

# Nota sull'esperienza pratica "Misura degli angoli interni di triangoli"

fuso@df.unipi.it; <http://www.df.unipi.it/~fuso/dida>

(Dated: versione 1 - 24 aprile 2011)

Queste pagine riportano una breve nota sull'esperienza pratica condotta in aula, in particolare sul trattamento dati. Esse intendono in particolare costituire un approfondimento degli argomenti riguardanti la misura di variabili stocastiche e anche una guida utile per la realizzazione della relazione relativa all'esperienza.

## I. MODALITÀ, SCOPO E STRUMENTI

L'esperienza prevede di disegnare un certo numero di triangoli di vario tipo, di misurarne gli angoli interni con un semplice goniometro e di registrare la somma dei tre angoli. La situazione presa in considerazione è più unica che rara: infatti il risultato della misura è noto a priori (la somma degli angoli interni è "per definizione" pari a 180 gradi) e quindi lo scopo non è quello di conoscere il valore incognito di una grandezza. Come ogni altra misura, anche quella degli angoli può essere affetta da diverse cause di incertezza, sia di tipo sistematico (per esempio un goniometro con calibrazione errata o usato in modo sistematicamente sbagliato), che di tipo stocastico (per esempio legate allo spessore delle linee tracciate, all'errore di parallasse, all'operazione di diversi sperimentatori (!), a fluttuazioni di varia natura). Normalmente in misure di questo tipo il carattere stocastico è prevalente, e ci si aspetta che esso domini la misura.

Lo scopo principale dell'esperienza è quindi verificare praticamente quali siano i metodi da adottare per l'analisi di dati influenzati da incertezza stocastica. Lo strumento di misura è il goniometro; in questa relazione si suppone che la sensibilità dello strumento sia 1 grado, cioè che esso permetta di determinare la misura dei singoli angoli con un'incertezza strumentale di  $\pm 0.5$  gradi. Inoltre si richiede di creare un campione di misure sufficientemente grande: si considera infatti di avere a disposizione  $N = 60$  misure, ognuna costituita da una misura indipendente della somma degli angoli interni. Questo campione può essere creato praticamente o disegnando e misurando 60 triangoli diversi, oppure attraverso misura di un insieme più piccolo di triangoli da parte di diversi sperimentatori.

## II. RAPPRESENTAZIONE E ANALISI DEI DATI

I dati raccolti in successione formano una tabella dei valori  $S_i$  della somma degli angoli interni, con  $i$ , indice che identifica misure diverse, che va da 1 a 60. Il valore della grandezza  $S_i$ , misurata sperimentalmente, deve essere corredato dell'incertezza di misura  $\Delta S_i$ . La determinazione quantitativa dell'incertezza è in parte affidata alla scelta dello sperimentatore.

Normalmente, per determinare l'incertezza di una misura si fa riferimento a tutte le cause che possono essere individuate facilmente. Questo vuol dire che in genere la valutazione dell'incertezza non tiene conto delle cause stocastiche, le quali possono essere evidenziate solo a posteriori, cioè dopo aver eseguito le misure e analizzato i dati. Inoltre può esserci un certo grado di arbitrarierà nello stabilire l'incertezza da attribuire alle misure, soprattutto quando si usano strumenti la cui correttezza di impiego è affidata alla capacità dello sperimentatore. In questo esempio, dove si fa uso di uno strumento semplicissimo, il goniometro, si suppone che lo sperimentatore sia in grado di determinare il valore di ogni singolo angolo con un'incertezza di  $\pm 0.5$  gradi, e si suppone che questa incertezza sia quella che prevale nella misura.

Poiché la misura di interesse, la somma degli angoli  $S_i$ , è la somma delle misure dei singoli angoli interni, per valutare  $\Delta S_i$  è necessario usare le regole della propagazione dell'errore. Si può ragionevolmente supporre che le misurazioni dei singoli angoli di ogni triangolo siano indipendenti fra loro[1]: dette  $\Delta\alpha_i, \Delta\beta_i, \Delta\gamma_i$  le incertezze nella misura dei singoli angoli interni del triangolo  $i$ -esimo,  $\alpha_i, \beta_i, \gamma_i$ , la propagazione dell'errore sulla somma  $S_i = \alpha_i + \beta_i + \gamma_i$  stabilisce:  $\Delta S_i = \sqrt{(\Delta\alpha_i)^2 + (\Delta\beta_i)^2 + (\Delta\gamma_i)^2}$ . Poiché  $\Delta\alpha_i = \Delta\beta_i = \Delta\gamma_i = 0.5$  gradi, è facile verificare che  $\Delta S_i = \sqrt{3} \times 0.5$  gradi  $\approx 0.9$  gradi. Di conseguenza attribuiamo a ogni singola misura  $S_i$  della somma degli angoli interni un'incertezza di  $\pm 0.9$  gradi.

