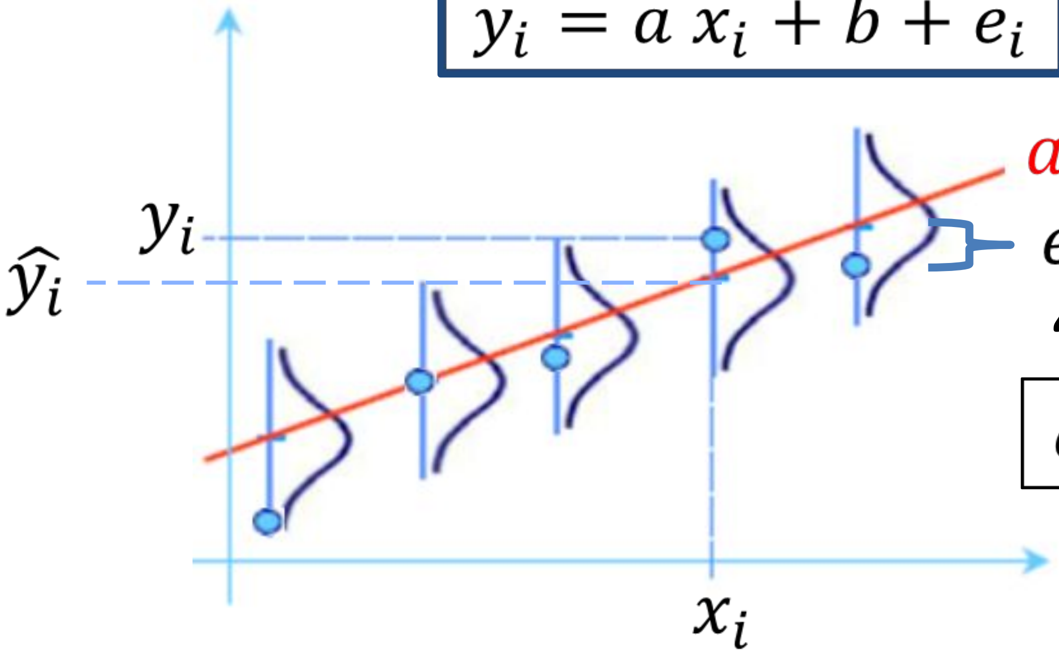
Independencia

- Los datos deben ser independientes.
- Una observación no debe dar información sobre las demás.

Normalidad

- Se asume que los datos son normales a priori.

$$\boxed{y_i = a\,x_i + b + e_i}$$



$a\,x_i + b = \hat{y}_i$

$e_i \sim N(0, \sigma^2)$

"Residuos"

