

Parte A

Richiami di analisi degli errori

- 1) Supponiamo di avere misurato sperimentalmente una qualche grandezza x molte volte, dopo aver individuato e ridotto a zero il più possibile le fonti di errore “sistematico” (ovvero le incertezze sperimentali che non possono essere rivelate ripetendo la misura; quelle che possono essere rivelate in questo modo sono invece chiamate incertezze sperimentali “casuali”). Ripetendo la misura N volte, si può dimostrare che la miglior stima di x è la media delle misure:

$$x_{best} = \bar{x} = \frac{1}{N} \sum_{i=1}^N x_i .$$

La deviazione standard è la stima dell’incertezza media delle singole misure ... , ed è definita come:

$$s_x = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2} .$$

La deviazione standard della media caratterizza, invece, l’incertezza media di ... , e si calcola nel seguente modo:

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{N}} .$$

- 2) Supponiamo di disporre di n misure di una grandezza x , denominate x_i ($i = 1, \dots, n$), ognuna caratterizzata da una incertezza σ_i ($i = 1, \dots, n$). La miglior stima per la grandezza x è data dalla media pesata:

$$x_{wav} = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i} ,$$

dove $w_i = 1/s_i^2$.

L’incertezza che affligge la misura è data da:

$$\sigma_{wav} = \frac{1}{\sqrt{\sum_{i=1}^n w_i}} .$$

- 3) Se le misure sono affette da piccole sorgenti di errori casuali e trascurabili errori sistematici, allora i valori misurati saranno distribuiti su una curva a campana detta “funzione di Gauss” (centrata in X , con larghezza ...):

$$f_{x,\sigma} = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-X)^2/2\sigma^2}.$$

- 4) Per verificare che i valori ricavati \bar{x} e σ_x siano valori ragionevoli rispetto al valore conosciuto della grandezza X si può calcolarne il livello di confidenza. Definiamo innanzitutto il numero di deviazioni standard per cui \bar{x} differisce da X :

$$t = |x_{best} - x_{vero}| / s_x.$$

La probabilità, date le ipotesi fatte, di ottenere un risultato che differisca da X per t o più deviazioni standard è:

$$P(\text{al di fuori di } t) = 1 - P(\text{entro } t).$$

Dove (vedere Allegati pag. I, C300):

$$P(\text{entro } t) = \frac{1}{\sqrt{2\pi}} \int_{-t}^t e^{-z^2/2} dz = \text{erf}\left(\frac{t}{\sqrt{2}}\right).$$

A questo punto $P(\text{al di fuori di } t)$ prende il nome di livello di confidenza (CL).

- 5) Una stima dell'accordo tra un insieme n di punti sperimentali (x_k, y_k) e una funzione $f(x)$ si può ottenere misurando χ^2 , definito come:

$$\chi^2 = \sum_{k=1}^n \frac{(y_k - f(x_k))^2}{s^2(y_k)}.$$

Se si confronta una distribuzione di eventi con una distribuzione di probabilità non normalizzata, ad esempio una gaussiana, la formula del χ^2 diventa:

$$\chi^2 = \sum_{k=1}^n \frac{(O_k - E_k)^2}{E_k},$$

dove è stato diviso l'intervallo dei possibili valori di x in n intervalli, per cui O_k rappresenta il numero di osservazioni che cadono nell'intervallo k -esimo e E_k il numero atteso delle misure dell'intervallo k -esimo. χ^2 è un indicatore ragionevole dell'accordo tra la distribuzione osservata e quella attesa. Se $\chi^2 = 0$ l'accordo è perfetto, cosa molto improbabile. In generale ci si aspetta che ogni termine della somma sia dell'ordine di 1. Il chi quadrato ridotto si definisce, invece, nel modo seguente:

,

essendo ν il numero di gradi di libertà, e N dove N è il numero di intervalli significativi e p è il numero di parametri, v il numero di vincoli.

