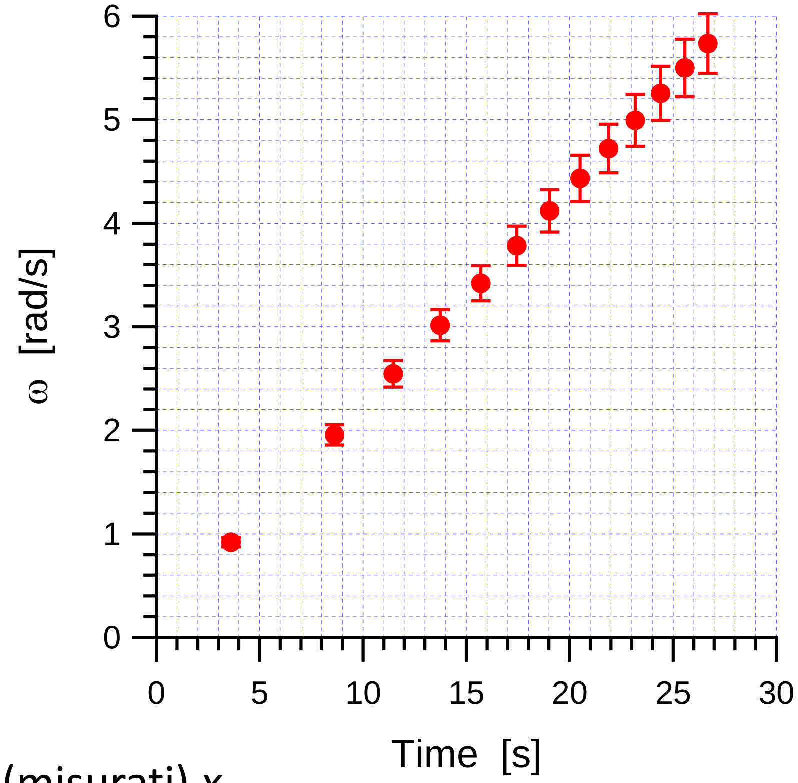
Best fit a una funzione

Problema:

1. Verificare la validità di una data legge (ovvero scegliere la legge giusta);
2. Determinare i valori dei parametri della legge (e la loro incertezza)

Esempio:
Nell'esperimento "Volano" voglio verificare la validità della legge $\omega(t) = \alpha_{tot} t$ e determinare il valore di α_{tot}



In generale:

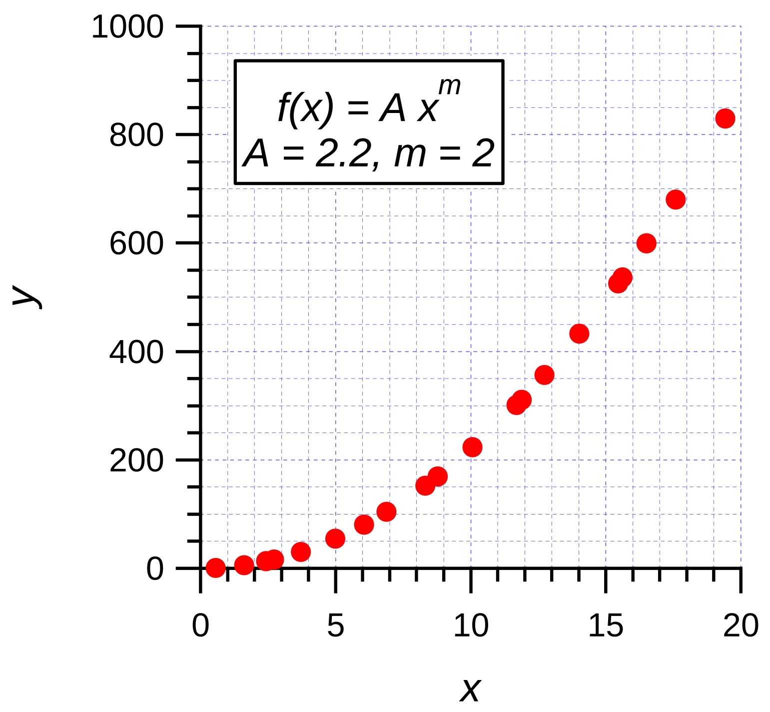
- ho un set di misure y_i corrispondenti a valori (misurati) x_i
-suppongo di conoscere una relazione funzionale $y = f(x)$
- voglio verificare la validità della legge e determinare i **parametri** della funzione $f(x)$

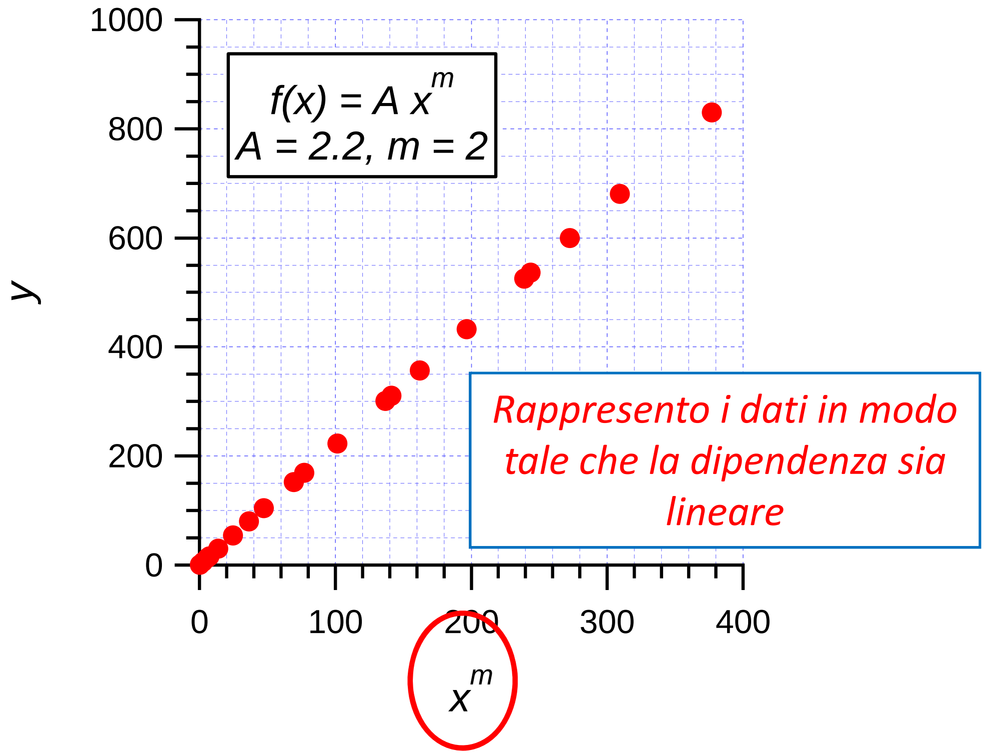
Nota: leggi di potenza

Caso più semplice: retta che passa per l'origine: $f(x) = ax$
Caso appena più complicato: retta con termine noto: $f(x) = ax + b$

<u>Attenzione</u>: talvolta è possibile "ricondursi" a forme lineari anche nel caso di funzioni nonlineari

Esempio: $f(x) = A\,x^m$
[e.g. legge oraria del moto in caduta libera con $v_0 = 0$: $\Delta s(t) = (a/2)t^2$
dipendenza del periodo di oscillazione di una molla con la massa m: $T(m) = 2\pi(m/k)^{1/2}$*]*



