

8 Elementi di Statistica

La conoscenza di alcuni elementi di statistica e di analisi degli errori è importante quando si vogliono realizzare delle osservazioni sperimentali significative, ed anche per poter leggere articoli di optometria pubblicati su riviste italiane e internazionali.

Nessuna misura può essere completamente priva di errore sperimentale. È importante conoscere l'entità di una tale incertezza e saperla ridurre al minimo. Occorre precisare che con la parola "errore" non si intende uno sbaglio nelle procedure effettuate, ma l'inevitabile incertezza connessa con l'uso di determinati strumenti e con la ripetizione della misura. Se ad esempio ripetiamo molte volte l'esperimento di misurare il tempo della caduta di una pallina lungo un piano inclinato con un orologio in grado di calcolare i millesimi, pur con tutte le attenzioni volte ad effettuare le misure nelle condizioni più simili possibili, molto difficilmente otterremo due volte lo stesso identico risultato. Il risultato del nostro esperimento sarà quindi dato dalla media di tutte le misure, che definiremo in seguito, e avrà un'incertezza, che intuitivamente diminuirà al crescere delle prove effettuate. Vediamo ora una formulazione più rigorosa dei concetti appena presentati.

8.1 La media e la deviazione standard

Immaginiamo di voler misurare una certa grandezza x . Non ci preoccupiamo in questa fase di quale grandezza si tratti, e con quale unità di misura debba essere calcolata.

Immaginiamo di aver svolto la misura cinque volte e di avere ottenuto i seguenti risultati

$$51, 52, 52, 53, 51$$

È lecito chiedersi quale sia la migliore stima per la grandezza x . Si può dimostrare che tale valore x_{sti} coincide con la media dei valori misurati, cioè

$$x_{sti} = \bar{x} = \frac{51 + 52 + 52 + 53 + 51}{5} = 51.8$$

Ovviamente tale procedimento può essere effettuato indipendentemente dal numero di prove effettuate. Se facciamo N misure della grandezza x e troviamo N valori

$$x_1, x_2, \dots, x_N$$

la migliore stima per x è ancora la media dei valori trovati che può essere definita come

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_N}{N} = \frac{\sum x_i}{N}$$

La sigma maiuscola Σ è una notazione matematica che corrisponde appunto alla somma dei vari fattori.

Una volta trovato il valore più attendibile per una certa grandezza, è importante capire quanto un tale valore sia affidabile. Una certa informazione ci viene fornita dalla differenza tra il valore di \bar{x} e ognuno dei valori x_i trovati sperimentalmente. Se la differenza è piccola rispetto al valore \bar{x} la misura risulta abbastanza attendibile. Se chiamiamo d_i la differenza $x_i - \bar{x}$, la somma di tutti i valori d_i fornisce sempre il valore 0. Per dare una stima dell'incertezza di una

data misura si ricorre a sommare il quadrato delle differenze d_i per poi farne la radice quadrata dopo aver diviso per il numero di prove effettuate meno uno. La corretta definizione per l'incertezza della misura è quindi

$$\sigma_x = \sqrt{\frac{1}{N-1} \sum (x_i - \bar{x})^2}$$

che per l'esempio precedente viene a valere 0.8.

Molte prove hanno mostrato che tale valore σ , detto "standard deviation" è una buona stima per l'errore sperimentale. Nel dare il valore finale della misura si tende quindi a scrivere che

$$x = \bar{x} \pm \sigma_x$$

Ad esempio nel caso visto prima scriveremmo

$$x = 51.8 \pm 0.8$$

Considerazioni matematiche ci permettono di affermare che se andassimo a ripetere la nostra misura di x , abbiamo circa il sessantotto per cento di probabilità che essa venga a cadere nell'intervallo indicato dalle espressioni precedenti.

8.2 Alcune definizioni

- Popolazione: è il gruppo di individui di interesse in un particolare studio.
Esempio: tutti i portatori di lenti a contatto miopi
- Campione: è un gruppo di individui, selezionato all'interno della popolazione e che può rappresentare la popolazione all'interno del nostro studio
Esempio: 100 portatori di lenti a contatto miopi

8.3 La distribuzione gaussiana

Immaginiamo di dover misurare una determinata grandezza (che potrebbe anche essere ad esempio lo spessore di una lente oftalmica, o il suo peso). Per i motivi precedentemente accennati se ripetiamo molte volte la misura otterremo sempre valori diversi, a causa dell'incertezza della misura stessa. Immaginiamo di aver svolto la misura 30 volte e di rappresentare i valori ottenuti in un istogramma che riporta sull'asse delle ascisse i valori ottenuti suddivisi in intervalli e sull'asse delle ordinate quante volte è stato ottenuto un valore all'interno di quell'intervallo. L'istogramma è riportato in Fig.36. Se poi ripetiamo la misura 100 volte il risultato potrebbe essere quello riportato in Fig.37.

