

Propagazione delle incertezze



Questo testo è distribuito con Licenza Creative Commons Attribuzione
Condividi allo stesso modo 4.0 Internazionale

Luca Mari, versione 18.5.16

Contenuti

Introduzione.....	1
Rappresentazione di grandezze in condizioni di incertezza.....	2
Inceteeze relative.....	2
Dall'inceteeza tipo all'inceteeza estesa.....	3
Propagazione dei valori misurati.....	3
Propagazione delle inceteeze: caso a un singolo argomento.....	4
Alcuni esempi.....	4
Propagazione delle inceteeze: caso a più argomenti.....	5
Alcuni casi semplici di propagazione delle inceteeze.....	5
Un esempio.....	5
Propagazione delle distribuzioni.....	6
Un esempio.....	6

I principali concetti introdotti in questo capitolo

deviazione standard campionaria.....	2
fattore di copertura.....	3
inceteeza estesa.....	3
inceteeza relativa.....	2
inceteeza tipo.....	2
inceteeza tipo combinata.....	4
intervallo di confidenza.....	3
legge di propagazione delle inceteeze.....	5
livello di confidenza.....	3
metodo Monte Carlo.....	6
valore misurato.....	2

Introduzione

Una grandezza X è stata valutata (per esempio misurandola) n volte, in condizioni di ripetibilità e di indipendenza statistica e come risultato di tali valutazioni è stato ottenuto il campione $\langle x_1, \dots, x_n \rangle$. Il fatto che gli elementi del campione siano diversi tra loro e che il valore x_i non sia precisamente prevedibile a partire dai valori x_1, \dots, x_{i-1} suggerisce di formalizzare la grandezza X come una variabile casuale, ottenuta da una distribuzione di probabilità che in generale è non nota (come d'abitudine in statistica, per semplicità adotteremo lo stesso simbolo, X , per indicare sia la grandezza sia la variabile casuale ad essa associata, e alterneremo liberamente il riferimento a X come grandezza e come variabile).

Nonostante la condizione statistico-probabilistica, è abituale rappresentare la grandezza mediante un singolo elemento x del suo insieme supporto A , $x \in A$, dunque:

$$X = x$$

come quando, per esempio, si scrive:

$$\text{lunghezza(oggetto)} = 1,23 \text{ m}$$

Un'importante ragione di questa rappresentazione è che le leggi (geometriche, fisiche, economiche, ...) che descrivono le relazioni tra grandezze sono generalmente scritte supponendo che le variabili abbiano come valori singoli numeri (con eventuali unità di misura, una precisazione non rilevante qui). Consideriamo il caso, sufficientemente generale, di una grandezza Y , al cui valore siamo interessati ma che non sappiamo valutare direttamente, ma solo attraverso una funzione f :

$$Y = f(X_1, \dots, X_k)$$

in cui le X_i sono variabili casuali. Anche Y è allora in generale una variabile casuale, le cui caratteristiche dovrebbero essere derivate dalle caratteristiche delle variabili casuali X_i "propagate" attraverso la funzione f . Poiché l'informazione sulle grandezze X_i è incerta, un problema di questo genere si chiama di *propagazione*

delle incertezze (il termine tradizionale è “propagazione degli errori”: non discuteremo qui l’eventuale differenza).

La più semplice situazione di propagazione delle incertezze è nel caso $Y = f(X)$. Potrebbe trattarsi per esempio del calcolo della superficie di un oggetto quadrato a partire da un campione di valori per la sua lunghezza, $Y = f(X) = X^2$. Un poco più complessa è per esempio la situazione in cui le variabili indipendenti da considerare sono due, come nel caso in cui si voglia calcolare la superficie di un oggetto rettangolare a partire da campioni di valori per le sue due dimensioni, $Y = f(X_1, X_2) = X_1 \times X_2$.

Ricordiamo prima di tutto come si può rappresentare l’informazione su una grandezza il cui valore è incerto. Vedremo poi come propagare le incertezze attraverso funzioni a un argomento e poi a $K > 1$ argomenti.

Rappresentazione di grandezze in condizioni di incertezza

Come sappiamo, in condizioni di incertezza l’informazione su una grandezza può essere sintetizzata in particolare mediante una statistica di posizione e una statistica di dispersione.