$$\boxed{e_i \equiv y_i - \hat{y}_i}$$

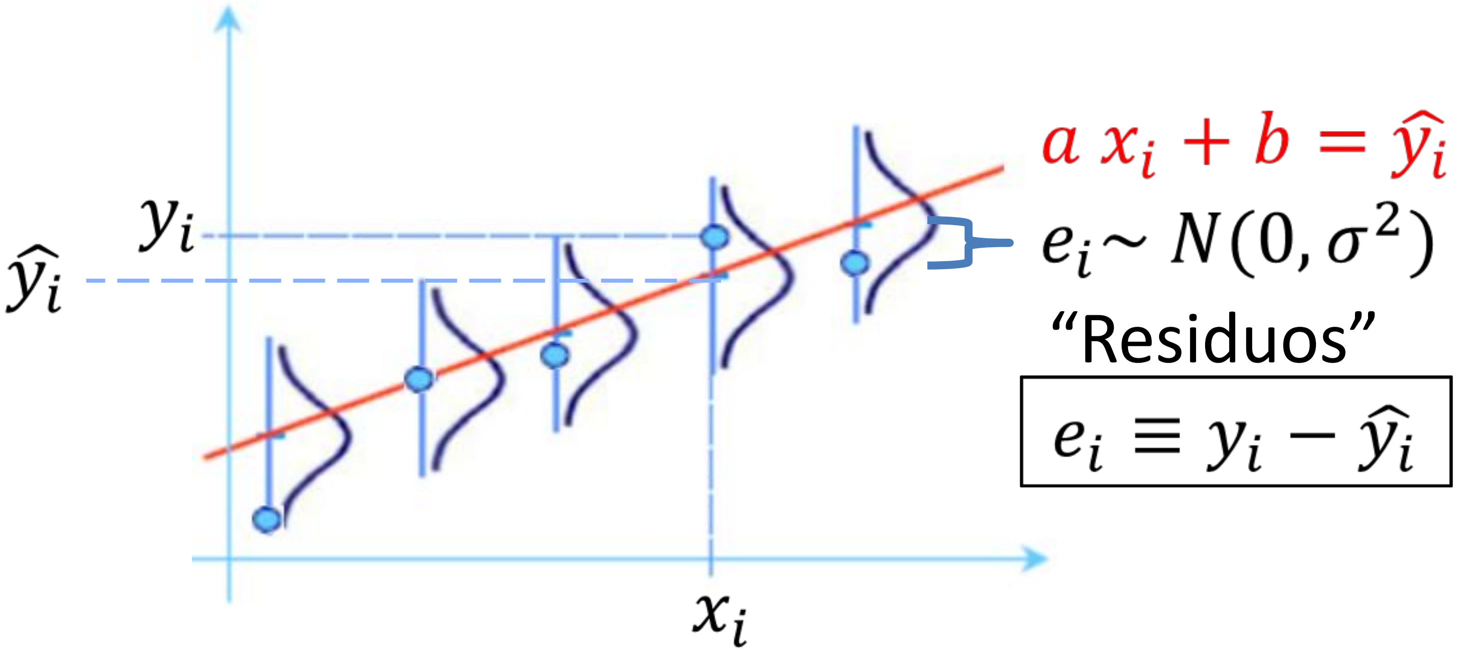
Estimadores de mínimos cuadrados

Gauss propuso en 1809 el método de mínimos cuadrados para obtener los valores $\boldsymbol{a}$ y **b** que mejor se ajustan a los datos:

$$\hat{y}_i = a\, x_i + b$$

El método consiste en minimizar la suma de los residuos al cuadrado (S^2):

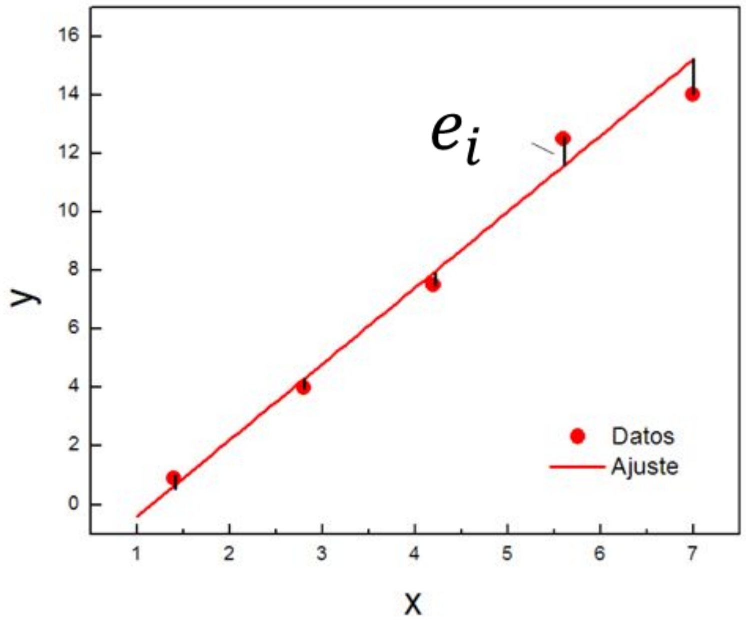
$$S^2 = \sum_i^N e_i^2 = \sum_i^N (y_i - \hat{y}_i)^2 = \sum_i^N [y_i - (a\, x_i + b)\,]^2$$



"Residuos"

$$\boxed{e_i \equiv y_i - \hat{y}_i}$$

Es decir, minimizar la suma de los cuadrados de las diferencias entre los valores reales observados (y_i) y los valores estimados ($\hat{y}_i$)



$$S^2 = \sum_i^N e_i^2 = \sum_i^N [y_i - (a.x_i + b)\,]^2$$

Busco minimizar $S^2(a, b)$

Incógnitas: *a* y *b*

¿Cómo cálculo a y b en el ajuste?
Minimizo S²

$$\left.\begin{aligned} \frac{\partial S^2}{\partial a}(a,b) &= 0 \\ \frac{\partial S^2}{\partial b}(a,b) &= 0 \end{aligned}\right\} \xrightarrow{\dots} \boxed{\begin{aligned} a &= \frac{N\sum(x_i\, y_i) - (\sum x_i).(\sum y_i)}{N\sum x_i{}^2 - (\sum x_i)^2} \\ b &= \frac{(\sum x_i{}^2)\,.\,(\sum y_i) - (\sum x_i).(\sum x_i y_i)}{N\sum x_i{}^2 - (\sum x_i)^2} \end{aligned}}$$

¿Cuáles son los errores de los parámetros obtenidos a y b?