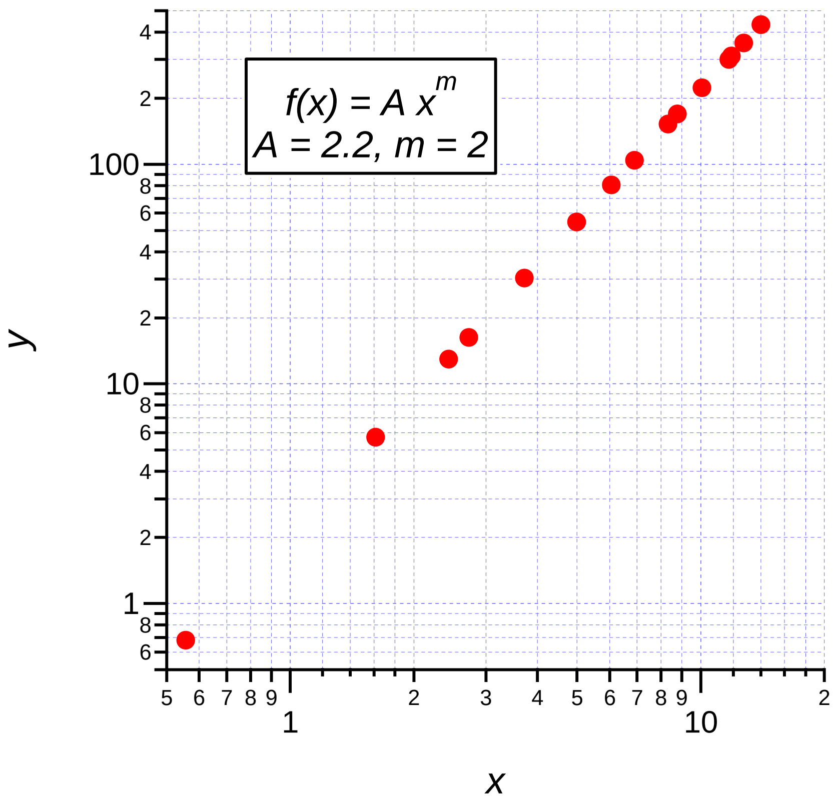


Nota: carta logaritmica (bilogaritmica)

Log Log

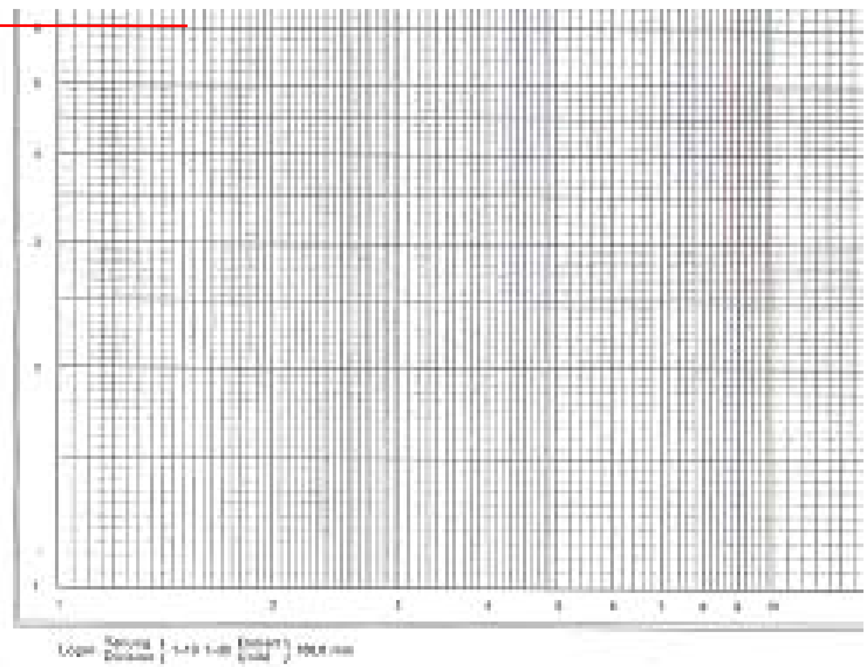
$f(x) = A\,x^m$

$\mathrm{Log}(f(x)) = \mathrm{Log}(A) + m\mathrm{Log}(x)$



Rappresentazione con andamento lineare dove
-termine noto è Log(A)
- pendenza è m

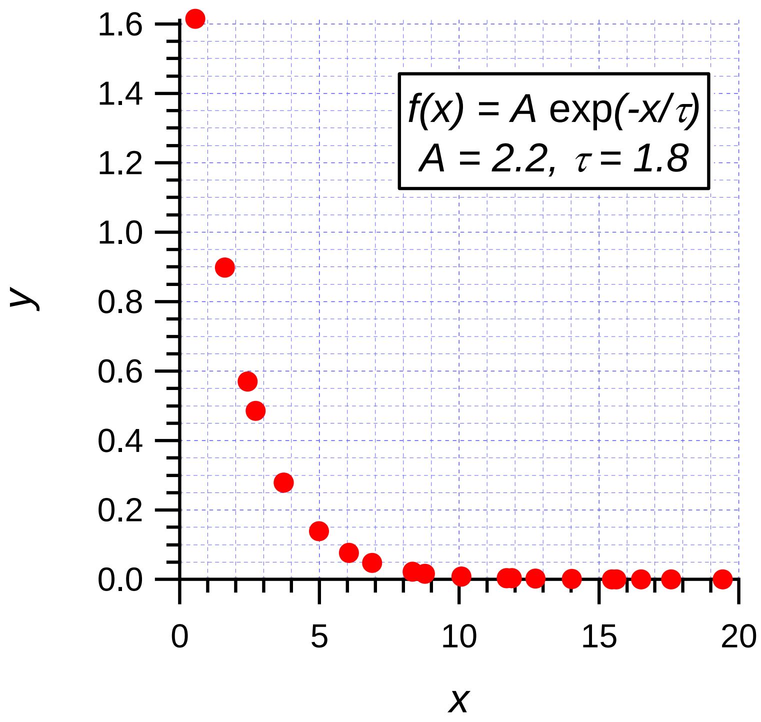
Carta (bi)logaritmica

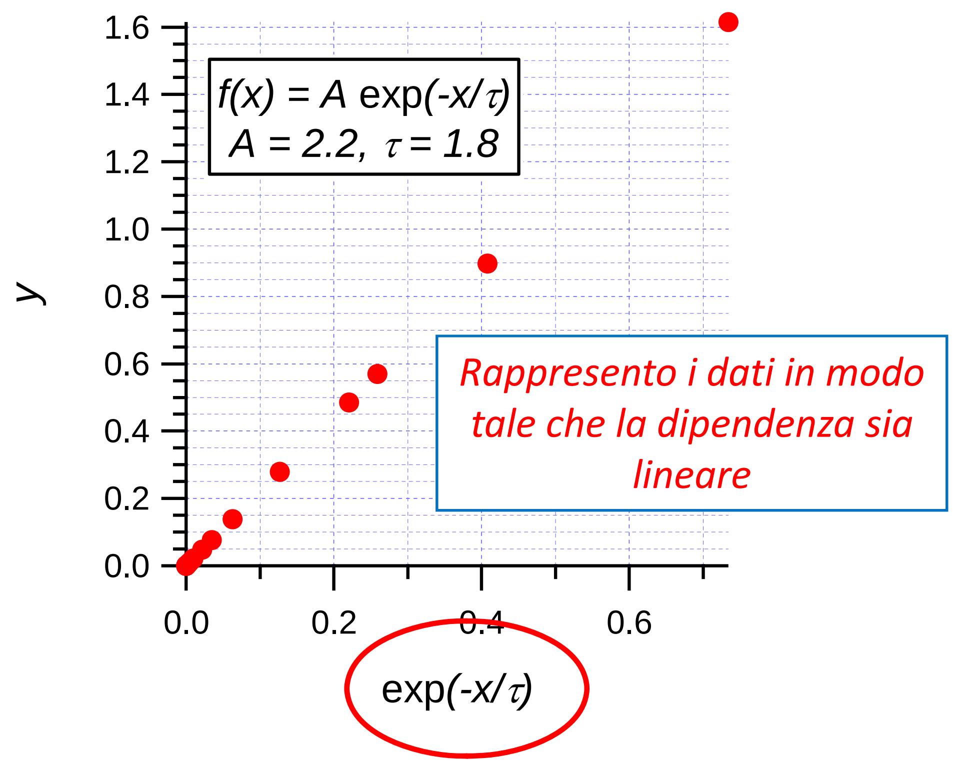


Usando la carta (bi)logaritmica rappresento $\mathrm{Log}(y_i)$ in funzione di $\mathrm{Log}(x_i)$