- 6) Supponiamo ora di aver misurato una o più grandezze x, y, \dots e calcolato gli errori corrispondenti $\Delta x, \Delta y, \dots$, e di dover calcolare la funzione $z = f(x, y, \dots)$. La miglior stima del valore della funzione sarà:

Per stimare l'incertezza di questo risultato si applica la propagazione degli errori delle misure x , y , quindi:

$$\sigma_z = \sqrt{\left(\frac{\partial f}{\partial x} \sigma_x\right)^2 + \left(\frac{\partial f}{\partial y} \sigma_y\right)^2 + \dots + 2\sigma_{xy} \frac{\partial f}{\partial x} \frac{\partial f}{\partial y}}$$

In particolare, nel caso in cui $\sigma_{xy} = 0$ e se gli errori su x ed y sono **indipendenti e casuali**, il calcolo della deviazione standard si riduce a:

$$\frac{\sigma_z}{z} = \sqrt{\left(a \frac{\sigma_x}{x}\right)^2 + \left(b \frac{\sigma_y}{y}\right)^2}$$

- 7) Supponiamo di avere una serie di misure (y_i) $i=1\dots N$ con gli x_i tutti esatti e gli y_i tutti ugualmente incerti, quindi assumiamo che le misure degli y_i siano governate da distribuzioni normali con larghezza σ_0 . Se le variabili x_i e y_i sono legate da una relazione **lineare** della forma (con A e B come incognite):

si può trovare la retta che meglio approssima la serie di misure, minimizzando la funzione:

$$\chi^2 = \sum_{i=1}^N \frac{(y_i - A - Bx_i)^2}{\sigma_0^2}$$

Differenziando rispetto ad A e B ed uguagliando a zero si ottiene:

$$A = \frac{\sum_{i=1}^N x_i^2 \sum_{i=1}^N y_i - \sum_{i=1}^N x_i \sum_{i=1}^N x_i y_i}{\Delta}$$

$$B = \frac{N \sum_{i=1}^N x_i y_i - \sum_{i=1}^N x_i \sum_{i=1}^N y_i}{\Delta}$$

dove:

$$\Delta = N \sum_{i=1}^N x_i^2 - \left(\sum_{i=1}^N x_i\right)^2$$

Le incertezze sui parametri calcolati sono le seguenti: $\sigma_A = \sigma_y \sqrt{\frac{\sum x_i^2}{\Delta}}$ e

Essendo: $\sigma_y = \sqrt{\frac{1}{N_{DF}} \sum_{i=1}^N (y_i - A - Bx_i)^2}$, con N_{DF} il numero di gradi di libertà (in questo caso $N-2$).

La covarianza $s(A, B)$ e la relativa correlazione $r(A, B) = \frac{s(A, B)}{s(A) s(B)}$ sono facilmente calcolabili come:

$$s(A, B) = \frac{-s_y^2 \sum x_i}{\Delta},$$

$$r(A, B) = \frac{-\sum x_i}{\sqrt{N \sum x_i^2}}.$$

- 8) Supponiamo ora, invece, che per ogni x_i il corrispondente valor vero di y_i è dato dalla **polinomiale** della variabile x (con A e C come incognite):

Assumendo ancora che le misure degli x_i siano governate da distribuzioni normali con larghezza σ_0 , si può trovare la miglior stima delle costanti A e C minimizzando la funzione:

$$\chi^2 = \sum_{i=1}^N \frac{(y_i - A - Cx_i^2)^2}{s_0^2}.$$

Differenziando rispetto ad A e C ed uguagliando a zero si ottiene:

$$A = \frac{\sum y_i \sum x_i^4 - (\sum x_i^2 y_i) \sum x_i^2}{D_p},$$

$$C = \frac{N \sum x_i^2 y_i - (\sum y_i) \sum x_i^2}{D_p},$$

Da qui si ricavano le migliori stime per A e C .

Attraverso la propagazione degli errori si possono ricavare le rispettive deviazioni standard:

$$\sigma_A = \sigma_y \sqrt{\frac{\sum x_i^4}{\Delta_p}} \quad \text{e} \quad \sigma_C = \sigma_y \sqrt{\frac{N}{\Delta_p}},$$

essendo:

$$s_y = \sqrt{\frac{1}{N_{DF}} \sum_{i=1}^N (y_i - A - Cx_i^2)^2} \quad (\text{dove } N_{DF} = N - 2).$$

L'equazione $y = A + Cx^2$ viene così denominata adattamento polinomiale dei minimi quadrati, o regressione lineare delle misure date. Analogamente al caso della retta, la covarianza

e la relativa correlazione $r(A, B) = \frac{s(A, B)}{s(A) s(B)}$ sono facilmente calcolabili come:

$$s(A, B) = \frac{-s_y^2 \sum x_i^2}{\Delta},$$

$$r(A, B) = \frac{-\sum x_i^2}{\sqrt{N \sum x_i^4}}.$$