Propagamos errores:

$$\sigma_a^2 = \sum \left(\frac{\partial a}{\partial y_i} \sigma_y \right)^2 \qquad \sigma_b^2 = \sum \left(\frac{\partial b}{\partial y_i} \sigma_y \right)^2$$

Si $\sigma_{y_1} = \sigma_{y_2} = \ldots\ldots = \sigma_{yN} = \sigma_y$

$$\boxed{\sigma_A = \sigma_y \sqrt{\frac{\sum x^2}{\Delta}}} \qquad \boxed{\sigma_B = \sigma_y \sqrt{\frac{N}{\Delta}}}$$

Donde: $\Delta = N \sum x^2 - \left(\sum x\right)^2$

Cuadrados mínimos solo considera errores en eje Y

- Se debe cumplir que:

$$\boxed{\Delta x \frac{dy}{dx} \ll \Delta y}$$

En caso contrario, **invertir los ejes**.

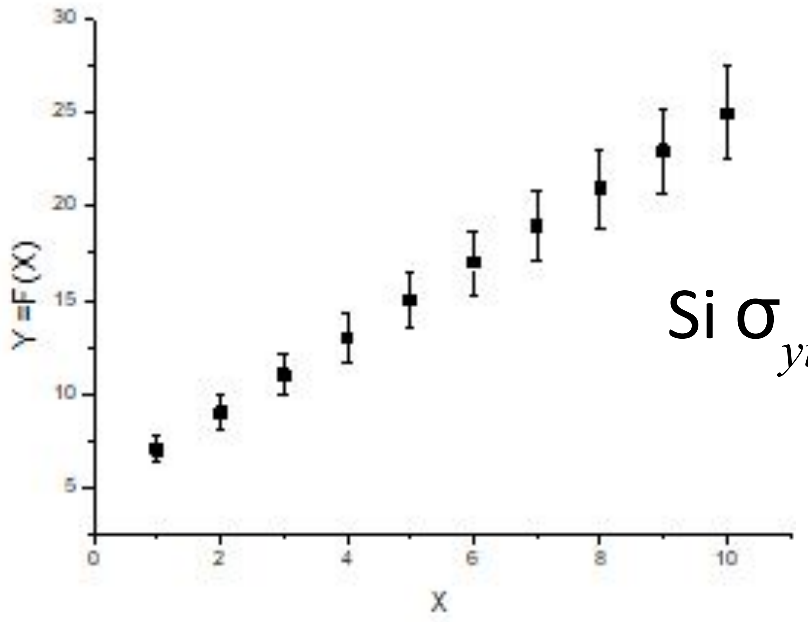
- Si aproximamos por una recta y **pasa por el origen**,

  entonces el criterio se puede expresar como: $\varepsilon_x \ll \varepsilon_y$

- Si los errores en x e y son del **mismo orden**, podemos "trasladar" el error en x al error en y:

$$\Delta y_{\mathrm{nuevo}}^2 = \left(\Delta x \frac{dy}{dx}\right)^2 + \Delta y_{original}^2$$

Hasta acá no tuvimos en cuenta el error de los datos. Algunos datos podrían ser más confiables que otros, y queremos que "pesen" más.



Si $\sigma_{yi} = \sigma_i$ son distintos:

Peso

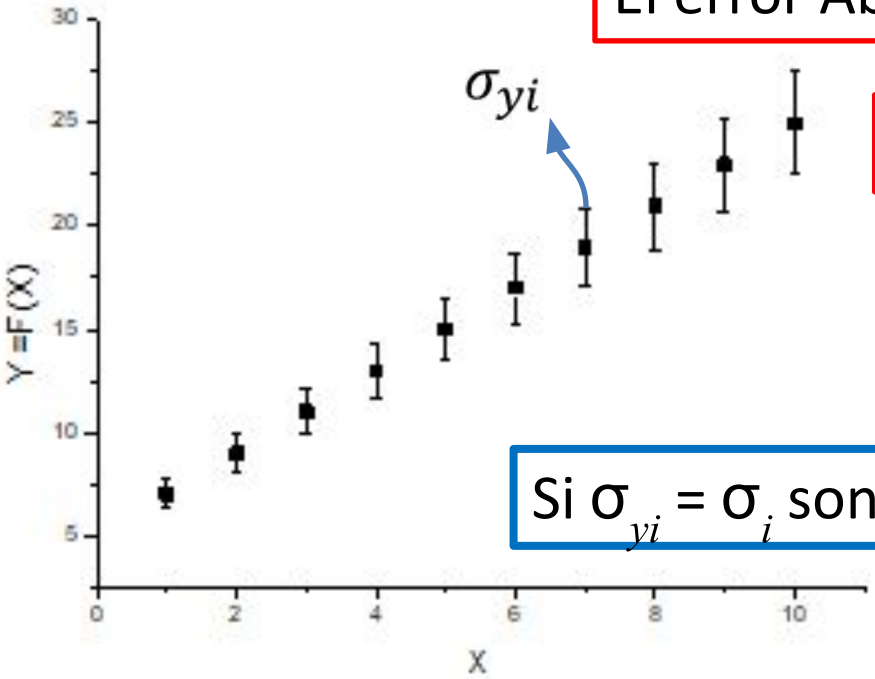
$$w_i = \frac{1}{\sigma_i^{\,2}}$$

Ajuste de cuadrados mínimos pesado con la incertidumbre

universidad de buenos aires - exactas
departamento de física

El error Absoluto cambia de dato a dato

¿Cómo tengo en cuenta eso?