Già dopo 100 misure ci si può accorgere che la distribuzione dei risultati assume una forma interessante, con un picco centrale e le misure che si diradano via via che ci allontaniamo dal centro. Questo tipo di distribuzione risulta ancora più evidente dopo 1000 misure (Fig.38).

Se poi decidiamo di svolgere centomila misure e di creare intervalli molto più piccoli ecco che l'istogramma assume una forma estremamente nota in statistica (Fig.39) e che si definisce una gaussiana. Si può dimostrare che il valore di picco della gaussiana coincide proprio con la media della misura. La gaussiana è una curva che non ha limiti sull'asse x , anche se la probabilità di ottenere una misura via via che ci allontaniamo dal valor medio scende sempre più. Si può dimostrare che circa il 68 % circa delle misure stanno tra $\bar{x} - \sigma_x$ e $\bar{x} + \sigma_x$, mentre il 95 %

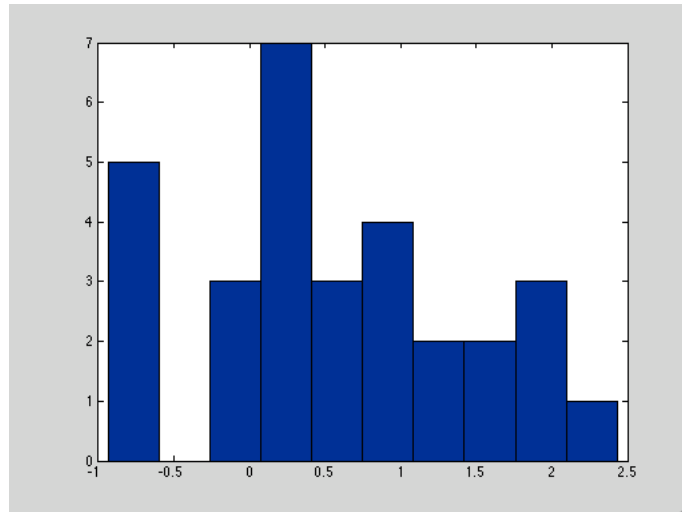


Figura 36: Istogramma di una misura ripetuta 30 volte (si tratta di una misura ipotetica generata al calcolatore con valore atteso zero)