La tabella riportata in Fig. 1 mostra la lista delle misure. Si noti come esse siano "risolte" al mezzo grado, in conseguenza del fatto che i dati di partenza sono ottenuti con uno strumento dotato di sensibilità 0.5 gradi.

### A. Rappresentazione grafica

La tabella contiene tutte le informazioni rilevanti dell'esperimento, ma la sua "leggibilità" è scarsa. Come di consueto, per mettere in evidenza i risultati di un esperimento conviene utilizzare rappresentazioni grafiche.

La prima rappresentazione grafica proposta, quella di Fig. 2, mostra i risultati sperimentali  $S_i$  in funzione dell'indice di misura  $i$ , cioè del triangolo considerato. Come

numero misura	$S_i$ [gradi]	$\Delta S_i$ [gradi]
1	181.5	0.9
2	179.5	0.9
3	181.0	0.9
4	181.0	0.9
5	179.0	0.9
6	180.0	0.9
7	178.0	0.9
8	179.5	0.9
9	179.0	0.9
10	180.5	0.9
11	180.5	0.9
12	179.5	0.9
13	179.0	0.9
14	181.5	0.9
15	180.0	0.9
16	179.5	0.9
17	181.0	0.9
18	177.0	0.9
19	179.5	0.9
20	180.5	0.9

numero misura	$S_i$ [gradi]	$\Delta S_i$ [gradi]
21	179.5	0.9
22	181.0	0.9
23	177.5	0.9
24	181.5	0.9
25	180.5	0.9
26	181.5	0.9
27	179.0	0.9
28	178.5	0.9
29	180.5	0.9
30	178.0	0.9
31	179.5	0.9
32	177.5	0.9
33	181.0	0.9
34	180.0	0.9
35	180.0	0.9
36	177.5	0.9
37	178.5	0.9
38	178.0	0.9
39	181.5	0.9
40	180.0	0.9

numero misura	$S_i$ [gradi]	$\Delta S_i$ [gradi]
41	182.5	0.9
42	179.5	0.9
43	179.5	0.9
44	181.5	0.9
45	180.0	0.9
46	181.5	0.9
47	181.0	0.9
48	180.0	0.9
49	178.0	0.9
50	180.0	0.9
51	181.0	0.9
52	177.5	0.9
53	180.0	0.9
54	180.0	0.9
55	178.5	0.9
56	179.5	0.9
57	179.0	0.9
58	181.0	0.9
59	181.5	0.9
60	180.0	0.9

Figura 1. Tabella delle misure.

necessario in ogni rappresentazione di questo tipo, i dati sperimentali sono corredati dalla barra di errore (in questo caso tutte le barre di errore sono uguali tra loro e pari a  $\Delta S_i = 0.9$  gradi). Il grafico mostra anche due righe tratteggiate orizzontali, che intercettano l'asse verticale in corrispondenza dei valori 179.1 gradi e 180.9 gradi: queste due righe sono in pratica i limiti inferiore e superiore delle misure ritenute “veridiche” sulla base dell'incertezza  $\Delta S_i$ . Infatti, in linea di principio, sapendo a priori che il valore della somma degli angoli interni è 180 gradi, ci aspetteremmo che tutte le misure cadano nell'intervallo definito dall'incertezza attorno a 180 gradi.

Come si vede dal grafico, questo non si verifica affatto. La spiegazione è ovvia: nella nostra determinazione dell'incertezza abbiamo trascurato ogni causa di tipo stocastico, che in effetti non possiamo valutare a priori, ma che evidentemente influenza la misura.