Si $\sigma_{yi} = \sigma_i$ son distintos

CM PONDERADOS

$$w_i = \frac{1}{\sigma_i^{\,2}}$$

CM sin ponderar

$$S^2 = \sum_i^N e_i^2 = \sum_i^N [y_i - f(x_i)]^2$$

CM PONDERADOS

$$S^2 = \sum_i^N \left[\frac{y_i - f(x_i)}{\sigma_{yi}}\right]^2 = \sum_i^N \frac{[y_i - f(x_i)]^2}{\sigma_{yi}^{\,2}}$$

$$S^2 = \sum_i^N \left[\frac{y_i - f(x_i)}{\sigma_{yi}}\right]^2 = \sum_i^N \frac{[y_i - f(x_i)]^2}{\sigma_{yi}^2}$$

$$a = \frac{1}{\Delta}\left[\sum \frac{1}{\sigma_{yi}^2}\sum \frac{x_i y_i}{\sigma_{yi}^2} - \sum \frac{x_i}{\sigma_{yi}^2}\sum \frac{y_i}{\sigma_{yi}^2}\right]$$

$$b = \frac{1}{\Delta}\left[\sum \frac{x_i^2}{\sigma_{yi}^2}\sum \frac{y_i}{\sigma_{yi}^2} - \sum \frac{x_i}{\sigma_{yi}^2}\sum \frac{x_i y_i}{\sigma_{yi}^2}\right]$$

$$\Delta = \sum \frac{1}{\sigma_{yi}^2}\sum \frac{x_i^2}{\sigma_{yi}^2} - \left(\sum \frac{x_i}{\sigma_{yi}^2}\right)^2$$

Ajuste de cuadrados mínimos pesado con la incertidumbre

Si $\sigma_{yi} = \sigma_i$ son distintos

$$w_i = \frac{1}{\sigma_i^{\ 2}}$$

$$\Delta = \sum w \sum wx^2 - \left(\sum wx\right)^2 \qquad \Delta = N\sum x^2 - \left(\sum x\right)^2$$

$$\sigma_A = \sigma_y \sqrt{\frac{\sum x^2}{\Delta}} \qquad \sigma_B = \sigma_y \sqrt{\frac{N}{\Delta}}$$

$$\sigma_A = \sigma_y \sqrt{\frac{\sum wx^2}{\Delta}} \qquad \sigma_B = \sigma_y \sqrt{\frac{\sum w}{\Delta}}$$

Estimadores de la validez del modelo

- r: permite evaluar el grado de correlación lineal entre las dos variables

- R^2: permite evaluar la validez de determinada dependencia funcional (=del ajuste que hicimos) entre las variables

- Chi^2 reducido: mide la suma cuadrática de residuos. permite comparar dos ajustes distintos realizados sobre el mismo set de datos.

Coeficiente r de correlación lineal de Pearson

- Covarianza: cuantificación de la varianza conjunta de dos variables

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}, \quad \overline{y} = \frac{\sum_{i=1}^{n} y_i}{n}, \quad S_X^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n}, \quad S_Y^2 = \frac{\sum_{i=1}^{n} (y_i - \overline{y})^2}{n}, \quad S_{XY} = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{n}.$$

Promedio en x

Promedio en y

Varianza en x (SD^2)

Varianza en y

Covarianza: aumenta cuando la dispersión “va de la mano” en ambas variables

Coeficiente r de correlación lineal de Pearson

- Covarianza: cuantificación de la varianza conjunta de dos variables

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}, \quad \bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}, \quad S_X^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}, \quad S_Y^2 = \frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n}, \quad S_{XY} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{n}.$$

Promedio en x

Promedio en y

Varianza en x (SD^2)

Varianza en y

Covarianza: aumenta cuando la dispersión “va de la mano” en ambas variables

$$r = \frac{S_{XY}}{S_X S_Y},$$

Coeficiente r de correlación lineal de Pearson

- Covarianza: cuantificación de la varianza conjunta de dos variables

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}, \quad \bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}, \quad S_X^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}, \quad S_Y^2 = \frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n}, \quad S_{XY} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{n}.$$

Promedio en x

Promedio en y

Varianza en x (SD^2)