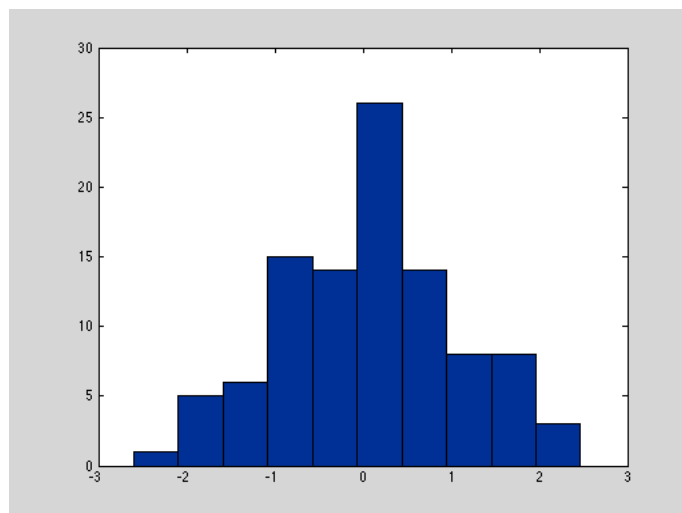


Figura 37: Istogramma di una misura ripetuta 100 volte

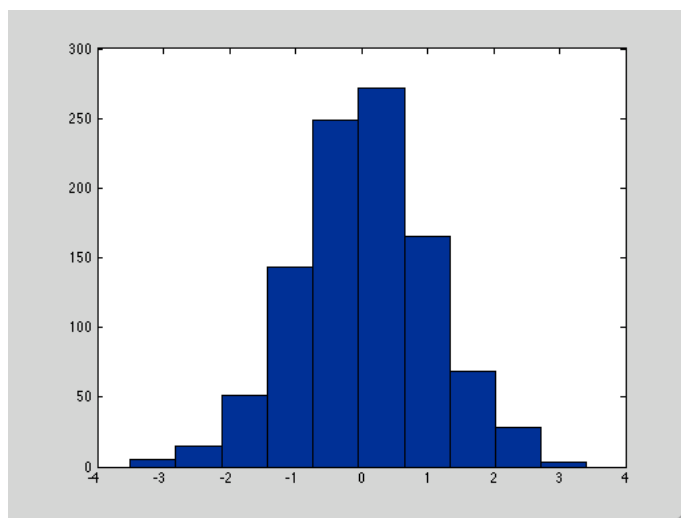


Figura 38: Istogramma di una misura ripetuta 1000 volte

circa stanno tra $\bar{x} - 2\sigma_x$ e $\bar{x} + 2\sigma_x$, e il 99.7% stanno tra $\bar{x} - 3\sigma_x$ e $\bar{x} + 3\sigma_x$. La standard deviation rappresenta in un certo modo la larghezza della gaussiana, quindi una gaussiana larga e bassa indicherà una misura poco precisa, mentre una gaussiana molto stretta indica generalmente una misura precisa.

8.4 Il Valore P

In molti casi statistici ci si può trovare di fronte a un problema di questo tipo: immaginiamo che un ricercatore stia esaminando un farmaco che deve influire sulla pressione oculare. Il metodo da utilizzare è quello di prendere due gruppi di pazienti con la stessa pressione oculare (anche questo concetto di “stessa pressione” andrebbe chiarito), dando a un gruppo un placebo e a un altro gruppo il farmaco da testare. Dopo l’assunzione del farmaco si va ad esaminare la pressione oculare media dei due gruppi. Si otterranno due risultati che saranno certamente diversi, per motivi statistici, ma questa differenza sarà dovuta al farmaco o ad un semplice caso statistico?

In pratica ci troviamo di fronte a due distribuzioni, che daranno luogo a due medie e due standard deviation diverse. Le possibilità che ci interessano sono dunque due

1. Le due popolazioni hanno in realtà la stessa media cioè le due distribuzioni derivano dalla stessa situazione e la differenza è dovuta esclusivamente al caso statistico: questo significa che il farmaco in realtà non ha avuto alcun effetto. È la cosiddetta ipotesi nulla
2. Le due popolazioni hanno medie differenti, cioè le due distribuzioni nascono da un’effettiva differenza tra i due campioni: questo significa che il farmaco ha avuto effetto. È la cosiddetta ipotesi alternativa.

Ovviamente rigetteremo l’ipotesi nulla solo quando ci troveremo di fronte ad un’evidenza che il secondo campione non coincide con il primo. Per esemplificare

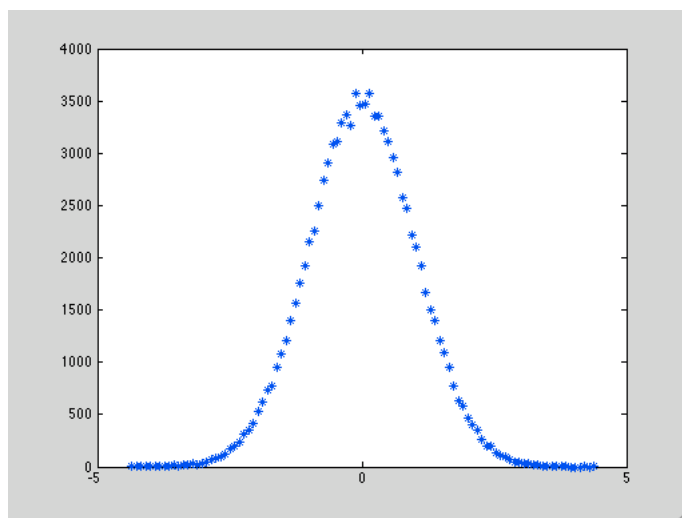


Figura 39: Istogramma di una misura ripetuta 1000 volte

chiamiamo D_1 (D_2) la distribuzione corrispondente al gruppo che ha ricevuto il placebo (farmaco). La loro pressione media oculare sarà P_1 (P_2) con un errore ΔP_1 (ΔP_2). Se la probabilità di ottenere il risultato P_2 nel caso della distribuzione D_1 è sufficientemente piccola, allora potremo rigettare l'ipotesi nulla e accettare l'ipotesi alternativa. Resta da chiedersi che cosa si intenda con probabilità sufficientemente piccola; nella maggior parte dei casi si sceglie lo 0,05. Per essere ancora più certi della falsità dell'ipotesi nulla si può scegliere lo 0,01. Se si è scelto il valore 0,05 questo significa che rigetteremo l'ipotesi nulla se, essendo vera la distribuzione D_1 , avremo meno del 5% di probabilità di ottenere il valore P_2 .