Una rappresentazione di questo tipo non è comunque particolarmente significativa. In effetti essa suggerisce qualcosa, ad esempio che l'incertezza  $\Delta S_i$  è sottostimata, secondo quanto già affermato. Inoltre essa potrebbe essere utile per verificare eventuali “tendenze” della misura: per esempio, se il goniometro modificasse in modo monotono la sua calibrazione (che so, un aumento di temperatura della stanza che comporta un aumento delle dimensioni fisiche del goniometro, dunque la perdita di calibrazione), allora questo potrebbe riflettersi in un trend sistematico in una direzione (sottostima o sovrastima del valore). Ciò evidentemente non si verifica nel campione di misure considerato, che invece appare dominato da incertezza di tipo stocastico.

## B. Istogramma delle ricorrenze

Un modo più significativo di rappresentare i dati consiste nel costruire il cosiddetto *istogramma delle ricorrenze*

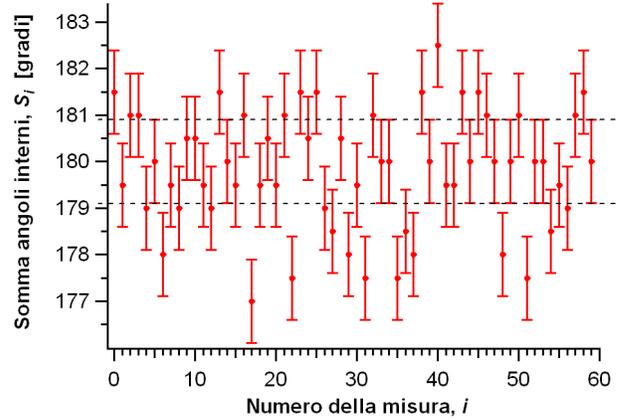


Figura 2. Rappresentazione grafica del campione di misure: l'asse verticale riporta il valore delle singole misure  $S_i$  con la rispettiva barra di errore (tutte le barre di errore sono uguali all'incertezza  $\Delta S_i = 0.9$  gradi), l'asse orizzontale riporta il numero d'ordine  $i$  della misura. Le righe tratteggiate mostrano l'intervallo nel quale ci aspetteremmo di trovare tutte le misure, come discusso nel testo.

(o occorrenze). Infatti è interessante poter capire in modo immediato (da un grafico!) quante volte, nell'ambito del campione considerato, la misura dà un certo valore, allo scopo di individuare se ci siano e quanto valgano valori della misura che risultano più probabili di altri.

L'istogramma è una forma di rappresentazione grafica che si presta benissimo a questo scopo. L'asse orizzontale di un'istogramma delle ricorrenze riporta il valore della misura, discretizzato in intervalli finiti detti “bins” (bin significa bidone, in pratica si distribuiscono i valori delle misure in tanti bidoni diversi); l'asse verticale riporta la ricorrenza del valore misurato, cioè, in pratica, quante volte la misura degli angoli cade nel bin considerato nell'ambito del campione che si sta esaminando.

Preliminarmente alla costruzione dell'istogramma occorre determinare l'estensione dei bins, cioè stabilire l'intervallo di valori di misura  $\Delta_k$  che definiscono ogni bin. Normalmente i bins hanno tutti la stessa estensione, cioè gli intervalli sono tutti uguali fra loro. Non esiste, però, una regola generale per la loro determinazione, che è affidata al buon senso dello sperimentatore. Se, come in questo caso, le misure sono affette da una data incertezza, allora può essere ragionevole identificare in prima battuta l'estensione del bin con l'incertezza della misura stessa. Infatti due misure sono indistinguibili fra loro se la loro differenza è minore della barra di errore e quindi è sensato che esse cadano nello stesso bin.

Cominciamo allora con il creare un'istogramma delle ricorrenze in cui l'estensione del bin è  $\Delta_k = 2\Delta S_i = 1.8$  gradi. Poiché in questo caso (eccezionale) sappiamo già qual è il valore “vero” della misura, 180 gradi, facciamo in modo che il valore 180 gradi cada al centro di un bin, cioè che ci sia un bin i cui estremi sono  $180 - 0.9 = 179.1$  gradi e  $180 + 0.9 = 180.9$  gradi. Troveremo gli estremi

degli altri bins adiacenti facendo delle semplici addizioni e sottrazioni. Per intenderci, avremo un bin compreso tra 177.3 gradi e 179.1 gradi, un altro tra 175.5 gradi e 177.3 gradi, un altro ancora tra 180.9 gradi e 182.7 gradi, e così via. Una volta determinati gli intervalli, analizzando la tabella è facile contare il numero  $n_k$  di volte in cui il valore della misura cade in un certo bin e quindi costruire l'istogramma. Il risultato è mostrato in Fig. 3.