Nota: andamenti esponenziali

Esempio: *f(x) = A exp(x/B)*
[e.g. velocità nel moto in attrito viscoso $v(t) = v_0 \exp(-t/\tau)$*]*



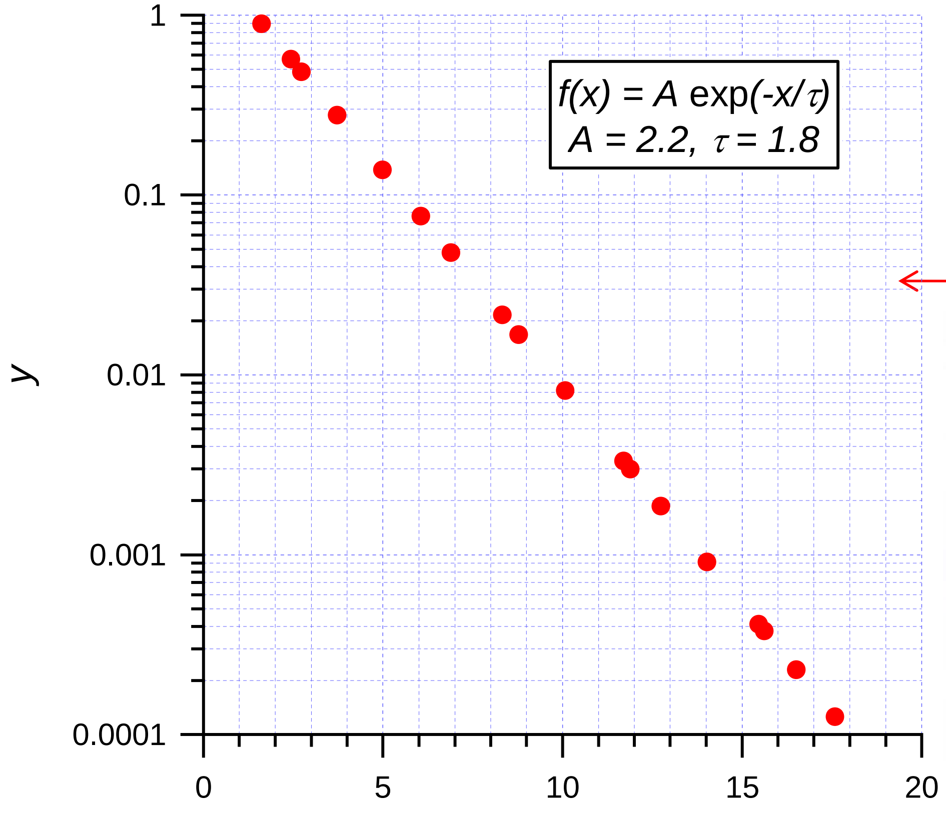


Nota: carta semilogaritmica

Log

$f(x) = A \exp(x/B)$

$\mathrm{Log}(f(x)) = \mathrm{Log}(A) + (1/B)\,x$

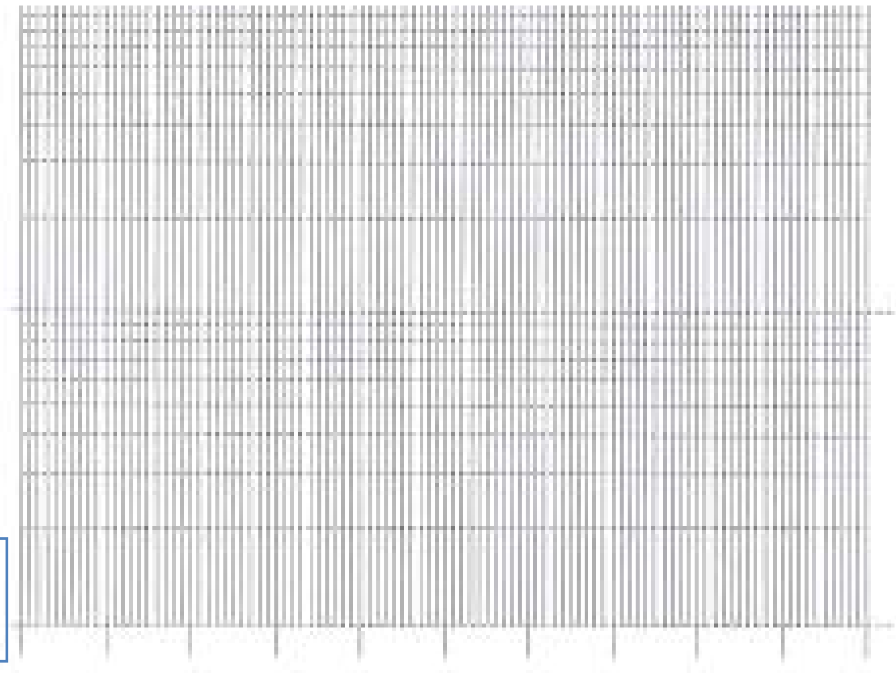


Rappresentazione con andamento lineare dove

-termine noto è Log(A)

- pendenza è 1/B

Carta semilogaritmica



Usando la carta semilogaritmica rappresento $\mathrm{Log}(y_i)$ in funzione x_i

Minimi quadrati

Definisco residuo: $\Delta_i = y_i - f(x_i)$

Costruisco: $S = \sum_i \Delta_i^2 = \sum_i (y_i - f(x_i))^2$

Cerco parametri di *f(x)* tali da minimizzare *S*

<u>Nota:</u> il metodo tratta tutte le y_i "allo stesso modo"

→ *Vale solo se gli errori sulle misure sono tutti paragonabili fra loro*

→ *(Vale solo se l'eventuale errore su x_i non si propaga "troppo" su y_i)*

Minimizzare → imporre derivata nulla rispetto ai parametri della funzione

In generale:

$$\begin{cases} \dfrac{\partial S}{\partial a} = 0 \\ \dfrac{\partial S}{\partial b} = 0 \\ \ldots. \end{cases}$$

a, b, … parametri della funzione

Non per tutte le funzioni f(x) è fattibile in modo analitico (servono computer e algoritmi numerici)

Minimi quadrati per una retta passante per l'origine

$$f(x) = ax$$

$$S = \sum_i (y_i - ax_i)^2$$

$$\text{Condizione di minimo}: \frac{dS}{da} = 0$$

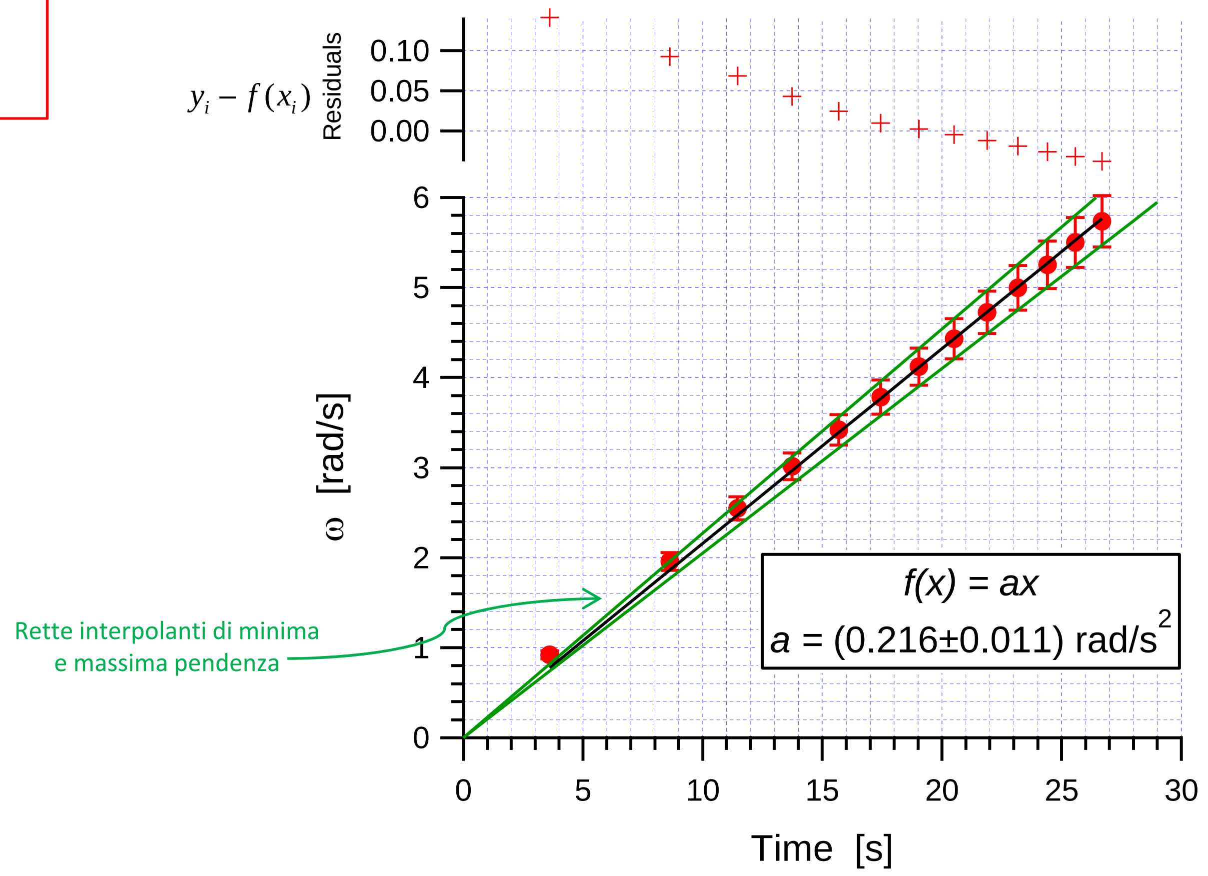
$$\rightarrow 0 = \sum_i -2\big(x_i(y_i - ax_i)\big) \rightarrow 0 = \sum_i x_i y_i - \sum_i a x_i^2$$

