Cenni alla teoria degli errori casuali

Sulla base di due serie di misurazioni numeriche, presentiamo le principali nozioni della teoria degli errori casuali di Gauss e, in particolare, intendiamo chiarire i significati di media aritmetica e di deviazione standard, nozioni che permettono di sintetizzare adeguatamente un intero processo di misura.

Media e scarti

Disponendo di uno strumento di misura sufficientemente sensibile quando eseguiamo la misura di una grandezza fisica X (per es. la durata di un certo fenomeno) nelle medesime, per quanto possibile, condizioni sperimentali, otteniamo di norma una serie $x_1, x_2, \ldots x_n$ di n valori numerici, generalmente diversi. Assodato che una tale serie di misure contenga un numero adeguato di misure cioè il numero n sia relativamente grande, si pone il problema, tipico della statistica, di quale valore assegnare alla grandezza X e, cosa altrettanto significativa, quale sia il suo grado di attendibilità. Poiché i singoli valori, pur essendo generalmente uno diverso dall'altro, sono stati rilevati nelle medesime condizioni, la teoria degli errori casuali ci assicura che la loro media aritmetica

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

risulta essere il valore più vicino al valore vero, valore questo che non si conosce né si può conoscere. La media aritmetica rappresenta quindi, non il valore vero, ma solo una sua stima, la migliore possibile.
In effetti la media aritmetica, se rappresentata come una retta orizzontale su un diagramma a dispersione delle misure (*scatter plot*, figure 1 e 2), dà la medesima importanza ai valori maggiori $x_i > \overline{x}$ così come ai valori che ne sono minori, $x_i < \overline{x}$: in termini più tecnici, la media aritmetica assicura che la somma degli scarti z_i

$$z_i = x_i - \overline{x}$$

definiti come la differenza tra la singola misura e la media, sia nulla.

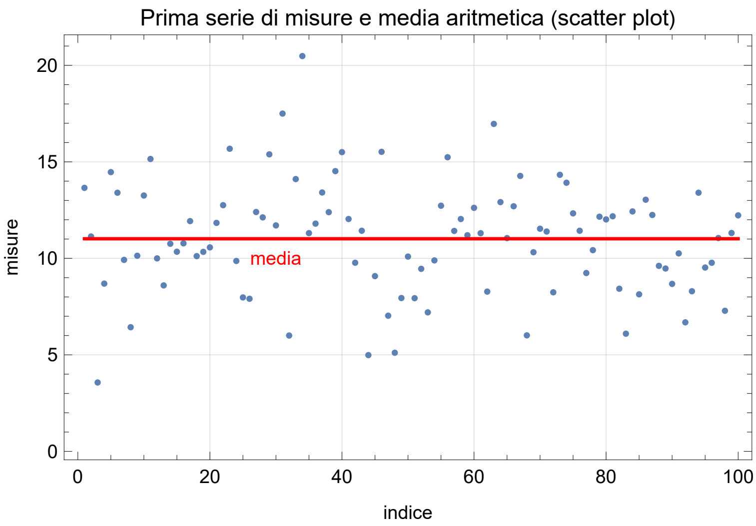

Figura 1

D'altra parte la media aritmetica non è in grado di distinguere tra la distribuzione di figura 1 e l'insieme di figura 2. Entrambe presentano $n = 100$ misure e le loro medie sono sostanzialmente uguali

$$\overline{x}_1 = 11.02 \qquad \text{e} \qquad \overline{x}_2 = 11.03$$

ma la seconda serie mostra una variabilità maggiore nei valori rilevati.

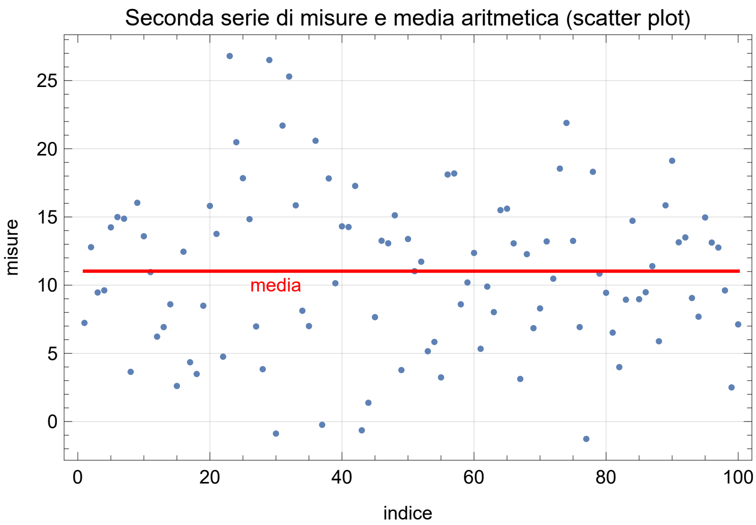

Figura 2

Difatti le misure nel secondo caso appaiono disperse in un intervallo di ampiezza maggiore rispetto al primo. Ovviamente, se dovessimo scegliere tra i due insiemi la scelta, per ora motivata solo dall'osservazione qualitativa dei due grafici, cadrebbe sulla prima serie nella quale probabilmente si è fatto uso di uno strumento più preciso. Nasce quindi l'esigenza di determinare una ulteriore, nuova, grandezza che ci informi sul grado di dispersione delle misure.