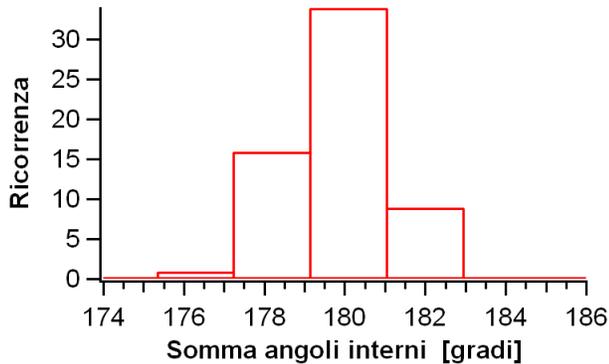


Figura 3. Iistogramma delle ricorrenze costruito come descritto nel testo: l'estensione di un bin è 1.9 gradi.

L'istogramma così costruito suggerisce che la maggior parte delle misure cade nell'intervallo compreso tra 179.1 gradi e 180.9 gradi (intervallo centrale nel grafico); quindi tali misure sono le più frequenti, o probabili, nel campione e ne costituiscono la cosiddetta *moda*. Pur essendo corretto, questo istogramma è ancora di limitata utilità. Infatti è evidente che la discretizzazione delle misure nei bins è poco accurata (solo 4 bins sono popolati). Possiamo cercare di ottenere una rappresentazione più significativa riducendo l'ampiezza  $\Delta_k$  del bin. Evidentemente, questo non può essere fatto in modo arbitrario. ad esempio, se scegliessimo degli intervalli di estensione  $\Delta_k = 0.1$  gradi, non potremmo mettere nulla nei bins compresi ad esempio tra 180.1 gradi e 180.2 gradi, tra 180.2 gradi e 180.3 gradi, etc. etc., dato che i dati di partenza sono risolti al mezzo grado. L'istogramma così costruito non avrebbe molto significato.

Potremmo però scegliere  $\Delta_k = 0.5$  gradi, cioè far coincidere l'estensione dei bins con la risoluzione dei dati, che in questo caso coincide con la sensibilità dello strumento. L'istogramma così costruito è mostrato in Fig. 4; a rigore, a causa dell'incertezza della misura, questo istogramma non ha molto significato, dato che, per esempio, le misure 180.0 gradi e 180.5 gradi differiscono tra di loro per meno dell'incertezza, 0.9 gradi, che abbiamo attribuito ai dati sperimentali; tuttavia esso è accettabile e, rispetto a quello mostrato in precedenza, ha il vantaggio di rappresentare con maggiore accuratezza il campione di misure.

Ovviamente le ricorrenze riportate sull'asse verticale dipendono dal numero totale  $N$  di misure del campione (in questo caso  $N = 60$ ). Qualche volta è comodo esprimere le ricorrenze in forma di *frazioni*, definite come rapporto tra le ricorrenze e il numero totale di misure: detto  $n_k$  il numero di ricorrenze che cadono nel bin  $k$ -esimo, la frazione è definita come  $F_k = n_k/N$ . I valori  $F_k$  rappresentano quante volte si ottiene, nell'ambito del campione considerato, una certa misura, relativamente alle dimensioni del campione stesso. Quindi essi possono essere interpretati come le *frequenze* di una certa misura: se tutto il campione fosse fatto della stessa misura, cioè se tutte le misure dessero lo stesso risultato, allora la frequenza sarebbe unitaria; viceversa, se una certa misura non uscisse mai, la sua frequenza sarebbe nulla. La frazione, o frequenza, è riportata nell'asse di destra del grafico di Fig. 4. Si vede come il risultato più frequente della misura, quello che fornisce valori compresi tra 179.75 gradi e 180.25 gradi, ha una frequenza superiore a 0.15, ovvero si verifica per oltre il 15% delle misure.