Varianza en y

Covarianza: aumenta cuando la dispersión "va de la mano" en ambas variables

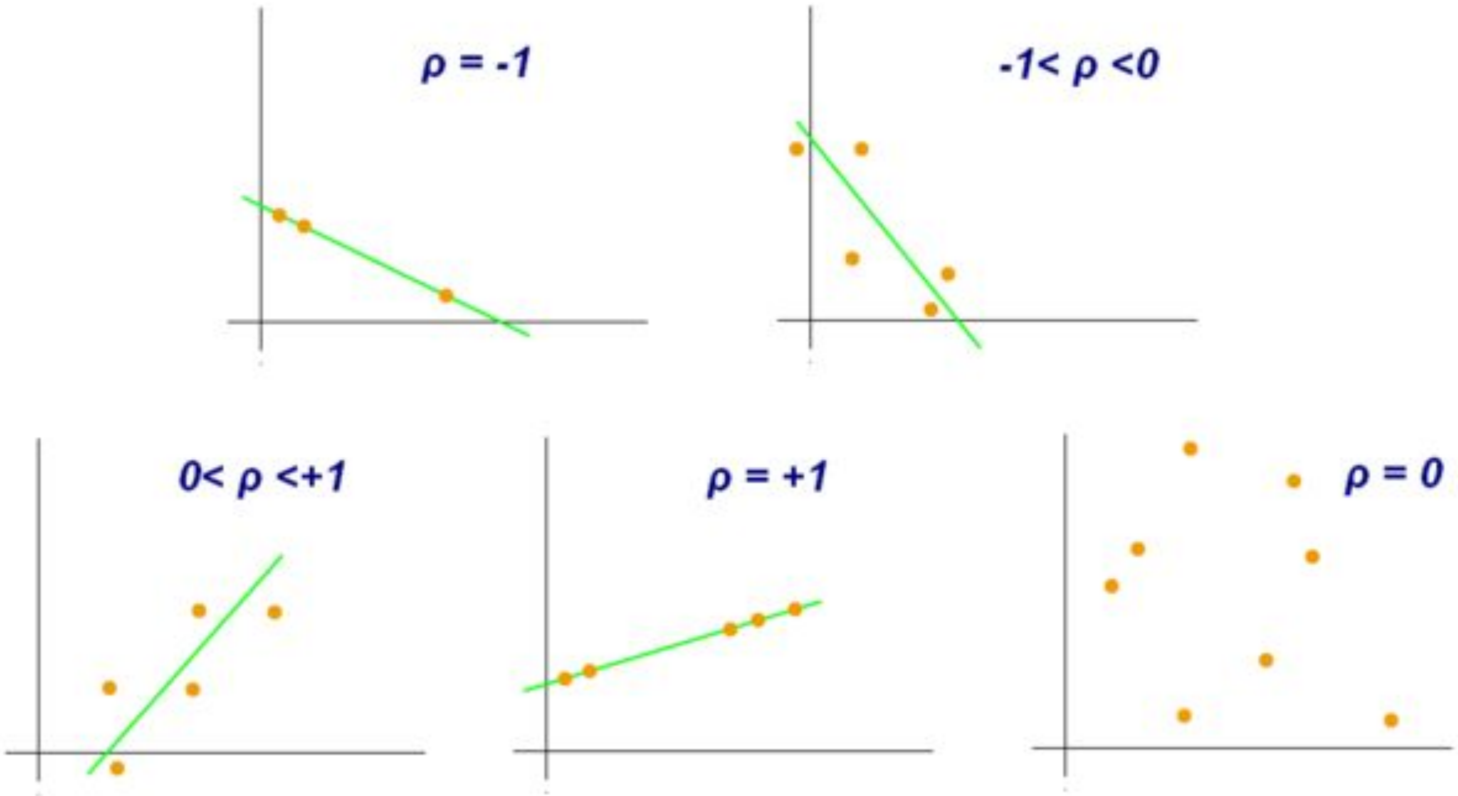
$$r = \frac{S_{XY}}{S_X S_Y}.$$

- No tiene dimensión, y siempre toma valores en [-1,1].
- Si r=1 o r=-1 la relación es perfectamente lineal
- Si las variables son independientes, entonces r=0, pero el inverso no tiene por qué ser cierto.
- Si $r > 0$, esto indica una relación directa entre las variables (si aumenta X, también aumenta Y).
- Si $r < 0$, la correlación entre las variables es inversa (si aumenta una, la otra disminuye).