Istogrammi e "distribuzione normale"

A tal fine, dato il numero di misure presenti nei due insiemi, possiamo costruire una rappresentazione grafica della loro distribuzione nella forma di un istogramma. Questa particolare rappresentazione si ottiene disponendo le misure lungo l'asse orizzontale e, suddivisa la regione dell'asse orizzontale compresa tra il minimo e il massimo in intervallini di ampiezza Δ opportuna (la scelta è arbitraria ma va fatta con attenzione), si riporta sull'asse verticale il numero di misure comprese in ciascun intervallino.

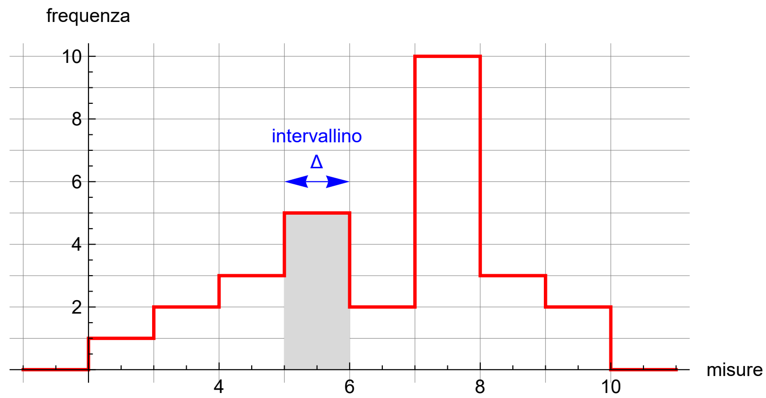

Figura 3

La figura precedente esemplifica le scelte e mostra come si sia scelta sull'asse delle misure (a loro volta comprese tra un minimo di 122 e un massimo di 130), un'ampiezza pari ad 1 degli intervallini e come, tra 125 e 126, cadano 5 loro valori mentre sono 10 quelli compresi in $[127, 128]$.

Riprese le due serie iniziali osserviamo come le misure oscillino tra un minimo prossimo allo zero e un massimo attorno a 28. Scegliamo pertanto intervallini di ampiezza unitaria e riportiamo in ordinata gli esiti di questo conteggio che indichiamo come *frequenza* (o *frequenza assoluta*) delle misure.