Talvolta, inoltre, fa comodo manipolare i dati in modo che l'area sottesa all'istogramma sia "normalizzata" a uno (numero puro): a questo scopo la frazione, o frequenza, va divisa per l'estensione del bin  $\Delta_k$ , ottenendo i valori  $f_k = F_k/\Delta_k$  che sono dimensionati come l'inverso della grandezza che si sta misurando ( $\text{gradi}^{-1}$ , nel nostro caso). L'asse aggiuntivo di destra del grafico di Fig. 4 si riferisce a tale grandezza, che chiamiamo *frequenza normalizzata* (a uno). Poiché la rappresentazione in frazione o frequenza, o frequenza normalizzata, non modifica i dati originari se non attraverso divisioni (si "riscala" l'asse verticale), la "forma" dell'istogramma non viene cambiata.

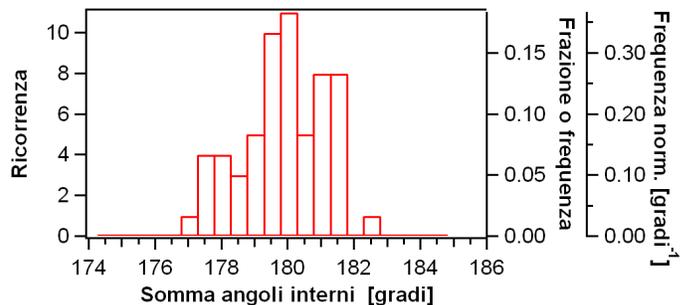


Figura 4. Iistogramma delle ricorrenze costruito come descritto nel testo: l'estensione di un bin è 0.5 gradi; gli assi di destra riportano frazione, o frequenza, e frequenza normalizzata a uno dei dati.

### C. Media e deviazione standard sperimentali

Affrontiamo ora il problema di derivare delle informazioni quantitative da un campione di misure affette da

incertezza stocastica. Infatti la rappresentazione con l'istogramma, benché utile per avere un'idea immediata di come sono distribuite le varie misure del campione e anche per stimare quale misura ha la massima frequenza, cioè la maggiore probabilità di verificarsi, non fornisce la risposta che generalmente si desidera ottenere quando si realizza un campione di misure. La domanda alla quale rispondiamo qui è quella di trovare un singolo valore che tenga conto dell'intero campione di misure e, possibilmente, di stabilire l'"affidabilità", cioè l'incertezza, di questo valore.

A questo scopo definiamo il *valore medio sperimentale*  $\mu_{exp}$  di un campione di misure  $x_i$  come:

$$\mu_{exp} = \frac{\sum_{i=1}^N x_i}{N}, \quad (1)$$

attribuendo a tale valore il significato di media delle misure contenute nel campione. Dunque  $\mu_{exp}$  può essere considerato come "risultato" sperimentale del campione di misure. Nel nostro caso si ha  $\mu_{exp} = 179.8$  gradi[2].

Può essere interessante notare che il valore medio sperimentale può essere dedotto sia dall'analisi dei dati di partenza (le  $N$  misure degli angoli interni) che da quella dell'istogramma delle frequenze. Infatti, ricordando che con  $\Delta_k, n_k, F_k = n_k/N$  indichiamo rispettivamente estensione del bin, numero di ricorrenze e frazione nel bin  $k$ -esimo, è facile verificare che:

$$\mu_{exp} = \frac{\sum_{k=1}^{N'} n_k x_k}{N} = \sum_{k=1}^{N'} F_k x_k, \quad (2)$$

dove con  $N'$  abbiamo indicato il numero totale di bins presenti nell'istogramma (per ovvie ragioni, sommando le ricorrenze che cadono in ogni bins si deve ritrovare il numero totale di misure, cioè deve essere  $\sum_{k=1}^{N'} n_k = N$ ).

L'affidabilità che questo valore ha nello stimare la misura considerata, cioè la valutazione dell'incertezza da attribuire alla stima, può essere dedotta valutando la discrepanza che le varie misure del campione hanno rispetto alla media. Dato che, come atteso nel caso di incertezza di origine stocastica, le discrepanze hanno segno sia negativo che positivo, occorre un qualche artificio matematico che consenta di considerare allo stesso modo sottostime e sovrastime. Questo può essere ottenuto calcolando la radice quadrata della somma degli scarti quadratici. Inoltre, per tenere conto del numero di misure che compongono il campione, è necessario che la somma degli scarti quadratici venga divisa per il numero di misure  $N$ . [3] In definitiva, si definisce una nuova grandezza, detta *deviazione standard* sperimentale, come:

$$\sigma_{exp} \approx \sqrt{\frac{\sum_{i=1}^N (x_i - \mu_{exp})^2}{N}}, \quad (3)$$

e si conclude che la media stimata  $\mu_{exp}$  ha un'incertezza pari a  $\pm\sigma_{exp}$ . [4] Notate che il simbolo  $\approx$  che compare in Eq. 3 è stato messo per ricordare quanto espresso nella nota, cioè che la definizione corretta prevede una divisione per  $N - 1$  e non per  $N$ , come qui considerato.

Svolgendo i conti, si vede che per il nostro campione è  $\sigma_{exp} = 1.3$  gradi. Di conseguenza possiamo concludere che il risultato dell'esperienza, condotto usando un campione di  $N = 60$  misure della stessa grandezza (somma degli angoli interni di triangoli diversi), è  $(179.8 \pm 1.3)$  gradi. L'esperienza è evidentemente "valida", dato che il valore atteso, 180 gradi, è bene all'interno dell'incertezza (deviazione standard sperimentale) determinata.

Infine vale la pena di ricordare che anche la valutazione della media sperimentale,  $\mu_{exp}$ , è affetta da un'incertezza, il cui valore è pari a  $\sigma_\mu = \sigma_{exp}/\sqrt{N}$ ; nel nostro caso tale errore della media, che indica l'accuratezza con la quale la media è stata determinata[5] vale 0.16 gradi.

### III. INTERPRETAZIONE: DISTRIBUZIONE DI DENSITÀ DI PROBABILITÀ

Esaminando le espressioni date sopra si vede come l'incertezza del valore medio sperimentale sia attesa migliorare con l'aumentare del numero di misure  $N$ , che sta a dividere nell'espressione di  $\sigma_{exp}$  (e soprattutto in quello di  $\sigma_\mu$ ). Questo risultato è del tutto ragionevole: ripetendo più e più volte una misura ci si aspetta di poter guadagnare in termini di accuratezza.

Alla base di questa semplice considerazione c'è una teoria che qui in questa nota riportiamo solo per sommi capi e per le conseguenze pratiche che essa comporta. La teoria parte da una constatazione che, sostanzialmente, stabilisce: *ogni campione contenente un numero finito di misure è un sottoinsieme di un campione fatto da tantissime (infinite) misure e di tale campione è rappresentativo.*

Vediamo di chiarire qual è il significato di questa affermazione. Partiamo con il notare che gli istogrammi che abbiamo disegnato rappresentano un modo per indicare come sono "distribuiti" i valori della misura. Per essere più precisi, essi sono una rappresentazione della *distribuzione di probabilità* delle misure. Infatti l'operazione di misura, essendo influenzata da fattori stocastici, conduce con una certa probabilità a un certo risultato, così come il lancio dei dadi conduce con una certa probabilità a un certo risultato[6]. Notate che gli istogrammi che abbiamo costruito, nascendo da un numero finito di misure affette da un'incertezza finita, hanno un carattere *discreto*. Infatti abbiamo in sostanza graficato delle ricorrenze  $n_k$ , frazioni o frequenze  $F_k = n_k/N$ , frequenze normalizzate  $f_k = F_k/\Delta_k$ , tutte grandezze che dipendono dall'indice (intero)  $k$  che identifica un bin (ovvero un valore di misura  $x_k$ ).

Si potrebbe supporre, idealmente, di aumentare a piacere il numero di misure che compongono il campione e di ridurre arbitrariamente l'incertezza a loro associata (basterebbe avere a disposizione tanto, tanto tempo e usare con molta attenzione un goniometro con una sensibilità altissima, cosa che nel mondo ideale è sicuramente possibile): così facendo sarebbe possibile costruire istogrammi in cui l'estensione dei bins è molto, molto piccola, cioè, in termini matematici,  $\Delta_k \rightarrow 0$ . Usando il linguag-

gio degli infinitesimi, e ricordando che l'estensione del bin ha, ovviamente, le dimensioni della grandezza  $x$  che si sta misurando, si potrebbe anche scrivere  $\Delta_k \rightarrow dx$ . La rappresentazione (di ricorrenze o frequenze) che se ne otterrebbe sarebbe più simile al grafico di una funzione *continua* che non a un istogramma.