Coeficiente r de correlación lineal de Pearson

$$r = \frac{S_{XY}}{S_X S_Y},$$

$$r = \frac{\Sigma (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\Sigma (x_i - \bar{x})^2 \Sigma (y_i - \bar{y})^2}}.$$



R^2 (R-square): Coeficiente de determinación

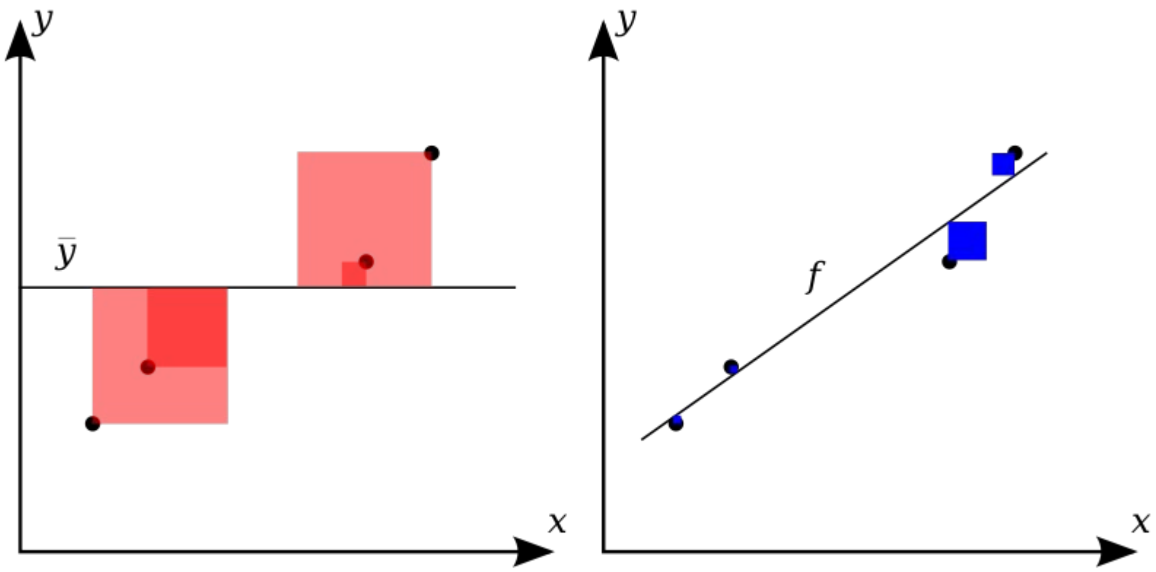
$$SS_{\text{tot}} = \sum_i (y_i - \bar{y})^2$$

Varianza en y (sin N)

$$SS_{\text{res}} = \sum_i (y_i - f_i)^2 = \sum_i e_i^2$$

Suma cuadrática de los residuos

$$\boxed{R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}}$$



Mide si pesa más la correlación entre los puntos o su dispersión

Mientras más cercano a 1 mejor es el ajuste

En una regresión lineal $R^2 = r^2$

χ_ν^2 (reduced chi-square)

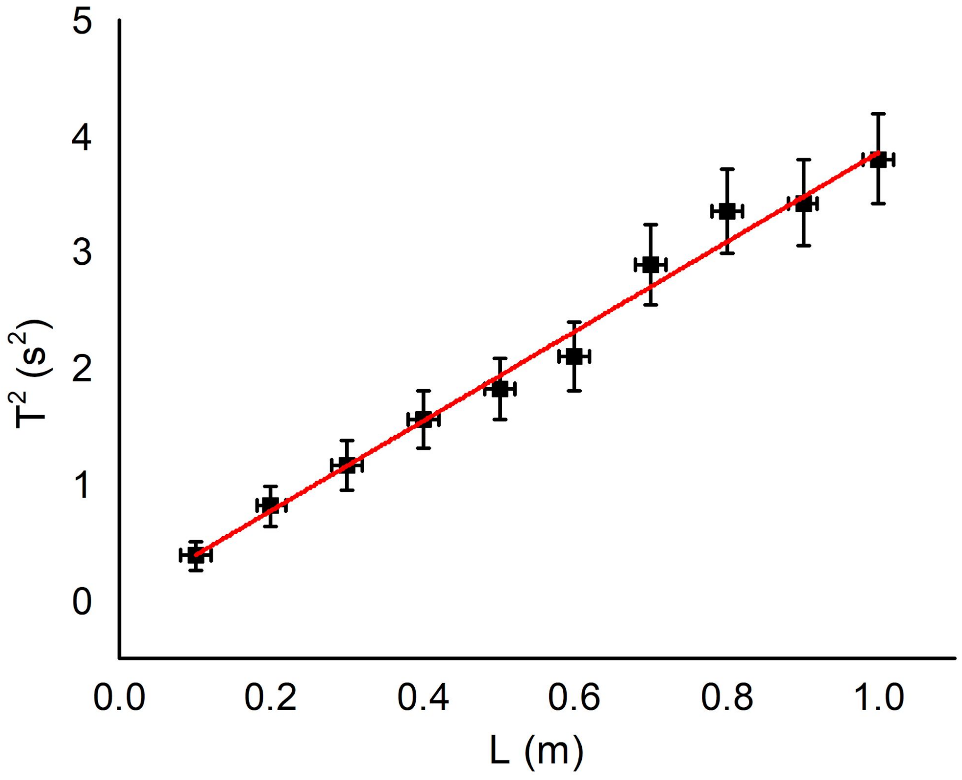
$$\chi_\nu^2 = \frac{1}{\nu}\sum_i \frac{(y_i - f_i)^2}{\sigma_{y_i}^2}$$