$$\boxed{\Rightarrow a = \frac{\sum_i x_i y_i}{\sum_i x_i^2}}$$

Se ho incertezza Δy_i:

$$\text{Propagazione degli errori (in quadratura)}:$$

$$(\Delta a)^2 = \sum_i \left(\frac{\partial a}{\partial y_i}\right)^2 (\Delta y_i)^2 = \frac{\sum_i x_i^2 \Delta y_i^2}{\left(\sum_i x_i^2\right)^2}$$



Minimi quadrati per una funzione costante

Supponiamo di eseguire tante misure della stessa grandezza *k* (costante):
f(x) = k *(k è l'unico parametro di questa funzione!)*

$$S = \sum_i (y_i - k)^2$$

$$\frac{dS}{dk} = -2\sum_i (y_i - k) = 0$$

$$\rightarrow \sum_i y_i = \sum_i k = Nk \quad \text{con } N \text{ numero delle misure}$$

$$\Rightarrow \boxed{k = \sum_i \frac{y_i}{N} = \overline{y}}$$

La media aritmetica è il miglior fit della funzione costante

$$(\Delta k)^2 = \sum_i \left(\frac{dk}{dy_i}\right)^2 (\Delta y_i)^2 = \sum_i \left(\frac{\Delta y_i}{N}\right)^2 = \left(\frac{\frac{\sum_i (y_i - \overline{y})}{N}}{N}\right)^2 = \left(\frac{\sigma_y}{N}\right)^2 = \left(\sigma_{\overline{y}}\right)^2$$

L'incertezza è l'"errore sulla media" $\sigma_{\overline{y}}$

Minimi quadrati per una retta che non passa per l'origine

$f(x) = ax + b$ *a, b sono due parametri)*

$$S = \sum_i (y_i - ax_i - b)^2$$

$$\text{Condizione di minimo}: \begin{cases} \dfrac{\partial S}{\partial a} = 0 = -2\sum_i (y_i - ax_i - b)x_i \\ \dfrac{\partial S}{\partial b} = 0 = -2\sum_i (y_i - ax_i - b) \end{cases}$$

$$\rightarrow \begin{cases} Nb + a\sum_i x_i = \sum_i y_i \\ b\sum_i x_i + a\sum_i x_i^2 = \sum_i x_i y_i \end{cases}$$

$$\Rightarrow \boxed{\begin{cases} a = \dfrac{N\sum x_i y_i - \sum x_i \sum y_i}{N\sum x_i^2 - \left(\sum x_i\right)^2} \\ b = \dfrac{\sum y_i \sum x_i^2 - \sum x_i \sum x_i y_i}{N\sum x_i^2 - \left(\sum x_i\right)^2} \end{cases}}$$

Formule complicate, ma ancora di tipo analitico!

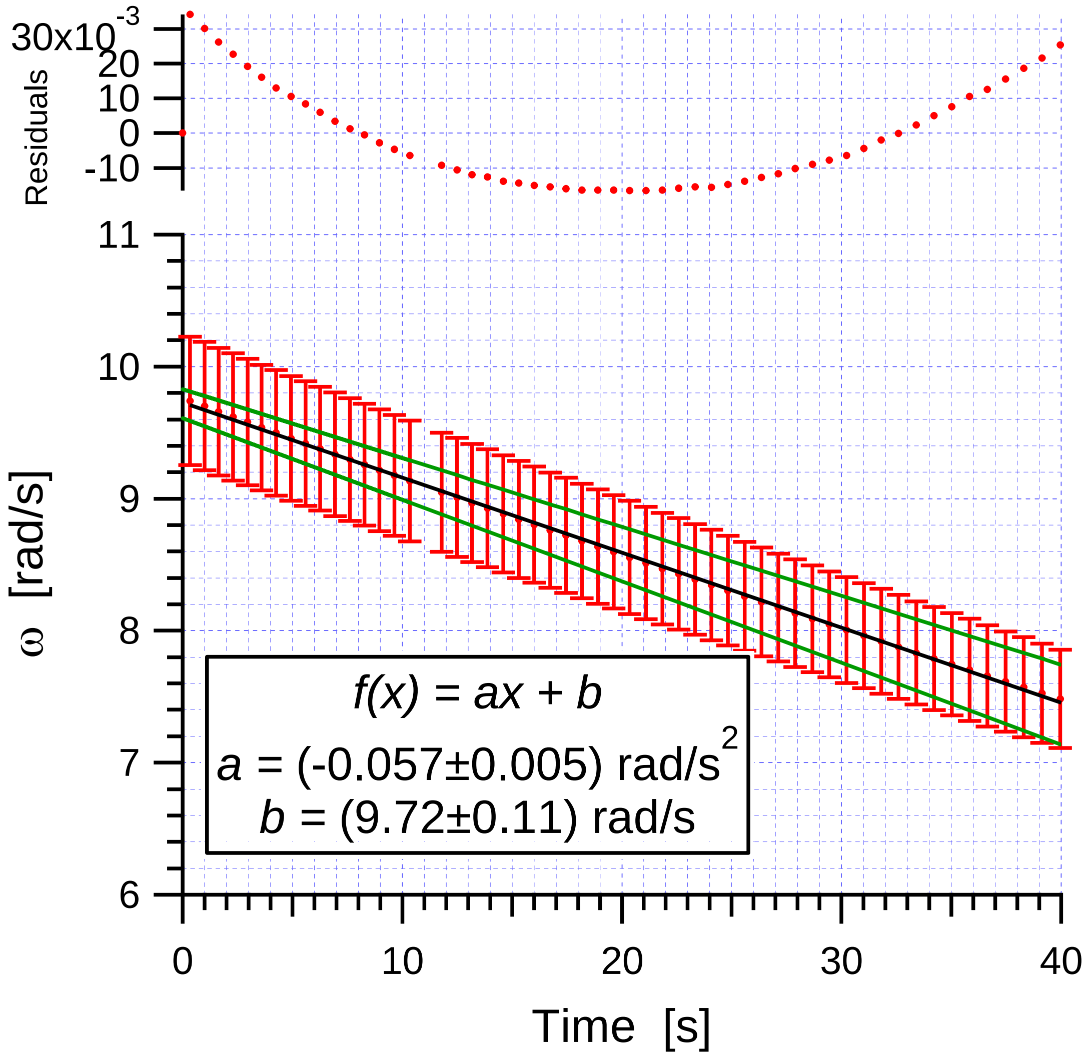
Propagazione degli errori (in quadratura)

$$\begin{cases} (\Delta a)^2 = \dfrac{N\sigma_y^2}{N\sum x_i^2 - \left(\sum x_i\right)^2} \\ (\Delta b)^2 = \dfrac{\sigma_y^2 \sum x_i^2}{N\sum x_i^2 - \left(\sum x_i\right)^2} \end{cases}$$

$$\text{con}: \ \sigma_y^2 = \frac{\sum_i (\Delta y_i)^2}{N}$$

Best fit per una retta che non passa per l'origine

Esempio:
Nell'esperimento "Volano" voglio verificare la validità della legge $\omega(t) = \omega_0 + \alpha_A t$ e determinare il valore α_A