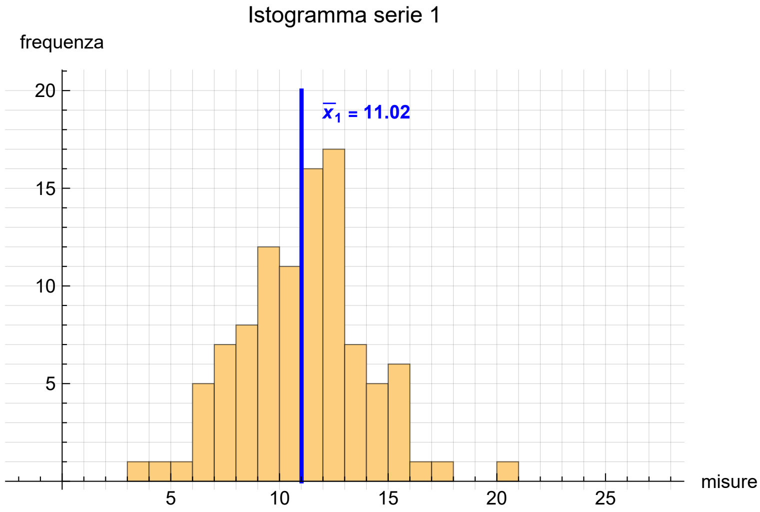

Figura 4

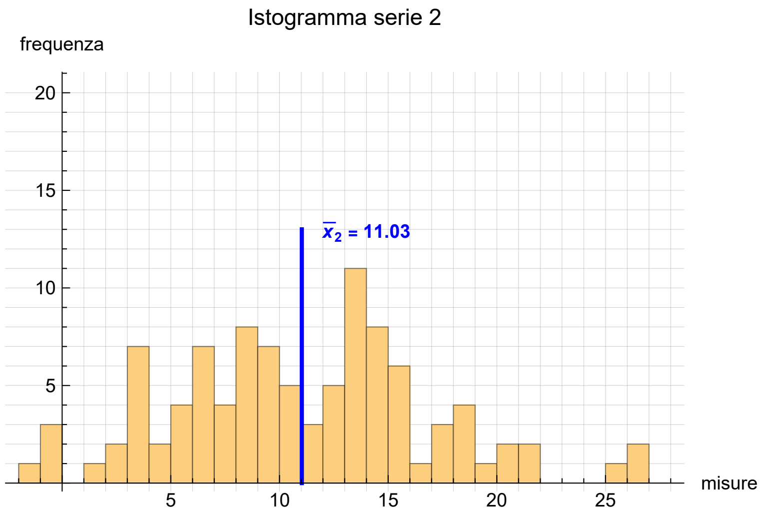

Figura 5

Quello che otteniamo è il tipico istogramma di un insieme di misure casuali (in statistica, detti *campioni*). Entrambi gli istogrammi (figg. 4 e 5), pur mostrando una evidente variabilità nella frequenza, condividono comunque alcune caratteristiche comuni: entrambi appaiono approssimativamente simmetrici rispetto alla retta verticale associata alla media aritmetica delle misure e, ancora entrambi, mostrano un addensamento delle frequenze maggiori in prossimità della media mentre queste tendono ad assumere valori sempre minori mano a mano che ci si allontana da questa.

I due istogrammi comunque si differenziano in modo significativo in quanto il secondo mostra, come aspettato, una larghezza maggiore assieme ad un altrettanto significativo appiattimento. La prima caratteristica è evidentemente associata alla maggiore dispersione delle misure mentre la seconda è collegata ad una caratteristica intrinseca che accomuna entrambi gli istogrammi: difatti entrambi condividono la medesima area (in colore nelle figure 4 e 5), area che per come sono stati costruiti è pari al numero delle misure che sappiamo essere 100. Di conseguenza *se la frequenza delle misure prossime alla media diminuisce, necessariamente l'istogramma si deve allargare*.

La grandezza che cerchiamo dev'essere quindi, in qualche modo, collegata alla larghezza di ciascun istogramma. Per comprenderne l'origine riportiamo una serie di istogrammi ciascuno relativo ad un numero n di misure sensibilmente crescente (fig. 6).

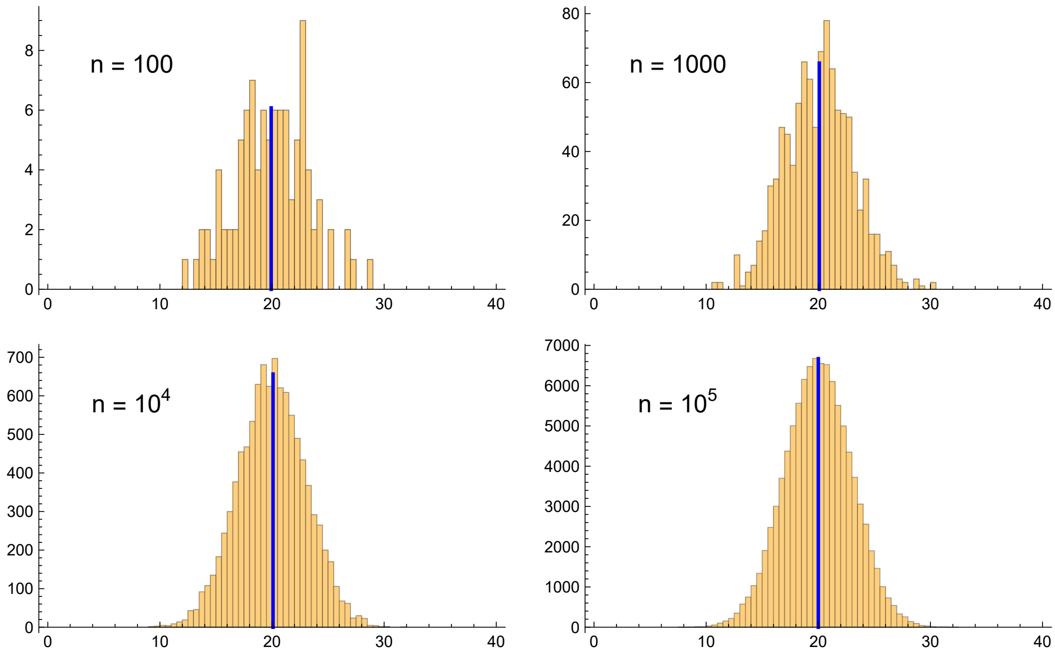

Figura 6

Emerge allora con evidenza come, più misure si facciano, più l'istogramma si avvicina a

- una curva dalla forma a campana,
- dotata di un massimo in corrispondenza della media (in blu in fig. 6),
- simmetricamente disposta attorno a questo valore e con un
- andamento i cui valori tendono allo zero mano a mano che ci si allontana da esso (figura 7).