Abbiamo già discusso circa i limiti di applicabilità della media campionaria. Quando essa risulta applicabile, e si ipotizza che la distribuzione da cui il campione è stato ottenuto sia “sufficientemente regolare”, il valore medio della variabile casuale è stimato mediante il valore medio campionario:

$$m_x = \frac{1}{n} \sum_{i=1}^n x_i$$

che porta un’informazione di posizione sulla distribuzione sui valori possibili della variabile casuale, e viene quindi usato come valore rappresentativo per la grandezza, chiamato in breve *valore misurato*. In assenza di altra informazione, è dunque il valore misurato che viene applicato alla funzione f per calcolare un valore y per la grandezza Y .

L’informazione sulla grandezza X dovrebbe però essere rappresentata non solo mediante il valore misurato ma anche attraverso una statistica di dispersione, che specifichi l’incertezza di tale valore, che è evidentemente dovuta alla dispersione dei valori della variabile casuale intorno al suo valor medio. Ricordiamo che la varianza campionaria è:

$$s_x^2 = \frac{\sum_{i=1}^n (x_i - m_x)^2}{n-1}$$

La sua radice quadrata, cioè la *deviazione standard campionaria*, è una statistica di incertezza relativa alla variabilità della grandezza. D’altra parte, ci interessa valutare l’incertezza non direttamente della variabile casuale ma del suo valor medio, cioè del valore misurato della grandezza, che è a sua volta una variabile casuale: eseguendo più campionamenti dalla stessa distribuzione, si otterrebbero infatti valori medi campionari diversi. Dunque la varianza del valore misurato è:

$$s_{m_x}^2 = \frac{s_x^2}{n} = \frac{\sum_{i=1}^n (x_i - m_x)^2}{n(n-1)}$$

la cui radice quadrata:

$$u_x = \frac{s_x}{\sqrt{n}} = \sqrt{\frac{\sum_{i=1}^n (x_i - m_x)^2}{n(n-1)}}$$

è una deviazione standard, dimensionalmente omogenea al misurando, chiamata *incertezza tipo* (in inglese *standard uncertainty*, da cui il simbolo u_x) del valore misurato.

In condizioni di incertezza l’informazione su una grandezza può essere dunque sintetizzata dalla coppia $\langle m_x, u_x \rangle$, valore misurato e incertezza tipo.

Incerteze relative

L’incertezza tipo u_x fornisce un’informazione sulla stabilità del valore misurato, ma mantiene una certa ambiguità: per esempio, una stessa incertezza di 1 Ω sul valore della resistenza di un resistore da 10 Ω e di uno da 10 k Ω influisce in modo ben diverso sulla qualità dei due valori misurati. Per ridurre questa ambiguità, si usa riportare l’*incertezza relativa*, definita come:

$$u_{xrel} = \frac{u_x}{|m_x|}$$

dunque mettendo in relazione il valore dell'incertezza tipo con il valore misurato. Poiché generalmente u_{xrel} è un numero piccolo, lo si indica specificandone solo la prima cifra significativa e la potenza negativa di dieci per cui è moltiplicato.

Per esempio, se (tralasciando l'indicazione dell'unità di misura) $m_x = 40,26$ e $u_x = 0,03$ allora l'incertezza relativa è $u_{xrel} = 0,03/40,26 \approx 0,0007$ cioè $u_{xrel} = 7 \times 10^{-4}$.

Un altro metodo per riportare le incertezze relative è di moltiplicarle per 10^3 , e quindi di comunicarle in "per mille" (nell'esempio, 0,7 ‰), oppure anche di moltiplicarle per 10^6 , e quindi di comunicarle in "parti per milione" (nell'esempio, 700 ppm).

Dall'incertezza tipo all'incertezza estesa

E' spesso utile esprimere le misure come intervalli di indifferenza, tali cioè che ogni elemento dell'intervallo possa essere scelto "con un'alta probabilità" come valore per la grandezza. Si può passare dalla rappresentazione mediante incertezze tipo a quella per intervalli moltiplicando l'incertezza tipo u_x per un **fattore di copertura** (in inglese *coverage factor*) k , ottenendo così l'**incertezza estesa** (in inglese *expanded uncertainty*) $U_x = k u_x$, tale dunque che $[m_x - U_x, m_x + U_x]$ sia l'intervallo di indifferenza cercato.