Best fit che tiene conto dell'incertezza sperimentale

$$\text{Definisco residuo "pesato" per gli errori}: \Delta_i = \frac{y_i - f(x_i)}{\sigma_i}$$

con σ_i deviazione standard (errore) delle misure

$$\text{Costruisco}: S = \sum_i \Delta_i^{\ 2} = \sum_i \frac{(y_i - f(x_i))^2}{\sigma_i^{\ 2}}$$

e minimizzo in funzione dei parametri

$$\begin{cases} \dfrac{\partial S}{\partial a} = 0 \\ \dfrac{\partial S}{\partial b} = 0 \\ \dots \end{cases}$$

Nel caso in cui si abbia incertezza anche per le grandezze "indipendenti" x_i è formalmente possibile costruire una deviazione standard σ_i che tiene conto anche di queste incertezze

<u>Attenzione</u>: l'incertezza su x non deve in nessun caso produrre effetti troppo marcati su y, cioè, per la propagazione degli errori:: $\left(\left(\frac{\partial f}{\partial x} \right)_{x_i} \Delta x_i \right)^2 < (\Delta y_i)^2$

Se si assumono y_i distribuite in modo normale (cioè secondo distribuzione Gaussiana), i residui Δ_i sono variabili Gaussiane con media nulla e varianza unitaria!

La funzione *S* segue allora la distribuzione del **chi-quadro**

Vantaggi del metodo del minimo chi-quadro

Il metodo del minimo χ^2 consente di tenere in conto incertezze differenti per le varie misure

Inoltre permette di usare il **test del chi-quadro** per valutare la "verosimiglianza" del fit

La funzione $S = \sum_{i=1}^{n} \left(\frac{y_i - f(x_i)}{\sigma_i} \right)^2$ è un chi-quadro ad *n* **gradi di libertà**

Pertanto *S* è distribuita secondo la distribuzione del chi-quadro ad *n* gradi di libertà, cioè:

$$\mathrm{p}(n, S) = C_n S^{\frac{n-2}{2}} e^{-\frac{S}{2}}$$

$$\text{con } C_n = \frac{1}{2^{n/2}\Gamma(n/2)}$$

$$\text{con}:$$

$$E(S) = n \quad \text{media}$$

$$\sigma_S{}^2 = 2n \quad \text{varianza}$$

Esistono delle tabelle che dicono qual è il valore γ di *S* tale che la probabilità sia p (per un dato valore di *n*): $\mathrm{p}(n, S \leq \gamma) = \mathrm{p}$

Le tabelle dunque dicono con quale probabilità il chi-quadro, cioè la somma dei residui pesata per le incertezze, è al di sotto di un certo **valore di confidenza**

Test del chi-quadro

	99.5%	99%	97.5%	95%	90%	75%	50%	25%	10%	5%	2.5%	1%	0.5%
n	$\chi^2_{.995}$	$\chi^2_{.99}$	$\chi^2_{.975}$	$\chi^2_{.95}$	$\chi^2_{.90}$	$\chi^2_{.75}$	$\chi^2_{.5}$	$\chi^2_{.25}$	$\chi^2_{.10}$	$\chi^2_{.05}$	$\chi^2_{.025}$	$\chi^2_{.01}$	$\chi^2_{.005}$
1	7.88	6.63	5.02	3.84	2.71	1.32	.455	.102	.0158	.0039	.0010	.0002	.0000
2	10.6	9.21	7.38	5.99	4.61	2.77	1.39	.575	.211	.103	.0506	.0201	.0100
3	12.8	11.3	9.35	7.81	6.25	4.11	2.37	1.21	.584	.352	.216	.155	.072
4	14.9	13.3	11.1	9.49	7.78	5.39	3.36	1.92	1.06	.711	.484	.297	.207
5	16.7	15.1	12.8	11.1	9.24	6.63	4.35	2.67	1.61	1.15	.831	.554	.412
6	18.5	16.8	14.4	12.6	10.6	7.84	5.35	3.45	2.20	1.64	1.24	.872	.676
7	20.3	18.5	16.0	14.1	12.0	9.04	6.35	4.25	2.83	2.17	1.69	1.24	.989
8	22.0	20.1	17.5	15.5	13.4	10.2	7.34	5.07	3.49	2.73	2.18	1.65	1.34
9	23.6	21.7	19.0	16.9	14.7	11.4	8.34	5.90	4.17	3.33	2.70	2.09	1.73
10	25.2	23.2	20.5	18.3	16.0	12.5	9.34	6.74	4.87	3.94	3.25	2.56	2.16
11	26.8	24.7	21.9	19.7	17.3	13.7	10.3	7.58	5.58	4.57	3.82	3.05	2.60
12	28.3	26.2	23.3	21.0	18.5	14.8	11.3	8.44	6.30	5.23	4.40	3.57	3.07
13	20.8	27.7	24.7	22.4	19.8	16.0	12.3	9.30	7.04	5.89	5.01	4.11	3.57
14	31.3	29.1	26.1	23.7	21.1	17.1	13.3	10.2	7.79	6.57	5.63	4.66	4.07
15	32.8	30.6	27.5	25.0	22.3	18.2	14.3	11.0	8.55	7.26	6.26	5.23	4.60
16	34.3	32.0	28.8	26.3	23.5	19.4	15.3	11.9	9.31	7.96	6.91	5.81	5.14
17	35.7	33.4	30.2	27.6	24.8	20.5	16.3	12.8	10.1	8.67	7.56	6.41	5.70
18	37.2	34.8	31.5	28.9	26.0	21.6	17.3	13.7	10.9	9.39	8.23	7.01	6.26
19	38.6	36.2	32.9	30.1	27.2	22.7	18.3	14.6	11.7	10.1	8.91	7.63	6.84
20	40.0	37.6	34.2	31.4	28.4	23.8	19.3	15.5	12.4	10.9	9.59	8.26	7.43
21	41.4	38.9	35.5	32.7	29.6	24.9	20.3	16.3	13.2	11.6	10.3	8.90	8.03
22	42.8	40.3	36.8	33.9	30.8	26.0	21.3	17.2	14.0	12.3	11.0	9.54	8.64
23	44.2	41.6	38.1	35.2	32.0	27.1	22.3	18.1	14.8	13.1	11.7	10.2	9.26
24	45.6	43.0	39.4	36.4	33.2	28.2	23.3	19.0	15.7	13.8	12.4	10.9	9.89
25	46.9	44.3	40.6	37.7	34.4	29.3	24.3	19.9	16.5	14.6	13.1	11.5	10.5
26	48.3	45.6	41.9	38.9	35.6	30.4	25.3	20.8	17.3	15.4	13.8	12.2	11.2
27	49.6	47.0	43.2	40.1	36.7	31.5	26.3	21.7	18.1	16.2	14.6	12.9	11.8
28	51.0	48.3	44.5	41.3	37.9	32.6	27.3	22.7	18.9	16.9	15.3	13.6	12.5
29	52.3	49.6	45.7	42.6	39.1	33.7	28.3	23.6	19.8	17.7	16.0	14.3	13.1

Esempio:
Suppongo n = 10
Se χ^2 = 18.3 allora, ripetendo il best fit (con altri parametri o con altra legge), nel 95% dei casi troverei χ^2 < 18.3
→ La "confidenza" del mio χ^2 è 100-95=5%