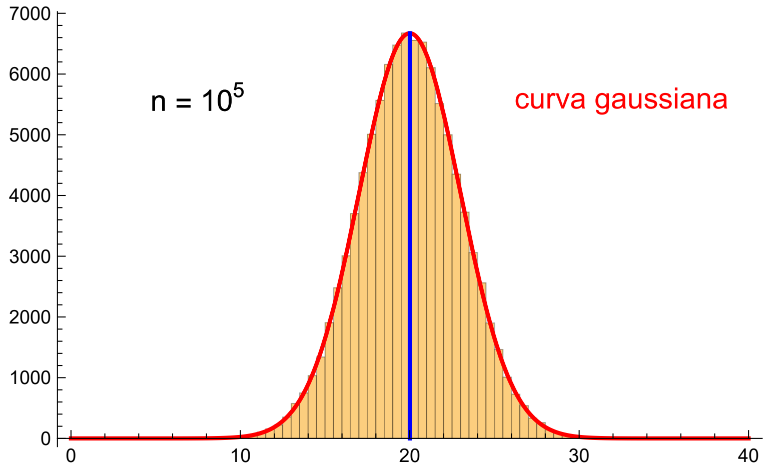

Figura 7

La funzione matematica che descrive l'andamento visualizzato sopra è chiamata *distribuzione normale* o *funzione di Gauss* o *curva gaussiana*. Questa è rappresentata matematicamente da una equazione che dipende unicamente da due parametri,

- il valore vero della grandezza che noi associamo alla *media aritmetica* $\overline{x}$ e
- la sua *deviazione standard* o *scarto quadratico medio* σ.

In corrispondenza della media tale curva raggiunge il suo massimo ed è simmetrica attorno ad essa mentre la deviazione standard definisce la sua larghezza. Questa larghezza è rappresentata graficamente dal segmento orizzontale di figura 8 che individua i punti di ascissa $\overline{x} \pm \sigma$ in corrispondenza dei quali la curva modifica la sua concavità/convessità.

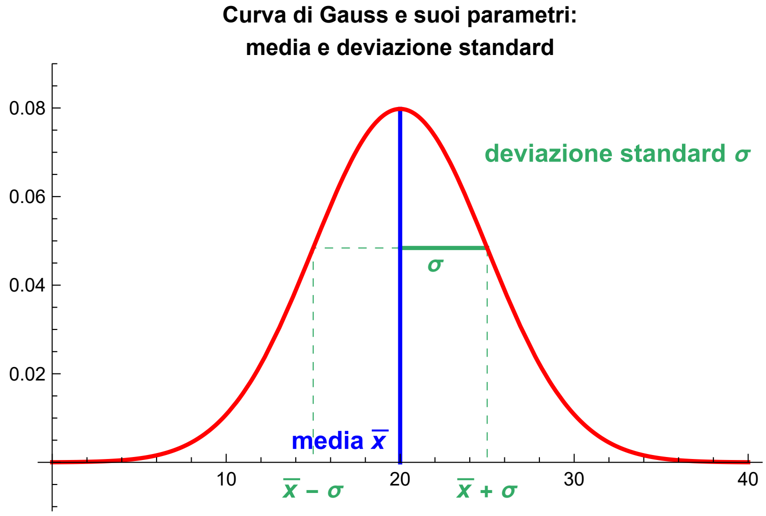

Figura 8

Deviazione standard e probabilità

Se come già accennato inizialmente, il significato di media aritmetica di un insieme di valori più o meno ripetuti è facilmente comprensibile, qui intendiamo chiarire ruolo e significato della deviazione σ.
Pertanto in figura 9 riportiamo tre gaussiane, tutte caratterizzate dalla stessa media (nulla) ma con deviazioni standard pari a 1, 2 e 3. Appare evidente come all'aumentare della deviazione standard la curva si appiattisca allargandosi e ciò accade in quanto, come diremo fra poco, l'area sottostante la curva deve rimanere costante.

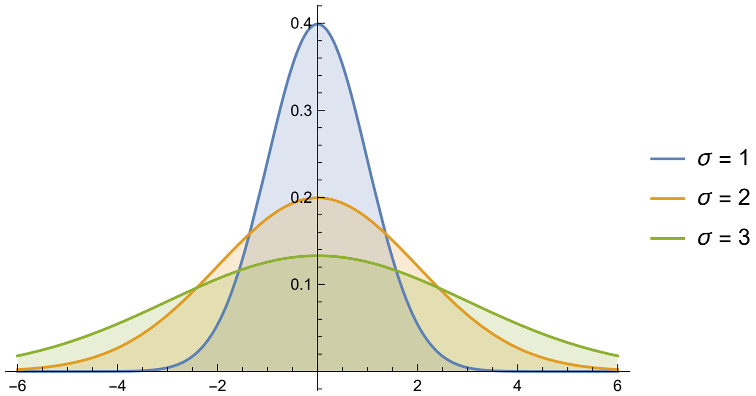

Figura 9

Se quindi la media ci fornisce un valore prossimo a quello vero, la deviazione standard ci informa sull'entità della dispersione delle misure. La stima di questa importante caratteristica è data dalla formula

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

dove sotto radice quadrata compare il rapporto tra la somma dei quadrati degli scarti con $n - 1$.
Ripresi i due insiemi iniziali, poiché è nota la numerosità n dei campioni e la larghezza Δ degli intervallini, nonché medie $\overline{x}_1 = 11.02$ (fig. 4) e $\overline{x}_2 = 11.03$ (fig. 5) e ora le deviazioni, $\sigma_1 = 2.88$ e $\sigma_2 = 5.93$, siamo in grado di associare la rispettiva curva gaussiana ai due istogrammi e, come conseguenza, confrontare anche numericamente le due serie: difatti la prima è evidentemente caratterizzata da una precisione che è il doppio della seconda essendo la sua deviazione circa la metà dell'altra (fig. 10).