Assumendo che sia nota la distribuzione di probabilità di cui m_x è il valor medio e u_x è l'incertezza tipo, può essere ricavata la relazione tra l'intervallo così ottenuto, chiamato **intervallo di confidenza**, e la probabilità dell'intervallo stesso, chiamata in tal caso **livello di confidenza**. C'è evidentemente una relazione di monotonicità diretta tra fattore di copertura e livello di confidenza, e quindi tra ampiezza dell'intervallo di confidenza e livello di confidenza, relazione che può essere espressa analiticamente se si conosce la distribuzione di probabilità sottostante. Nel caso di distribuzione gaussiana, in particolare:

fattore di copertura	livello di confidenza
1	0,68
1,645	0,9
1,960	0,95
2	0,9545
2,576	0,99
3	0,9973

Propagazione dei valori misurati

Data una relazione funzionale $Y = f(X_1, \dots, X_K)$ se le X_i sono variabili casuali, ognuna con valor medio m_i e incertezza tipo u_i , evidentemente anche la Y sarà una variabile casuale, dipendente dalle X_i attraverso la funzione f . Dalle K coppie $\langle m_i, u_i \rangle$ e conoscendo l'espressione analitica di f si pone dunque il problema di come calcolare la coppia $\langle m_Y, u_Y \rangle$ per la variabile casuale Y , cioè propriamente la misura per la grandezza Y .

Per quanto riguarda m_Y , la scelta abituale è:

$$m_Y = f(m_1, \dots, m_K)$$

che però è non problematica solo nel caso in cui f sia lineare. Solo in questo caso, infatti indicando con $E[X]$ il valore atteso della variabile casuale X , vale che $f(E[X_1], \dots, E[X_K]) = E[f(X_1, \dots, X_K)]$, cioè non è rilevante se prima si calcolano i valori medi delle X_i e poi ad essi si applica f , come appena supposto, oppure si applica f agli n elementi dei K campioni e poi si calcola il valor medio di Y sul risultato, un'opzione di principio altrettanto ammissibile.

Vediamo un semplice esempio, nel caso di una funzione non lineare mono-argomentale, cioè con $K = 1$, come:

$$Y = f(X) = X^2$$

con cui potremmo valutare l'area di una superficie quadrata a partire dalla misura del suo lato. Supponiamo che X sia stato valutato tre volte, ottenendo $x_1=1.00$, $x_2=1.10$, $x_3=1.30$ (i numeri sono evidentemente artificiali e stiamo tralasciando l'indicazione dell'unità di misura). Le due procedure forniscono allora:

$$m_Y = f(E[X]) \approx (m_X)^2 = ((x_1+x_2+x_3)/3)^2 = 1.28$$

e:

$$m_Y = E[f(X)] = E[X^2] \approx (x_1^2 + x_2^2 + x_3^2)/3 = 1.30$$

Il fatto che i risultati siano diversi mostra che la scelta della procedura non è in generale indifferente. D'altra parte, il criterio di scelta è complesso: potremo quindi continuare ad assumere che i valori misurati si propagano secondo la procedura tradizionale.

Propagazione delle incertezze: caso a un singolo argomento

Il calcolo di u_Y è più complesso. Cominciamo a considerare il caso semplice delle funzioni a un solo argomento, $K = 1$, in cui dunque $Y = f(X)$. La presenza di un'incertezza su X fa sì che i valori x_i siano in generale non coincidenti con il valor medio m_X e quindi che a ogni valore sia associato uno scarto $x_i - m_X$.