Esempio:
Suppongo n = 20
Se χ^2 = 12.4 allora, ripetendo il best fit (con altri parametri o con altra legge), nel 10% dei casi troverei χ^2 < 12.4
→ La "confidenza" del mio χ^2 è 100-10=90%

Conoscere il valore del chi-quadro permette di conoscere la probabilità che, rifacendo il best fit (altri valori dei parametri, altra legge) questo venga "meglio" di quello che si è fatto, cioè con un valore del chi-quadro minore di quello originariamente determinato → si attribuisce un "valore di confidenza" al best fit

Il best fit con minimo chi-quadro è noto anche come "maximum likelihood"

Alcune formule per chi-quadro

❑ **Esempio**: Prendiamo come funzione $f(x) = a + bx$.

$$S = \sum_{i=1}^{n} \left(\frac{Y_i - a - bX_i}{\sigma_i} \right)^2 = \text{minimo}$$

retta

$$\begin{cases} \dfrac{\partial S}{\partial a} = -2 \sum \dfrac{1}{\sigma_i} \left(\dfrac{Y_i - a - bX_i}{\sigma_i} \right) = 0 \\ \dfrac{\partial S}{\partial b} = -2 \sum \dfrac{x_i}{\sigma_i} \left(\dfrac{Y_i - a - bX_i}{\sigma_i} \right) = 0 \end{cases} \qquad \begin{cases} a \sum \dfrac{1}{\sigma_i^2} + b \sum \dfrac{X_i}{\sigma_i^2} = \sum \dfrac{Y_i}{\sigma_i^2} \\ a \sum \dfrac{X_i}{\sigma_i^2} + b \sum \dfrac{X_i^2}{\sigma_i^2} = \sum \dfrac{Y_i X_i}{\sigma_i^2} \end{cases}$$

$$\begin{cases} a = \dfrac{\sum \frac{Y_i}{\sigma_i^2} \cdot \sum \frac{X_i^2}{\sigma_i^2} - \sum \frac{X_i}{\sigma_i^2} \cdot \sum \frac{X_i Y_i}{\sigma_i^2}}{\sum \frac{1}{\sigma_i^2} \cdot \sum \frac{X_i^2}{\sigma_i^2} - \left(\sum \frac{X_i}{\sigma_i^2} \right)^2} \\ \\ b = \dfrac{\sum \frac{1}{\sigma_i^2} \cdot \sum \frac{X_i Y_i}{\sigma_i^2} - \sum \frac{Y_i}{\sigma_i^2} \cdot \sum \frac{X_i}{\sigma_i^2}}{\sum \frac{1}{\sigma_i^2} \cdot \sum \frac{X_i^2}{\sigma_i^2} - \left(\sum \frac{X_i}{\sigma_i^2} \right)^2} \end{cases}$$

❑ **Esempio**: Sia $f(x) = \text{Cte} = a$. Applichiamo il metodo appena imparato:

costante

$$S = \sum_{i=1}^{n} \left(\frac{Y_i - a}{\sigma_i} \right)^2 = \text{minimo}$$

$$\frac{dS}{da} = -2 \sum_{i=1}^{n} \left(\frac{Y_i - a}{\sigma_i} \right) \frac{1}{\sigma_i} = 0$$

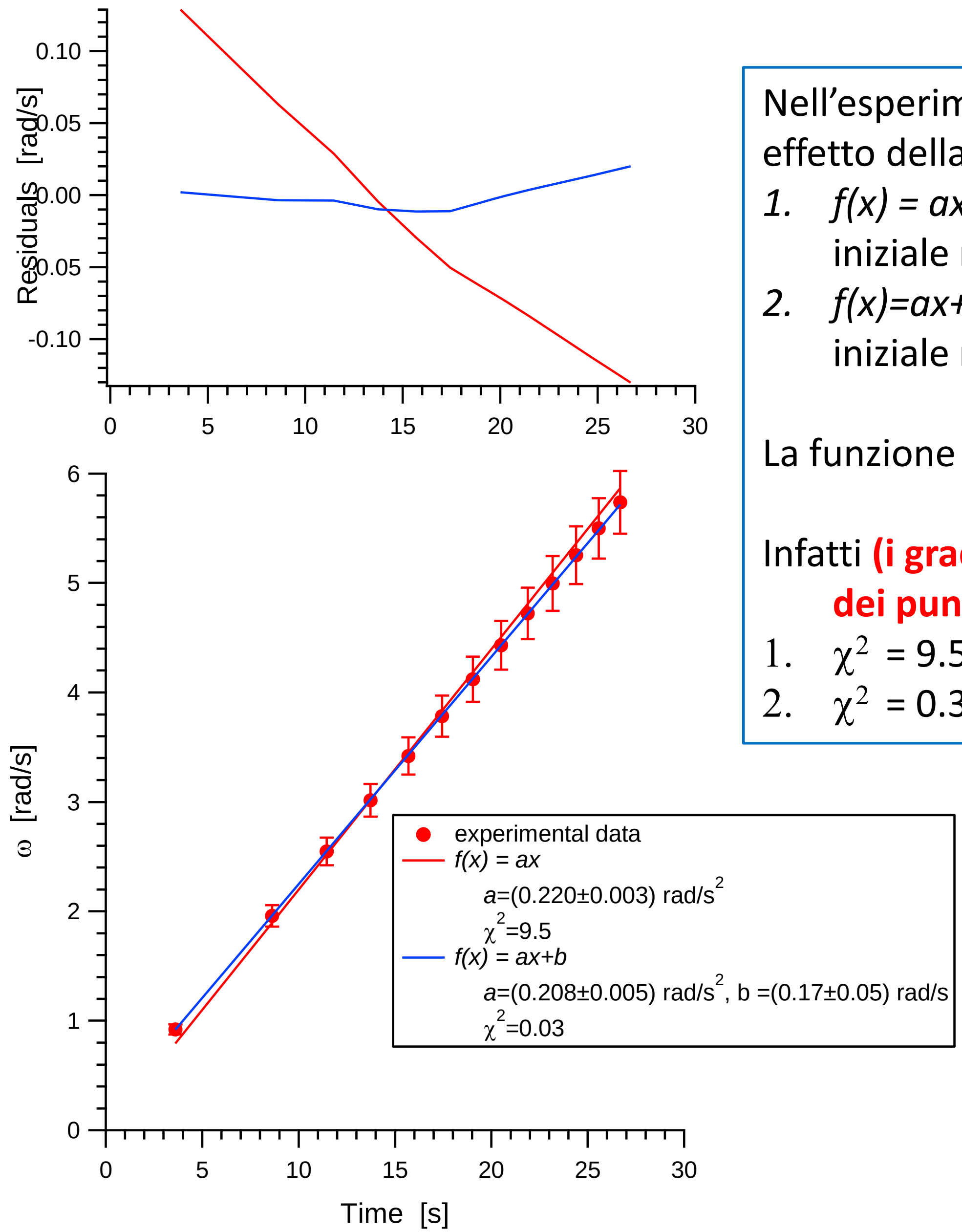
$$\Rightarrow \quad a = \frac{\sum Y_i/\sigma_i^2}{\sum 1/\sigma_i^2}, \qquad (\Delta a)^2 = \frac{1}{\sum 1/\sigma_i^2}.$$

cioè i valori delle Y_i sono pesati con l'inverso dei quadrati dell'errore. Si usa tutte le volte che le misure hanno precisione differente. Questa importante formula è detta *media pesata* delle Y_i.

$$\begin{cases} (\Delta a)^2 = \sum \left(\dfrac{\partial a}{\partial Y_i} \right)^2 \sigma_i^2 = \dfrac{\sum \frac{X_i^2}{\sigma_i^2}}{\sum \frac{1}{\sigma_i^2} \cdot \sum \frac{X_i^2}{\sigma_i^2} - \left(\sum \frac{X_i}{\sigma_i^2} \right)^2} \\ \\ (\Delta b)^2 = \sum \left(\dfrac{\partial b}{\partial Y_i} \right)^2 \sigma_i^2 = \dfrac{\sum \frac{1}{\sigma_i^2}}{\sum \frac{1}{\sigma_i^2} \cdot \sum \frac{X_i^2}{\sigma_i^2} - \left(\sum \frac{X_i}{\sigma_i^2} \right)^2} \end{cases}$$