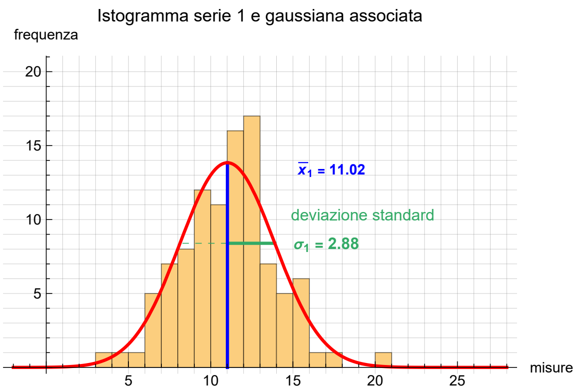

Figura 10

Inoltre, come entrambi gli istogrammi definiscono regioni del piano con la stessa area in quanto relativi al medesimo numero di misure, così accade per la gaussiana. Difatti la regione compresa tra la gaussiana e l'asse orizzontale possiede un'area uguale a quella degli istogrammi cosicché, ad elementi degli istogrammi che superano la gaussiana devono corrispondere dei vuoti dei medesimi sotto di essa.
In definitiva, il ruolo della deviazione formalizza quanto già osservato sugli istogrammi: un piccolo valore della deviazione σ è indice di una distribuzione con un massimo accentuato e stretto cui si accompagna una maggiore precisione mentre un valore di σ maggiore indica una distribuzione appiattita e allargata e quindi di minor precisione.
Il ruolo dell'area sottostante la gaussiana assume notevole importanza in quanto, nella teoria, il suo valore è direttamente associato alla probabilità. In particolare, poiché la curva gaussiana si estende sull'intero insieme di misure poste in ascissa, l'area dell'intera regione compresa tra la curva e l'asse orizzontale esprime la certezza di ottenere, in una rilevazione, un qualsivoglia valore e ciò equivale ad affermare che la sua probabilità è del 100%. Se invece ci limitiamo ad aree comprese tra due estremi qualsiasi dell'asse delle misure, per esempio a e b, allora l'area sottostante alla gaussiana dà la probabilità P che una misura x dia come risultato un valore compreso nell'intervallo $a \le x \le b$ (figura 11). Esprimiamo questa probabilità nella forma

$$P(a \le x \le b).$$

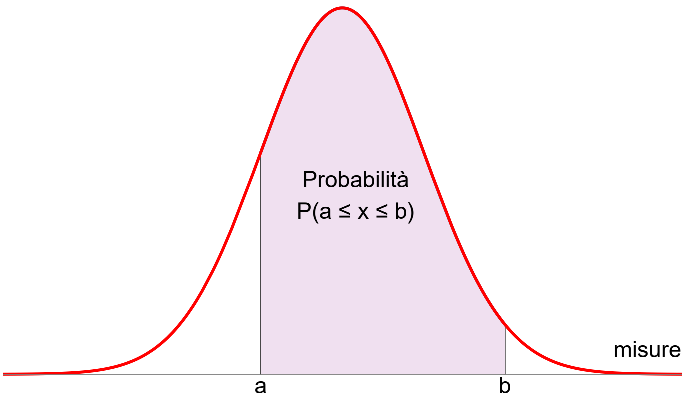

Figura 11

Se poi colleghiamo gli estremi del precedente intervallo alla media $\bar{x}$ e alla deviazione σ, allora la probabilità che una misura cada entro una deviazione standard dalla media è del 68% ossia risulta

$$P(\bar{x} - \sigma \le x \le \bar{x} + \sigma) = 68\%,$$

probabilità che associamo all'area della regione nella figura 12. Detto in altro modo, siamo confidenti al 68% che una misura cada internamente all'intervallo $[\bar{x} - \sigma, \bar{x} + \sigma]$.