In questo passaggio da discreto a continuo si finirebbe per costruire la *funzione di distribuzione della probabilità*  $f(x)$ , che avrebbe tutte le caratteristiche di una funzione matematica  $f$  valutata sui valori continui  $x$  possibili per la misura. Più precisamente, a partire dalla frequenza normalizzata  $f_k$  potremmo costruire la *distribuzione di densità di probabilità* (o densità di distribuzione di probabilità, la denominazione è così complicata che si possono tollerare inversioni tra i termini che la compongono!)  $f(x)$  tale che, in sostanza, il prodotto  $f(x)dx$  rappresenta la *probabilità* (infinitesima)  $dP$  di trovare un valore della misura compreso nell'intervallo (infinitesimo)  $[x, x + dx]$ . Se ben costruita, la funzione  $f(x)$  godrebbe delle tipiche proprietà della probabilità: ad esempio, per la normalizzazione si avrebbe  $\int_{-\infty}^{+\infty} f(x)dx = 1$ , cioè la probabilità integrata su tutti i possibili eventi (risultati della misura) varrebbe uno. Inoltre la probabilità che l'esito della misura sia contenuto nell'intervallo  $[x_1, x_2]$  sarebbe  $P(x_1, x_2) = \int_{x_1}^{x_2} f(x)dx$ .

La teoria della cui esistenza abbiamo prima fatto menzione stabilisce che, nel caso di misure affette da incertezza stocastica:

- la funzione che descrive la densità di distribuzione è una *Gaussiana*, cioè una funzione che ha la caratteristica forma “a campana” centrata su un certo valore e dotata di una certa larghezza;
- il valore centrale della campana e la sua larghezza sono *approssimabili* con i valori della media e della deviazione standard trovati sperimentalmente.

Queste affermazioni sono conseguenza di un teorema, detto teorema di Tschebyscheff (simile, nei suoi contenuti essenziali, alla popolare “legge dei grandi numeri”), secondo il quale la distribuzione di probabilità (discreta) determinata sperimentalmente è “figlia” di una funzione “genitrice” (*parent distribution*) che si esprime come:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (4)$$

con  $\mu = \mu_{exp}$  e  $\sigma = \sigma_{exp}$ . La funzione di distribuzione Gaussiana, detta anche *distribuzione normale*, o *standard*, ha alcune caratteristiche, contenute nella parte della funzione esponenziale (la frazione che sta a moltiplicare serve solo per garantire la corretta normalizzazione e non porta informazioni significative) che sono di particolare rilievo:

- come già affermato, la forma è a campana, cioè simmetrica attorno al valore medio  $\mu$ ;
- mediana, moda e valore più probabile coincidono con  $\mu$ ;

- come si può facilmente verificare, la larghezza della campana dipende dal valore di  $\sigma$ ; in particolare la *larghezza a metà altezza* (la distanza tra i punti nei quali la funzione assume valore metà del suo valore massimo) è pari a  $2\sigma\sqrt{2\ln(2)} \approx 2.4\sigma$ .

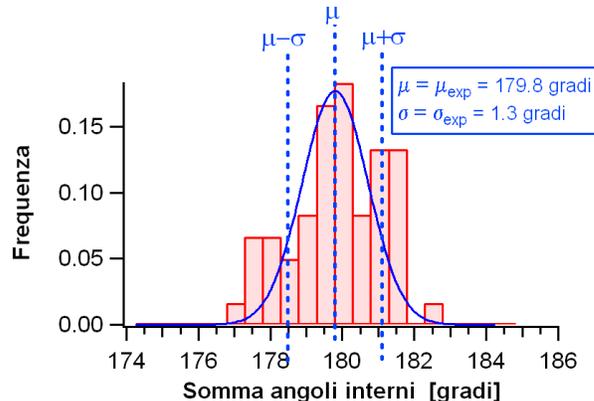


Figura 5. Istogramma delle frequenze con sovrapposta una funzione di distribuzione Gaussiana costruita con i valori di media e deviazione standard trovati dai dati sperimentali (i valori di media e l'intervallo definito dalla deviazione standard sono rappresentati con righe verticali tratteggiate).