Spesso si usano algoritmi numerici implementati in programmi di calcolo

Esempi di chi-quadro I



Nell'esperimento "volano" i dati del moto sotto effetto della forza peso possono essere fittati con:

1. *f(x) = ax* (retta passante per origine, velocità iniziale nulla)
2. *f(x)=ax+b* (retta con termine noto di velocità iniziale non nulla)

La funzione 2 "fitta meglio"

Infatti **(i gradi di libertà sono calcolati come numero dei punti sperimentali meno parametri di fit)**:

1. χ^2 = 9.5, *n = 12*
2. χ^2 = 0.3, *n = 11*

Esempi di chi-quadro I: test di verosimiglianza

n	99.5% $\chi^2_{.995}$	99% $\chi^2_{.99}$	97.5% $\chi^2_{.975}$	95% $\chi^2_{.95}$	90% $\chi^2_{.90}$	75% $\chi^2_{.75}$	50% $\chi^2_{.5}$	25% $\chi^2_{.25}$	10% $\chi^2_{.10}$	5% $\chi^2_{.05}$	2.5% $\chi^2_{.025}$	1% $\chi^2_{.01}$	0.5% $\chi^2_{.005}$
1	7.88	6.63	5.02	3.84	2.71	1.32	.455	.102	.0158	.0039	.0010	.0002	.0000
2	10.6	9.21	7.38	5.99	4.61	2.77	1.39	.575	.211	.103	.0506	.0201	.0100
3	12.8	11.3	9.35	7.81	6.25	4.11	2.37	1.21	.584	.352	.216	.155	.072
4	14.9	13.3	11.1	9.49	7.78	5.39	3.36	1.92	1.06	.711	.484	.297	.207
5	16.7	15.1	12.8	11.1	9.24	6.63	4.35	2.67	1.61	1.15	.831	.554	.412
6	18.5	16.8	14.4	12.6	10.6	7.84	5.35	3.45	2.20	1.64	1.24	.872	.676
7	20.3	18.5	16.0	14.1	12.0	9.04	6.35	4.25	2.83	2.17	1.69	1.24	.989
8	22.0	20.1	17.5	15.5	13.4	10.2	7.34	5.07	3.49	2.73	2.18	1.65	1.34
9	23.6	21.7	19.0	16.9	14.7	11.4	8.34	5.90	4.17	3.33	2.70	2.09	1.73
10	25.2	23.2	20.5	$\boxed{18.3}$	16.0	12.5	9.34	6.74	4.87	3.94	3.25	2.56	2.16
11	26.8	24.7	21.9	19.7	17.3	13.7	10.3	7.58	5.58	4.57	3.82	3.05	2.60
12	28.3	26.2	23.3	21.0	18.5	14.8	11.3	8.44	6.30	5.23	4.40	3.57	3.07
13	20.8	27.7	24.7	22.4	19.8	16.0	12.3	9.30	7.04	5.89	5.01	4.11	3.57
14	31.3	29.1	26.1	23.7	21.1	17.1	13.3	10.2	7.79	6.57	5.63	4.66	4.07
15	32.8	30.6	27.5	25.0	22.3	18.2	14.3	11.0	8.55	7.26	6.26	5.23	4.60
16	34.3	32.0	28.8	26.3	23.5	19.4	15.3	11.9	9.31	7.96	6.91	5.81	5.14
17	35.7	33.4	30.2	27.6	24.8	20.5	16.3	12.8	10.1	8.67	7.56	6.41	5.70
18	37.2	34.8	31.5	28.9	26.0	21.6	17.3	13.7	10.9	9.39	8.23	7.01	6.26
19	38.6	36.2	32.9	30.1	27.2	22.7	18.3	14.6	11.7	10.1	8.91	7.63	6.84
20	40.0	37.6	34.2	31.4	28.4	23.8	19.3	15.5	12.4	10.9	9.59	8.26	7.43
21	41.4	38.9	35.5	32.7	29.6	24.9	20.3	16.3	13.2	11.6	10.3	8.90	8.03
22	42.8	40.3	36.8	33.9	30.8	26.0	21.3	17.2	14.0	12.3	11.0	9.54	8.64
23	44.2	41.6	38.1	35.2	32.0	27.1	22.3	18.1	14.8	13.1	11.7	10.2	9.26
24	45.6	43.0	39.4	36.4	33.2	28.2	23.3	19.0	15.7	13.8	12.4	10.9	9.89
25	46.9	44.3	40.6	37.7	34.4	29.3	24.3	19.9	16.5	14.6	13.1	11.5	10.5
26	48.3	45.6	41.9	38.9	35.6	30.4	25.3	20.8	17.3	15.4	13.8	12.2	11.2
27	49.6	47.0	43.2	40.1	36.7	31.5	26.3	21.7	18.1	16.2	14.6	12.9	11.8
28	51.0	48.3	44.5	41.3	37.9	32.6	27.3	22.7	18.9	16.9	15.3	13.6	12.5
29	52.3	49.6	45.7	42.6	39.1	33.7	28.3	23.6	19.8	17.7	16.0	14.3	13.1

Esaminiamo la tabella:
Per *n=12* il valore χ^2 = 9.5, calcolato per il best fit 1, cade tra la colonna del 50% e quella del 25%. Quindi, ripetendo il best fit con altri parametri o altra legge, nel 25-50% dei casi otterrei un c^2 minore, cioè un best fit migliore. Il livello di confidenza è dunque compreso tra 100-50% e 100-25%, cioè è compreso tra 50-75% (niente male, ma niente di che…)

Per *n*=11 il valore χ^2 = 0.3, calcolato per il best fit 2, è minore di tutti i valori riportati. Quindi, ripetendo il best fit con altri parametri o altra legge, in meno dello 0.5% dei casi otterrei un c^2 minore, cioè un best fit migliore. Il livello di confidenza è dunque migliore del 99.5%!! Il best fit 2 è nettamente migliore del best fit 1

Attenzione: valori di χ^2 molto bassi sono sospetti! A parte "manipolazioni" dei dati, essi possono venire da una sovrastima dell'incertezza delle misure (l'incertezza σ_i^2 compare a dividere!)