Assumendo che gli scarti $x_i - m_X$ siano sufficientemente piccoli e che il comportamento di f intorno al punto m_X sia sufficientemente lineare, si può sviluppare f in serie di Taylor intorno a m_X arrestandosi al termine del primo ordine:

$$y = f(x) = f(m_X) + \frac{df}{dx}(x - m_X)$$

avendo indicato con df/dx la derivata df/dX della funzione f calcolata nel punto m_X . D'altra parte, poiché abbiamo assunto che $f(m_X) = m_Y$:

$$y - m_Y = \frac{df}{dx}(x - m_X)$$

relazione che stabilisce la dipendenza dei (piccoli) scarti di y intorno a m_Y dai (piccoli) scarti di x intorno a m_X (ricordiamo che lo sviluppo in serie di Taylor consente di calcolare il valore $f(x + \Delta x)$ a partire dal valore di $f(x)$ e delle derivate di f calcolate in x):

$$f(x + \Delta x) = f(x) + \frac{df}{dx} \Delta x + \frac{d^2 f}{dx^2} \frac{\Delta x^2}{2!} + \frac{d^3 f}{dx^3} \frac{\Delta x^3}{3!} + \dots,$$

dove appunto $\frac{d^i f}{dx^i}$ è la derivata i -esima della funzione f calcolata in x .

Tali scarti possono essere immediatamente trasformati in incertezze tipo, ottenendo:

$$u_Y = \left| \frac{df}{dx} \right| u_X$$

(espressione che tra l'altro giustifica l'idea che questo non sia altro che un problema di cambio di variabili). Il valore u_Y , chiamato *incertezza tipo combinata* della grandezza Y dipende dall'incertezza tipo u_X della grandezza X attraverso il termine $|df/dx|$, che rappresenta dunque un coefficiente di sensibilità della variazione di X mediante f nell'intorno del punto m_X .

Per esempio, se $Y = f(X) = X^2$ allora $\left| \frac{df}{dx} \right| = 2|m_X|$ e quindi $u_Y = 2|m_X|u_X$ (naturalmente queste equazioni sono dimensionalmente corrette, nel senso che $[u_Y] = [Y]$; infatti assumendo che $[X] = [u_X]$ e considerando che $[dX] = d[X]$, $\frac{d[f]}{d[X]}[u_X] = [f] = [Y]$).

Alcuni esempi

(Si noti: i valori numerici riportati in questi esempi non sono realistici e le grandezze sono indicate omettendo l'unità di misura)

- $Y = X + k$

Poiché $df/dX = 1$, allora $u_Y = u_X$

→ L'incertezza non si modifica per traslazione.

- $Y = kX$

Poiché $df/dX = k$, allora $u_Y = k u_X$

→ Se l'incertezza assoluta di Y è k volte superiore a quella di X , le incertezze relative, $u_{Xrel} = u_X/|m_X|$ e $u_{Yrel} = u_Y/|m_Y| = k u_X/|k m_X|$, sono uguali.

- $Y = X^2$

Poiché $df/dX = 2X$, allora $u_Y = 2|m_X|u_X$

→ Dunque in questo caso l'incertezza dipende dal valore della grandezza di ingresso. Per esempio, se $m_X = 10$ e $u_X = 2$, allora $m_Y = m_X^2 = 100$ e $u_Y = 2m_X u_X = 2 \times 10 \times 2 = 40$. In termini di incertezze relative: $u_{Xrel} = 2/10 = 0,2$ mentre $u_{Yrel} = 40/100 = 0,4$: la relazione quadratica tra X e Y peggiora anche l'incertezza relativa.

- $Y = \sin(X)$
Poiché $df/dX = \cos(X)$, $u_Y = |\cos(m_X)|u_X$
→ Dunque anche in questo caso l'incertezza dipende dal valore della grandezza di ingresso. Per esempio, se $m_X = 0$ e $u_X = 0,1$, allora $m_Y = \sin(m_X) = 0$ e $u_Y = \cos(m_X)u_X = 1 \times 0,1 = 0,1$. Nell'intorno di $m_X = 0$, la funzione $Y = \sin(X)$ è approssimata come $Y \approx X$: l'incertezza tipo di Y è uguale all'incertezza tipo di X .