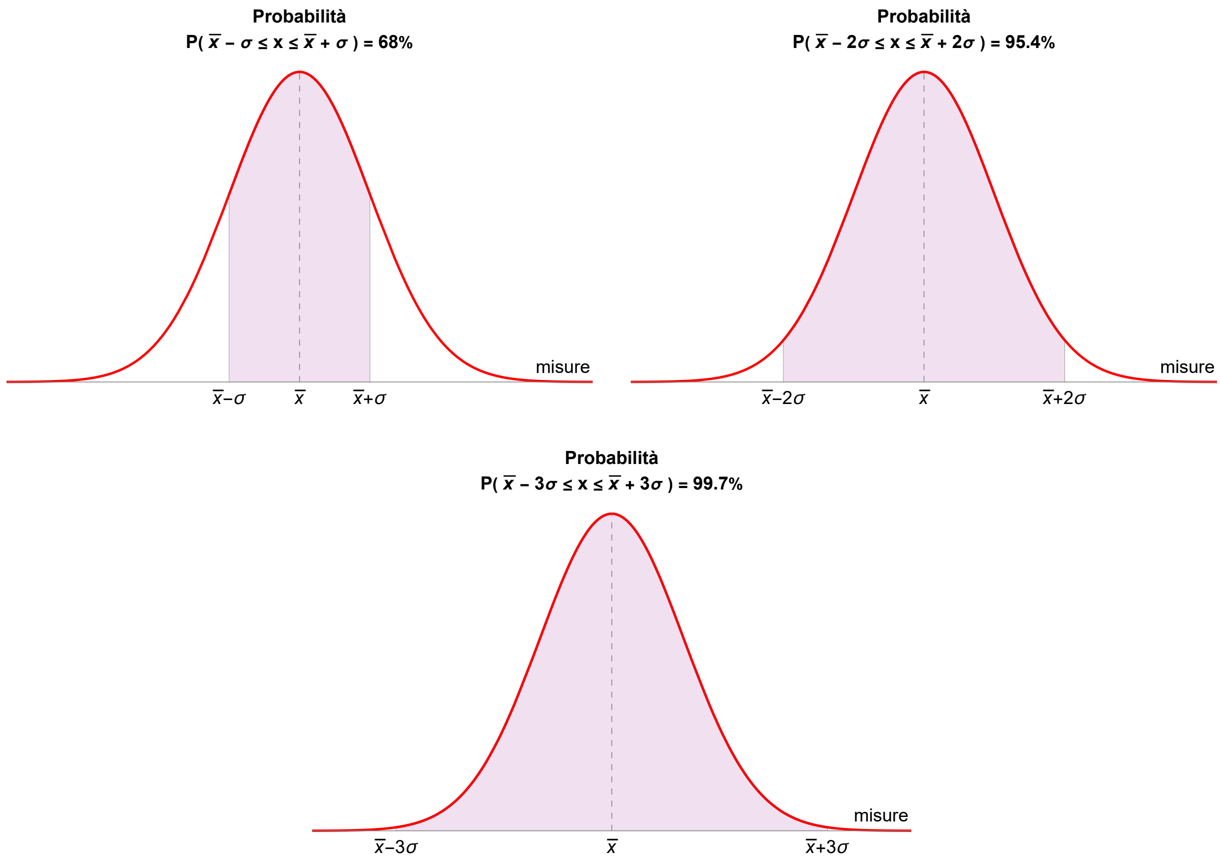

Figura 12

Se estendiamo la semiampiezza dell'intervallo a 2σ (fig. 12), la probabilità che una misura vi cada all'interno sale al 95.4% e così il nostro livello di confidenza

$$P(\overline{x} - 2\sigma \leq x \leq \overline{x} + 2\sigma) = 95.4\%,$$

mentre se aumentiamo la semiampiezza a 3σ

$$P(\overline{x} - 3\sigma \leq x \leq \overline{x} + 3\sigma) = 99.7\%$$

possiamo aspettarci con quasi certezza che la misura vi sia compresa.

Per esprimere in modo alternativo questi risultati, la probabilità che una misura cada al di fuori di questi intervalli è ancora apprezzabile nel primo caso (32%) mentre scende a meno del 5% nel secondo ed è praticamente trascurabile nel terzo.

È sulla base di queste osservazioni che riprendiamo le due serie iniziali associando alle rispettive gaussiane le regioni aventi aree corrispondenti a questi tre livelli di probabilità (figg. 13 e 14).

Per la prima serie solo una misura esce dall'intervallo di semiampiezza 3σ ma ciò nonostante non va rigettata in quanto la sua frequenza è compatibile con la gaussiana e non è elevata. Diversa sarebbe la conclusione se si presentasse a questi estremi una maggiore frequenza quale quella mostrata in figura 15. Se ciò accadesse vorrebbe dire che durante il processo di misura dev'essere intervenuto un qualche fattore estraneo tale da modificare le condizioni preesistenti.

L'istogramma della seconda serie è, a maggior ragione, coerente con la gaussiana in quanto tutte le misure cadono nell'intervallo di estremi $[\overline{x}-3\sigma, \overline{x}+3\sigma]$ cosicché entrambe le serie non presentano anomalie e sono del tutto accettabili anche se, con tutta evidenza, la seconda dev'essere stata acquisita con strumenti meno precisi.