Tenendo conto della definizione di probabilità che abbiamo riportato sopra, è facile valutare la probabilità  $P$  di trovare dei valori della misura all'interno di un dato intervallo attorno al valore medio. Per esempio, eseguendo calcoli numerici che si trovano tabulati in apposite tavole, si verifica facilmente che  $P[\mu - \sigma, \mu + \sigma] = \int_{\mu - \sigma}^{\mu + \sigma} f(x)dx = 0.68$ , cioè la probabilità di trovare valori della misura che si discostano dal valore medio per meno di  $\sigma$  (in valore assoluto) è del 68%. Analogamente, si può dimostrare che la probabilità di trovare valori compresi nell'intervallo attorno alla media definito da  $2\sigma$  è del 96%. Questo vuol dire che ben oltre la metà delle misure di un campione fornisce un valore che si discosta (in valore assoluto) dal valore medio per meno di una deviazione standard e quasi tutte le misure si discostano (in valore assoluto) dalla media per meno di due deviazioni standard.

Per rendersi conto di questo e, più in generale, per concludere elegantemente questa nota, si riporta in Fig. 5 lo stesso istogramma di Fig. 4 con sovrapposta l'indicazione della media sperimentale  $\mu_{exp} = 179.8$  gradi e l'intervallo definito dalla deviazione standard  $\sigma_{exp} = 1.3$  gradi, rappresentati con righe verticali tratteggiate, e una funzione Gaussiana, calcolata numericamente secondo l'Eq. 4 usando media e deviazione standard sperimentali. In sostanza, in questo grafico confrontiamo la “forma” della distribuzione (discreta, dunque rappresentata con un istogramma)  $F_k$  con la funzione genitrice  $f(x)$ , continua e descritta con una Gaussiana, dato che ci aspettiamo che

le misure del nostro campione seguano una distribuzione normale (ovvero Gaussiana). Notate che, a rigore, il confronto dovrebbe essere effettuato con la frequenza normalizzata a uno,  $f_k$ , però, dato che la forma dell'istogramma non cambia passando da  $F_k$  a  $f_k = F_k/\Delta_k$ , si può benissimo fare così. L'accordo quantitativo con la  $f(x)$  sarà infatti garantito dal termine che moltiplica l'esponenziale, che in pratica verrà aggiustato in modo da "fittare" (cioè adattarsi) ai dati sperimentali.

Come si può vedere, l'istogramma mostra che la mag-

gior parte dei risultati, cioè delle misure, si trova nell'intervallo attorno alla media definito dalla deviazione standard. Inoltre la funzione Gaussiana, rappresentata con una linea continua di colore blu, riproduce in modo abbastanza soddisfacente i dati sperimentali descritti dall'istogramma. Ci si può aspettare che, aumentando il numero di misure del campione (dedicando più tempo all'esperimento!), l'accordo tra istogramma e funzione Gaussiana tenda a migliorare e che media e deviazione standard sperimentale siano approssimazioni sempre migliori dei corrispondenti valori della funzione genitrice.

- 
- [1] Notate che questa è un'affermazione non del tutto corretta; essa stessa presuppone che lo sperimentatore sia tanto bravo da rendere indipendenti tali misure, cosa non sempre possibile. Ad esempio, nel caso in cui un triangolo sia mal disegnato, con linee storte e spesse, potrebbe verificarsi interdipendenza nella misura dei singoli angoli.
- [2] Notate che abbiamo qui scelto di usare quattro cifre significative; l'operazione aritmetica di media fornisce ulteriori cifre, che vengono qui trascurate perché inutili.
- [3] Osserviamo che in maniera più rigorosa la divisione deve essere fatta per  $N - 1$ , espressione che tiene conto del fatto che il calcolo della media "riduce di uno il numero di gradi

di libertà del campione", secondo quanto stabilito dalla teoria della misura. Tuttavia la differenza tra dividere per  $N$  o per  $N - 1$  è irrilevante già per  $N < 10$ .

- [4] Per completezza, si ricorda che si definisce *varianza* sperimentale  $\sigma^2$  il quadrato della deviazione standard  $\sigma$ .
- [5] Attenzione: questa accuratezza non è interpretabile come incertezza della misura, che è invece data da  $\sigma_{exp}$ !
- [6] Naturalmente il lancio dei dadi segue una distribuzione di probabilità ben diversa da quella delle misure che stiamo qui considerando!