Propagazione delle incertezze: caso a più argomenti

Possiamo ora generalizzare il discorso precedente al caso in cui la funzione f ha $K > 1$ argomenti, ancora sviluppando f in serie di Taylor intorno al punto m_1, \dots, m_K e arrendendosi ai termini del primo ordine. Nell'ipotesi che le covarianze tra le grandezze di ingresso X_i siano trascurabili, dopo alcuni passaggi analoghi a quelli compiuti nel caso mono-argomentale si ottiene:

$$(y - m_Y)^2 = \sum_{i=1}^K \left(\frac{\partial f}{\partial x_i} \right)^2 (x_i - m_i)^2$$

e quindi:

$$u_Y = \sqrt{\sum_{i=1}^K \left(\frac{\partial f}{\partial x_i} \right)^2 u_i^2}$$

espressione che consente di calcolare l'incertezza tipo di Y in funzione delle incertezze tipo delle X_i (nota: anche in questo caso, i coefficienti di sensibilità $\partial f / \partial x_i$ si intendono calcolati nel valor medio m_1, \dots, m_K).

Più in generale, considerando le covarianze $u_{i,j}$ (e indicando con $u_{i,i} = u_i^2$ la varianza di X_i):

$$u_Y = \sqrt{\sum_{i=1}^K \sum_{j=1}^K \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u_{i,j}}$$

la versione generale della cosiddetta *legge di propagazione delle incertezze*.

Alcuni casi semplici di propagazione delle incertezze

... cioè di applicazione a casi particolari della legge: $u_Y = \sqrt{\sum_{i=1}^K \left(\frac{\partial f}{\partial x_i} \right)^2 u_i^2}$ (dunque nell'ipotesi di covarianze nulle):

- se $Y = X_1 + X_2$ allora $u_Y^2 = u_1^2 + u_2^2$;
- se $Y = X_1 X_2$ allora $u_Y^2 = m_2^2 u_1^2 + m_1^2 u_2^2$ o anche, più espressivamente $u_{Yrel}^2 = u_{1rel}^2 + u_{2rel}^2$;
- se $Y = X^k$ allora $u_Y^2 = (k m_X^{k-1})^2 u_X^2$ o anche, più espressivamente $u_{Yrel} = |k| u_{Xrel}$.

Nell'esempio semplice $Y = X_1 + X_2$, applichiamo anche la formula per covarianze non nulle (ricordando naturalmente che $u_{i,j} = u_{j,i}$):

$$u_Y^2 = \sum_{i=1}^K \sum_{j=1}^K \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u_{i,j} = \frac{\partial f}{\partial x_1} \left[\frac{\partial f}{\partial x_1} u_1^2 + \frac{\partial f}{\partial x_2} u_{1,2} \right] + \frac{\partial f}{\partial x_2} \left[\frac{\partial f}{\partial x_1} u_{1,2} + \frac{\partial f}{\partial x_2} u_2^2 \right]$$

e poiché in questo caso $\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = 1$ si ha finalmente che $u_Y^2 = u_1^2 + u_2^2 + 2u_{1,2}$.

Un esempio

Si vuole valutare la potenza P dissipata ai capi di un resistore a cui è applicata una tensione, ma non si dispone di un sistema per misurare direttamente P . Si ricorda, d'altra parte, che $P = V^2/R$, dove V è la tensione applicata al resistore e R è la sua resistenza. Disponendo di un sistema di misura che consente di valutare V e R , si potrà allora calcolare, cioè "misurare indirettamente", P .

Supponiamo che la misurazione di V e R sia stata ripetuta. Dai corrispondenti campioni si calcolano i valori medi m_V e m_R e le incertezze tipo u_V e u_R , così che l'informazione per V e R è rappresentata dalle coppie $\langle m_V, u_V \rangle$ e $\langle m_R, u_R \rangle$ rispettivamente. Il problema è dunque di calcolare $\langle m_P, u_P \rangle$ per la potenza dissipata.

Per quanto riguarda m_P , semplicemente $m_P = m_V^2 / m_R$.

Per calcolare u_P si utilizza la legge di propagazione delle incertezze, nell'ipotesi di covarianze nulle:

$$u_P^2 = \left(\frac{\partial f}{\partial V} \right)^2 u_V^2 + \left(\frac{\partial f}{\partial R} \right)^2 u_R^2 = \left(2 \frac{m_V}{m_R} \right)^2 u_V^2 + \left(\frac{-m_V^2}{m_R^2} \right)^2 u_R^2$$

Il passo successivo potrebbe essere di riconoscere che il “modello della misurazione” prevede la dipendenza di R dalla temperatura t , $R(t)=R_0[1+\alpha(t-t_0)]$, dove R_0 è la resistenza del resistore alla temperatura t_0 .