In pratica il valore P è la risposta alla domanda:

se le due popolazioni hanno in realtà la stessa media, quale è la probabilità che il caso statistico abbia condotto ad una differenza tra le medie quale quella che abbiamo ottenuto o addirittura più grande?

Per questo motivo se otteniamo un valore P di 0,35 questo significa che nel 35% dei casi noi avremmo potuto ottenere una differenza tra i risultati di tale entità o anche maggiore, nonostante che in realtà le due popolazioni abbiano la stessa media. In questo caso il farmaco non ha fatto effetto. Se invece il valore P è 0,001 allora una tale differenza di risultati sarebbe stata possibile solo nello 0,1% dei casi se non ci fosse differenza tra le popolazioni. Questo significa che le due popolazioni sono diverse: il farmaco ha fatto effetto.

Vediamo ora alcuni esempi

1. Un ricercatore misura lo spessore del film lacrimale degli abitanti di Sovigliana e poi lo spessore del film lacrimale degli abitanti di Vinci, e confronta poi le due distribuzioni così ottenute. Egli ottiene un valore $p=0,02$. Che conseguenza può trarre da questo risultato?

Risposta: se egli ha posto il livello di significatività della sua misura a 0,05, il valore P è risultato minore. Questo significa che l'ipotesi nulla è falsa.

In pratica, questo significa che abitare a Vinci o a Sovigliana influisce sullo spessore del film lacrimale dei soggetti. Se il valore P fosse risultato essere 0,58 allora egli avrebbe tratto la conseguenza che non vi è differenza tra il vivere nei due paesi per ciò che riguarda lo spessore del film lacrimale.

2. Lo spessore del film lacrimale della popolazione italiana è 211 micron con una standard deviation di 46 micron. Prendiamo un campione di 12 persone e gli somministriamo un farmaco che pensiamo abbia l'effetto di cambiare lo spessore del film lacrimale. In questo campione otteniamo una media di 217 micron. Ci chiediamo se questo dato sperimentale ci autorizza a pensare che il farmaco ha avuto effetto.

Risposta: Per prima cosa è necessario sapere a quante standard deviation si trova il risultato trovato sul campione dalla media generale. Per questo motivo ci calcoliamo la standard deviation della media.

$$\sigma_x = \frac{\sigma_x}{\sqrt{N}} = \frac{46}{\sqrt{12}} \approx 13.3$$

A questo punto possiamo calcolarci quante standard deviation ci sono tra le due medie

$$z = \frac{217 - 211}{13.3} = 0.45$$

Ora possiamo andare a vedere sulle apposite tabelle quale la probabilità di ottenere un tale risultato (le tabelle sono rintracciabili in vari siti web, come ad esempio [?] o sui testi specifici, come [?]. Grazie a tali tabelle si può vedere che la probabilità che la media caschi all'interno di 0.45 standard deviation è 34.73%. Questo significa che la probabilità che il risultato fosse a una distanza maggiore è di 65.27%, cioè un valore P di circa 0.65. Si noti che così facendo noi abbiamo considerato sia il caso di uno spessore più alto della media che quello di uno spessore più basso della media, dato che non indicavamo l'effetto del farmaco: si parla di test di ipotesi bilaterali. Dato l'elevato valore P non possiamo escludere l'ipotesi nulla, e dobbiamo concludere che il farmaco non ha avuto effetto.

3. Vediamo un esempio preso da un articolo di una rivista di optometria [?]. Nell'articolo sono esaminati 239 soggetti, 184 portanti lenti a contatto e 55 no. Si è chiesto ai soggetti se hanno mai avuto sintomi tipici dell'occhio secco. 151 hanno risposto no (63%) e 88 si (37%). Se però esaminiamo i soli non portatori il risultato è 48 no (87%) 7 si (13%), mentre tra i portatori 103 no (56%) e 81 si (44%). Questa differenza tra portatori e non portatori è un caso statistico, o corrisponde a un'effettiva differenza? Guillon calcola il valore P ottenendo $p = 0.001$ il che conferma che siamo di fronte a una differenza reale, cioè ad un effetto reale delle lenti a contatto.