Equivale a la suma cuadrática de residuos ponderada por la incerteza en y de cada punto y dividido el número de grados de libertad (ν)
ν= número de mediciones (N) -número de parámetros

Los métodos de regresión no-lineal buscan minimizar esa cantidad

Retomando el ejemplo:

- Aplicamos una regresión lineal por cuadrados mínimos y obtuvimos: f(x)=a.x+b



Parameters

		Value	Standard Error
T^2	Intercept	0.00197	0.05303
	Slope	3.85968	0.12218

Statistics

	T^2
Number of Points	10
Degrees of Freedom	8
Reduced Chi-Sqr	0.20082
Residual Sum of Squares	1.60659
Pearson's r	0.99602
R Value	0.99602
R-Square(COD)	0.99205
Adj. R-Square	0.99105

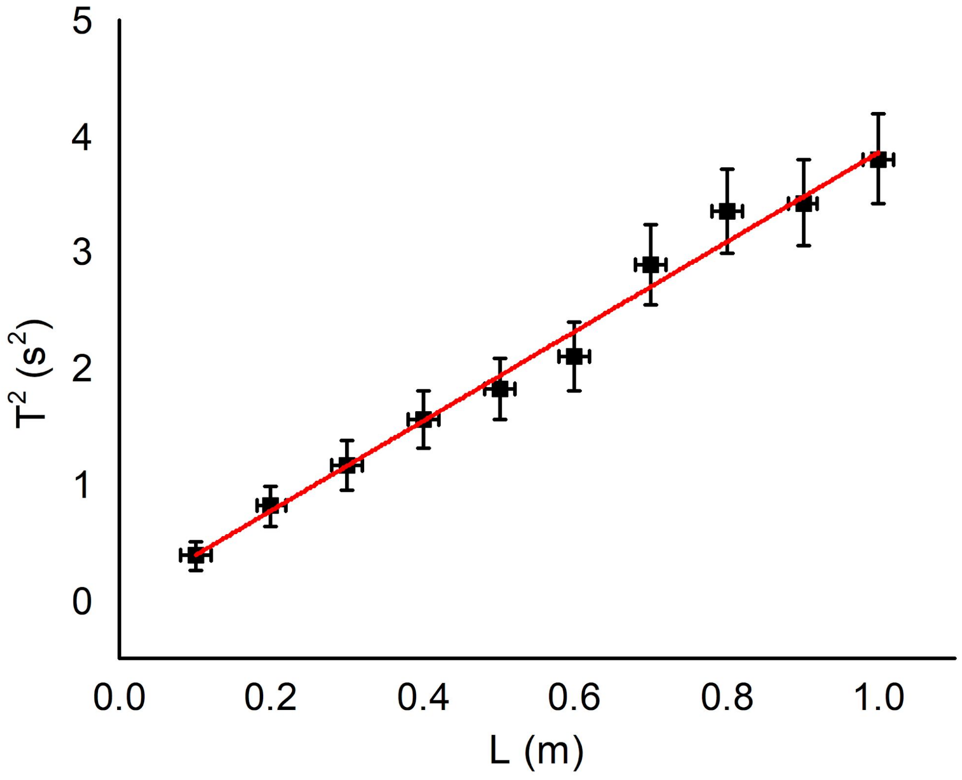
a= (3.86±0.12)
b= (0.002±0.053)

pregunta:¿unidad?

r= 0.99602
R^2= 0.99205

Retomando el ejemplo:

- Aplicamos una regresión lineal por cuadra obtuvimos: f(x)=a.x+b

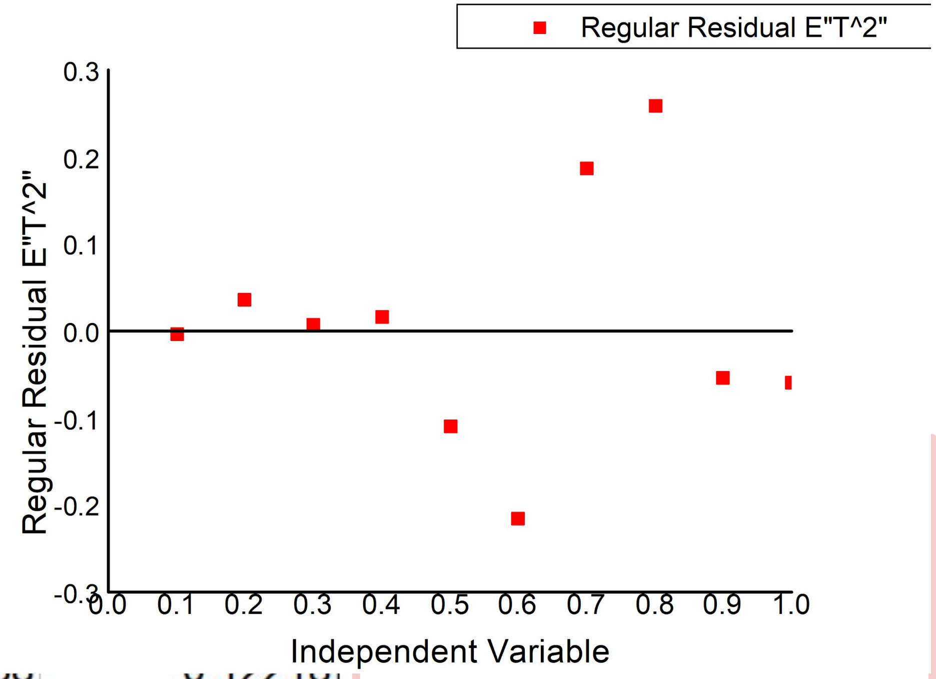


Parameters

		Value
T^2	Intercept	0.001
	Slope	3.859

Statistics

	T^2
Number of Points	10
Degrees of Freedom	8
Reduced Chi-Sqr	0.20082
Residual Sum of Squares	1.60659
Pearson's r	0.99602
R Value	0.99602
R-Square(COD)	0.99205
Adj. R-Square	0.99105



pregunta:¿unidad?

r= 0.99602

R^2= 0.99205

Retomando el ejemplo:

a= (3.86±0.12) s^2/m
b= (0.002±0.053)s^2

$$T = 2\pi\sqrt{\frac{L}{g}}$$

- ¿Cómo obtener **g** a partir de la pendiente? Valor central y prop. de errores
- ¿Qué rol cumplen **a** y **b** en el modelo de péndulo simple? $T^2 = \frac{4\pi^2}{g}L$

- Cálculo g: $g(a) = \frac{4\pi^2}{a}$
- Propagación de errores: $\Delta g = \left|\frac{d\,g(a)}{da}\right| \Delta a$

$$\Delta g = \frac{4\pi^2}{a^2}\,\Delta a$$

Retomando el ejemplo:

a= (3.86±0.12) s^2/m
b= (0.002±0.053)s^2

$$T = 2\pi\sqrt{\frac{L}{g}}$$

- ¿Cómo obtener **g** a partir de la pendiente? Valor central y prop. de errores
- ¿Qué rol cumplen **a** y **b** en el modelo de péndulo simple? $T^2 = \frac{4\pi^2}{g}L$

El parámetro **b** no interviene en el cálculo.

De todos modos da información sobre la validez del modelo.

- Cálculo g: $g(a) = \frac{4\pi^2}{a}$
- Propagación de errores: $\Delta g = \left|\frac{d\,g(a)}{da}\right| \Delta a$

$$\Delta g = \frac{4\pi^2}{a^2}\,\Delta a$$

Comparación entre regresión lineal y cálculo dato a dato:

- Opción 1: calcular g para cada una de las N mediciones. Promediar y calcular SD.

$$T = 2\pi\sqrt{\frac{L}{g}}$$

$$g(L,T) = \frac{4\pi^2}{T^2}L$$

$$\Delta g_{inst} = \sqrt{\left(\frac{d\,g(L,T)}{dL}\,\Delta L\right)^2 + \left(\frac{d\,g(L,T)}{dT}\,\Delta T\right)^2}$$

$$\Delta g_{inst} = \sqrt{\left(\frac{4\pi^2}{T^2}\,\Delta L\right)^2 + \left(2\frac{4\pi^2}{T^3}L\,\Delta T\right)^2}$$

$$\Delta g_{inst} = \sqrt{\left(\frac{4\pi^2}{T^2}\,\Delta L\right)^2 + \left(2\frac{g}{T}\,\Delta T\right)^2}$$

$$\Delta g = \sqrt{\Delta g_{inst}{}^2 + \left(\frac{\sigma}{\sqrt{N}}\right)^2}$$

- Nota: calcular g con una <u>única medición</u> de T y L ofrece con poca exactitud. No tiene en cuenta la fluctuación estadística.
- Opción 2: obtenerlo a partir de la regresión lineal.

Experimento 4: Determinación de g

- Leemos guía

- Debate:
  - Criterios para medir T y L
  - Incertezas de T y L
  - ¿Cuántas configuraciones distintas tomar para L?