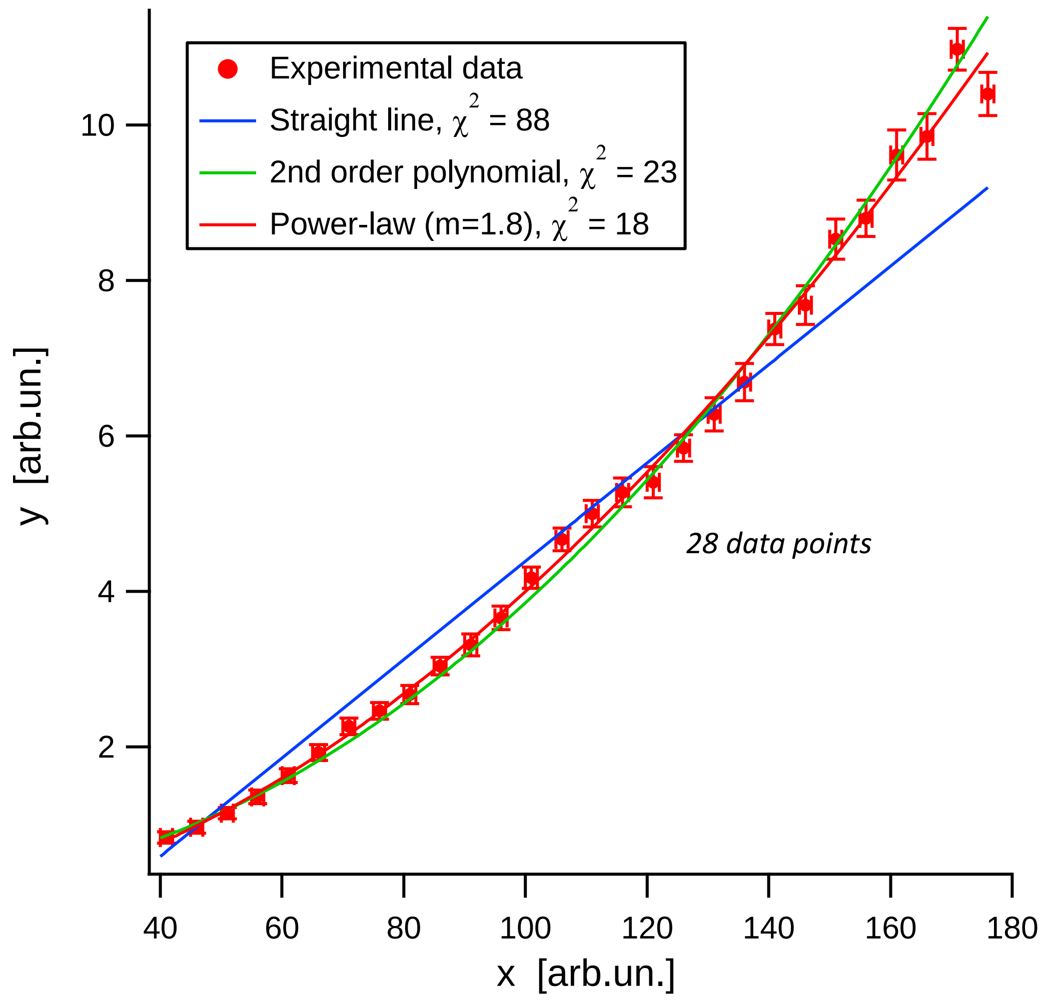
Esempi di chi-quadro II

Simulazione di dati che seguono una legge di potenza: $y = a+bx^m$

1. Fittando con una retta con termine costante, ho $\chi^2 = 88$ *(il numero di gradi di libertà n è pari a numero misure – numero parametri della legge di fit, cioè 28-2=26)*

2. Fittando con polinomio di grado 2 (3 parametri, cioè n = 25) ho $\chi^2 = 23$

3. Fittando con legge di potenza (3 parametri, cioè n = 25 anche in questo caso) ho ho $\chi^2 = 18$



Esempi di chi-quadro II: test di verosimiglianza

n	99.5% $\chi^2_{.995}$	99% $\chi^2_{.99}$	97.5% $\chi^2_{.975}$	95% $\chi^2_{.95}$	90% $\chi^2_{.90}$	75% $\chi^2_{.75}$	50% $\chi^2_{.5}$	25% $\chi^2_{.25}$	10% $\chi^2_{.10}$	5% $\chi^2_{.05}$	2.5% $\chi^2_{.025}$	1% $\chi^2_{.01}$	0.5% $\chi^2_{.005}$
1	7.88	6.63	5.02	3.84	2.71	1.32	.455	.102	.0158	.0039	.0010	.0002	.0000
2	10.6	9.21	7.38	5.99	4.61	2.77	1.39	.575	.211	.103	.0506	.0201	.0100
3	12.8	11.3	9.35	7.81	6.25	4.11	2.37	1.21	.584	.352	.216	.155	.072
4	14.9	13.3	11.1	9.49	7.78	5.39	3.36	1.92	1.06	.711	.484	.297	.207
5	16.7	15.1	12.8	11.1	9.24	6.63	4.35	2.67	1.61	1.15	.831	.554	.412
6	18.5	16.8	14.4	12.6	10.6	7.84	5.35	3.45	2.20	1.64	1.24	.872	.676
7	20.3	18.5	16.0	14.1	12.0	9.04	6.35	4.25	2.83	2.17	1.69	1.24	.989
8	22.0	20.1	17.5	15.5	13.4	10.2	7.34	5.07	3.49	2.73	2.18	1.65	1.34
9	23.6	21.7	19.0	16.9	14.7	11.4	8.34	5.90	4.17	3.33	2.70	2.09	1.73
10	25.2	23.2	20.5	$\boxed{18.3}$	16.0	12.5	9.34	6.74	4.87	3.94	3.25	2.56	2.16
11	26.8	24.7	21.9	19.7	17.3	13.7	10.3	7.58	5.58	4.57	3.82	3.05	2.60
12	28.3	26.2	23.3	21.0	18.5	14.8	11.3	8.44	6.30	5.23	4.40	3.57	3.07
13	20.8	27.7	24.7	22.4	19.8	16.0	12.3	9.30	7.04	5.89	5.01	4.11	3.57
14	31.3	29.1	26.1	23.7	21.1	17.1	13.3	10.2	7.79	6.57	5.63	4.66	4.07
15	32.8	30.6	27.5	25.0	22.3	18.2	14.3	11.0	8.55	7.26	6.26	5.23	4.60
16	34.3	32.0	28.8	26.3	23.5	19.4	15.3	11.9	9.31	7.96	6.91	5.81	5.14
17	35.7	33.4	30.2	27.6	24.8	20.5	16.3	12.8	10.1	8.67	7.56	6.41	5.70
18	37.2	34.8	31.5	28.9	26.0	21.6	17.3	13.7	10.9	9.39	8.23	7.01	6.26
19	38.6	36.2	32.9	30.1	27.2	22.7	18.3	14.6	11.7	10.1	8.91	7.63	6.84
20	40.0	37.6	34.2	31.4	28.4	23.8	19.3	15.5	12.4	10.9	9.59	8.26	7.43
21	41.4	38.9	35.5	32.7	29.6	24.9	20.3	16.3	13.2	11.6	10.3	8.90	8.03
22	42.8	40.3	36.8	33.9	30.8	26.0	21.3	17.2	14.0	12.3	11.0	9.54	8.64
23	44.2	41.6	38.1	35.2	32.0	27.1	22.3	18.1	14.8	13.1	11.7	10.2	9.26
24	45.6	43.0	39.4	36.4	33.2	28.2	23.3	19.0	15.7	13.8	12.4	10.9	9.89
25	46.9	44.3	40.6	37.7	34.4	29.3	24.3	19.9	16.5	14.6	13.1	11.5	10.5
26	48.3	45.6	41.9	38.9	35.6	30.4	25.3	20.8	17.3	15.4	13.8	12.2	11.2
27	49.6	47.0	43.2	40.1	36.7	31.5	26.3	21.7	18.1	16.2	14.6	12.9	11.8
28	51.0	48.3	44.5	41.3	37.9	32.6	27.3	22.7	18.9	16.9	15.3	13.6	12.5
29	52.3	49.6	45.7	42.6	39.1	33.7	28.3	23.6	19.8	17.7	16.0	14.3	13.1

Esaminiamo la tabella:

1. Per *n=26* il valore χ^2 = 88 calcolato per il best fit 1 (retta) è maggiore del massimo numero riportato nella linea, per cui la confidenza del best fit è peggiore di 100-99.5% = 0.5% (il best fit è pessimo!)

2. Per *n*=25 il valore χ^2 = 23, calcolato per il best fit 2 (polinomio), cade tra la colonna del 25% e quella del 50%, per cui la confidenza del best fit è 50-75%

3. Per *n*=25 il valore χ^2 = 18, calcolato per il best fit 3 (legge di potenza), cade tra la colonna del 10% e quella del 25%, per cui la confidenza del best fit è 75-90%: la legge di potenza è quella che descrive meglio (con maggior livello di confidenza) i risultati sperimentali!

Il confronto tra valori di χ^2 calcolati per varie scelte della funzione di best fit permette di determinare la legge che descrive meglio (con maggior livello di confidenza) i dati sperimentali

Esempi di chi-quadro III

Decadimento della fotoluminescenza di un cromoforo in matrice polimerica (PMMA)

La legge a doppio esponenziale descrive meglio il comportamento (infatti ha una spiegazione fisica dovuta alla compresenza di due canali di decadimento)

<ins>Nota</ins>: con un alto numero di punti sperimentali, il numero di gradi di libertà *n* diventa troppo alto per essere trovato nelle tabelle e il valore del χ^2 risulta anche molto elevato.
Spesso si definisce e si usa il **chi-quadro ridotto** $\chi^2_R = \chi^2/n$ con $n = n_{DATI} - n_{PARAMETRI}$: il best fit è "ragionevole" se $\chi^2_R \leq 1$, come nel caso del best fit al doppio esponenziale