Propagazione delle distribuzioni

La procedura indicata per la propagazione delle incertezze può non essere sempre applicabile, in particolare perché si ritiene troppo complessa, o non appropriata (per esempio perché f è sensibilmente non lineare intorno al valor medio delle grandezze di ingresso, e quindi un'approssimazione solo al primo ordine nello sviluppo in serie di Taylor è critica ma si giudica troppo complesso un trattamento analitico dell'approssimazione a ordini superiori), o non fattibile (per esempio perché f non è derivabile intorno al valor medio delle grandezze di ingresso). Si può allora adottare una strategia alternativa per valutare l'incertezza per la grandezza Y , mediante una tecnica di tipo non analitico ma numerico, basata sul *metodo Monte Carlo*.

Supponiamo che siano noti la funzione $Y=f(X_1, \dots, X_K)$ e le distribuzioni delle grandezze di ingresso X_i (e quindi non solo le coppie $\langle m_i, u_i \rangle$). Un campione di input è una K -upla di valori $\langle x_1, \dots, x_K \rangle$, in cui ogni x_i è ottenuto dalla distribuzione associata a X_i (se le variabili casuali X_i non fossero indipendenti la K -upla dovrebbe essere ottenuta per campionamento dalla distribuzione congiunta). Evidentemente da $\langle x_1, \dots, x_K \rangle$ si può calcolare un valore y per Y , $y=f(x_1, \dots, x_K)$. Supponiamo ora che queste due operazioni – campionamento dalle K variabili di input a f e calcolo di f sul campione così ottenuto – siano ripetute un numero W sufficientemente elevato di volte. Si genera in questo modo un campione $\langle y_1, \dots, y_W \rangle$ di valori per Y . Se W è appunto sufficientemente elevato (dell'ordine delle decine di migliaia di elementi o più), la sua distribuzione a frequenze relative si può affidabilmente considerare non distinguibile dalla distribuzione di probabilità di Y . Il valor medio e la deviazione standard del campione forniscono dunque i valori m_Y e u_Y cercati, ma da tale distribuzione si possono anche calcolare direttamente gli intervalli di confidenza per ogni dato livello di confidenza richiesto.

Questa tecnica dunque “propaga le distribuzioni”, e non solo le incertezze, e ha anche il merito di non richiedere la conoscenza dei coefficienti di sensibilità di f .

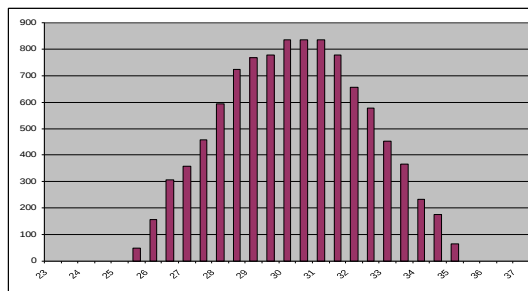
Un esempio

Supponiamo (valori numerici non realistici e senza indicazione di unità di misura):

$$Y=f(X_1, X_2)=X_1+X_2$$

con entrambe le grandezze di ingresso distribuite uniformemente, negli intervalli 10 ± 2 e 20 ± 3 rispettivamente e statisticamente indipendenti.

Utilizzando per esempio un foglio di calcolo, generiamo W campioni di input, cioè W coppie di valori distribuiti come richiesto (confidando dunque nell'uniformità del generatore di numeri casuali del programma) e sommiamo i valori di ognuna delle W coppie. La distribuzione dei valori y ottenuti può essere allora visualizzata in istogramma:



e da essa possono essere calcolati m_Y e u_Y .

Naturalmente la qualità dell'operazione dipenderà sia dalla qualità del generatore di numeri casuali adottato, sia dalla dimensione W del campione (nel caso raffigurato, $W=10000$) (in questo caso non è difficile confrontare i risultati ottenuti con Monte Carlo con quelli analitici: le incertezze tipo sono $u_1=2/\sqrt{3}$ e $u_2=3/\sqrt{3}$, e quindi $u_Y=\sqrt{u_1^2+u_2^2}\approx 2,08$).