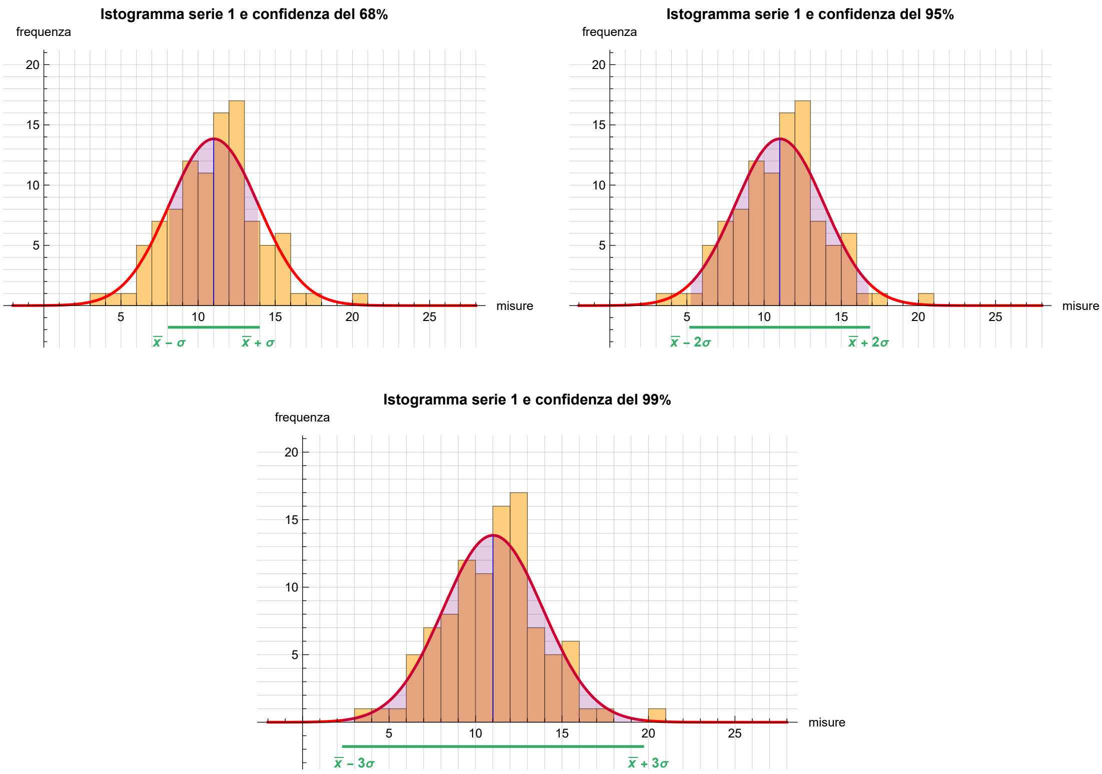

Figura 13

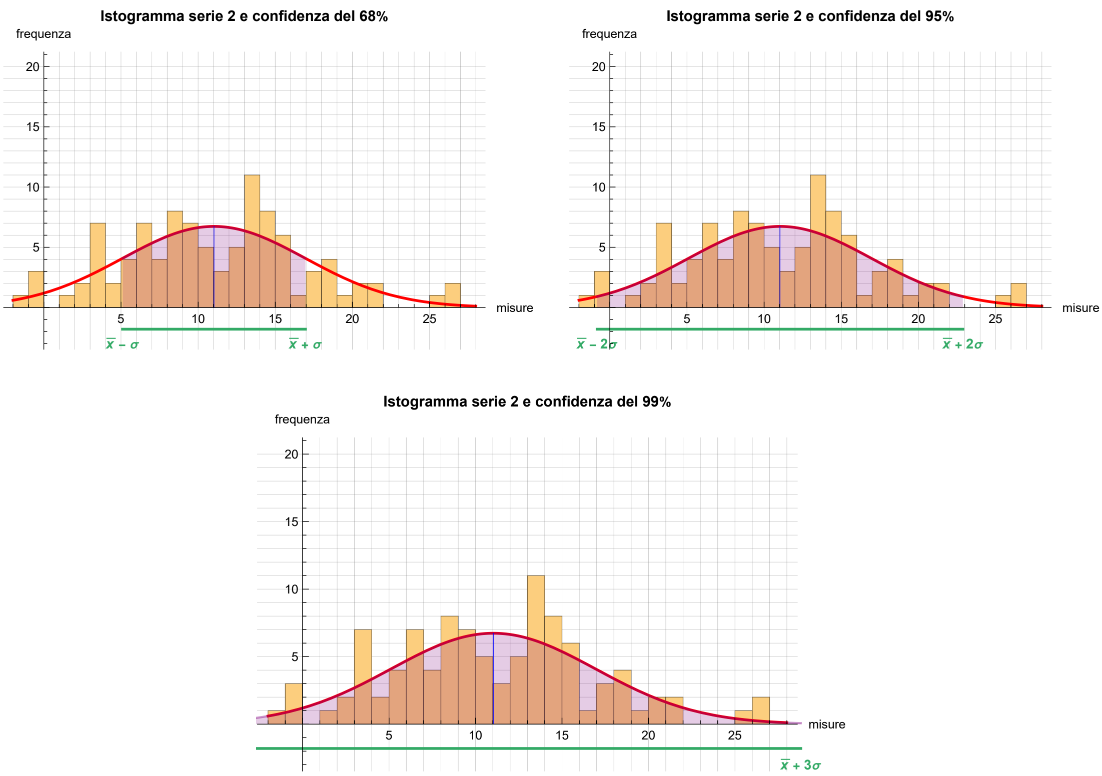

Figura 14

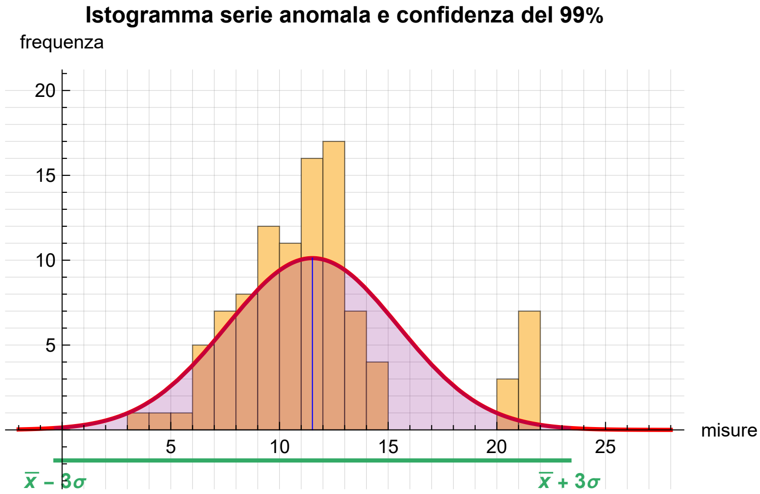

Figura 15

La deviazione standard della media

Il percorso finora seguito ci ha permesso di associare ad un insieme di misure, rilevate tutte nelle medesime condizioni sperimentali, due importanti grandezze statistiche, la media $\bar{x}$ e la deviazione standard σ. Sappiamo che la media rappresenta la miglior stima del valore vero della grandezza che si vuole misurare mentre la deviazione standard ci permette di stimare la bontà di una misura in quanto possiamo collegarla alla sua probabilità di presentarsi.
D'altra parte, se dovessimo utilizzare per qualche scopo la media vorremmo anche conoscere quale sia la sua attendibilità o, il che è lo stesso, il grado di confidenza che possiamo assegnarle. Per rispondere a tale esigenza (ovvia in ambito scientifico, dove una grandezza assume significato solo se è associata ad una stima del suo errore!), potremmo seguire il medesimo percorso che, dalla singola misura ci ha portato ad un insieme di misure e quindi alla media e alla deviazione. In sostanza, trattando la media ottenuta come l'espressione di una singola misura questa volta però appartenente ad una nuova serie costituita da altrettante medie, dalla quale calcolare la "media" delle medie e la sua deviazione.
È evidente comunque l'impraticabilità di una tale procedura ma gli sviluppi nella teoria ci permettono, nonostante tale difficoltà, di rispondere positivamente all'esigenza iniziale e quindi di disporre di una stima della bontà dell'unico valor medio effettivamente ottenuto. In altre parole siamo in grado di associare alla media la sua deviazione che indichiamo come *deviazione standard della media* o *scarto quadratico medio della media* $\bar{\sigma}$ (o anche σ_m). Questa nuova grandezza statistica è strettamente collegata alla deviazione standard σ dalla formula

$$\bar{\sigma} = \frac{\sigma}{\sqrt{n}}$$

e mostra come l'attendibilità della media cresca (e quindi $\bar{\sigma}$ diminuisca) con l'aumento del numero n delle misure. Risultato ovvio si potrebbe pensare, ma ora disponiamo pure di come l'attendibilità dipenda da n: se quindi, per esempio, intendiamo raddoppiarla dovremo aumentare n di 4 volte!
Per le due serie iniziali, composte da $n = 100$ misure, con deviazioni standard, rispettivamente $\sigma_1 = 2.88$ e $\sigma_2 = 5.93$, le rispettive deviazioni delle due medie sono $\bar{\sigma}_1 = 0.288$ e $\bar{\sigma}_2 = 0.593$.
Poiché valgono le medesime osservazioni circa la probabilità di ottenere una nuova media x interna a opportuni intervalli, possiamo riaffermare le seguenti probabilità

$$P(\bar{x} - \bar{\sigma} \leq x \leq \bar{x} + \bar{\sigma}) = 68\%$$
$$P(\bar{x} - 2\bar{\sigma} \leq x \leq \bar{x} + 2\bar{\sigma}) = 95.4\%$$
$$P(\bar{x} - 3\bar{\sigma} \leq x \leq \bar{x} + 3\bar{\sigma}) = 99.7\%$$

dove alla deviazione standard si è sostituita la deviazione standard della media.
In forma alternativa com'è usuale in ambito scientifico, possiamo esprimere l'esito complessivo dell'intero processo di misura della grandezza X,

- con confidenza del 68%, $\quad X = \bar{x} \pm \bar{\sigma}$,

- con confidenza del 95.4%, $X = \overline{x} \pm 2\overline{\sigma}$,
- con confidenza del 68%, $X = \overline{x} \pm 3\overline{\sigma}$.

Per concludere con i nostri esempi, scelta una confidenza del 95.4% e deviazioni standard delle medie approssimate a due cifre significative, gli esiti finali sono

$$X_1 = 11.02 \pm 0.58 \qquad \text{e} \qquad X_2 = 11.03 \pm 1.2.$$