Probabilità, statistica e false credenze

Marco Caliari
Dipartimento di Informatica
Università di Verona

Simone Zuccher
Liceo Scientifico Statale "E. Medi"
Villafranca di Verona

Piano Lauree Scientifiche, a.s. 2010–2011

Capitolo 1

Errori, media, scarto, deviazione standard

Il compito essenziale della fisica è la scoperta e l'uso di relazioni tra osservazioni quantitative di fenomeni naturali.

1.1 PowerBalance

Proiezione video PowerBalance. Test della rotazione.

1.2 Pendolo

Misura di cinque periodi di un pendolo, per un totale di 60 misure, con cronometro a centesimi di secondo.

1.3 Errori

Purtroppo, le osservazioni sperimentali contengono sempre qualche inesattezza. Gli errori possono essere classificati in

sistematici, se dipendono dai particolari strumenti e tecniche di misura usati;

accidentali (o casuali), se dipendono dalle incontrollabili e sconosciute variazioni delle condizioni sperimentali.

Esistono poi gli sbagli di trascrizione, le sviste nel leggere gli strumenti e le distrazioni nell'eseguire i calcoli. Questi ultimi si possono eliminare, mentre gli errori sistematici o accidentali sono ineliminabili.

1.4 Media e scarto

Supponiamo di eseguire dieci misure diverse di una certa quantità x e di ottenere dieci valori eventualmente diversi $x_1, x_2, \ldots, x_{10}$. Calcolare la *media*

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_{10}}{10}$$

appare intuitivamente come un procedimento logico, anche se non c'è nessuna garanzia, ovviamente, che la media $\bar{x}$ sia uguale al valore vero x. Di più, potrebbe succedere che una o più delle misure x_n sia migliore della media, nel senso che l'errore $|x_n - x|$ è minore di $|\bar{x} - x|$. Per ogni misura x_i, chiamiamo *scarto* la quantità

$$d_n = x_n - \bar{x}$$

1.5 Deviazione standard

Supponiamo di misurare il peso di un telefono cellulare. Affidiamo a due persone diverse il compito e chiediamo loro di effettuare tre pesate. La prima persona ottiene i valori 70.5, 70.6 e 71.9 grammi e la seconda 52.1, 69.9 e 91.0 grammi. Di quale persona vi fidereste di più? Eppure, in entrambi i casi la media dei valori vale 71 grammi. Serve allora una misura di quanto siano disperse le misure. Si vede subito come nel primo caso, gli scarti siano mediamente più piccoli che nel secondo. In generale, date N misure, potremmo considerare la media degli scarti

$$\frac{1}{N}\sum_{n=1}^{N} d_n$$

Con semplici manipolazioni, si trova

$$\begin{aligned}
\frac{1}{N}\sum_{n=1}^{N} d_n &= \frac{1}{N}\sum_{n=1}^{N}(x_n - \bar{x}) = \frac{1}{N}\left(\sum_{n=1}^{N} x_n - \sum_{n=1}^{N} \bar{x}\right) = \\
&= \frac{1}{N}\left(\sum_{n=1}^{N} x_n - N\bar{x}\right) = \frac{1}{N}\left(\sum_{n=1}^{N} x_i - N\frac{1}{N}\sum_{n=1}^{N} x_n\right) = \\
&= \frac{1}{N}\left(\sum_{n=1}^{N} x_n - \sum_{n=1}^{N} x_n\right) = 0
\end{aligned}$$

Dunque, la media degli scarti è sempre zero e pertanto risulta essere poco significativa. Come indice della dispersione delle misure, si usa invece la *deviazione standard*, definita come

$$\sigma = \sqrt{\frac{1}{N}\sum_{n=1}^{N} d_n^2}$$

Il quadrato della deviazione standard, σ^2, è detto *varianza*. Talvolta la deviazione standard viene definita come

$$\sigma = \sqrt{\frac{1}{N}\sum_{n=1}^{N} x_n^2 - \left(\frac{1}{N}\sum_{n=1}^{N} x_n\right)^2}$$

Le due definizioni sono equivalenti: infatti

$$\begin{aligned}
\frac{1}{N}\sum_{n=1}^{N} d_n^2 &= \frac{1}{N}\sum_{n=1}^{N}(x_n - \bar{x})^2 = \frac{1}{N}\sum_{n=1}^{N}(x_n^2 - 2x_n\bar{x} + \bar{x}^2) = \\
&= \frac{1}{N}\sum_{n=1}^{N} x_n^2 - \frac{1}{N}\sum_{n=1}^{N} 2x_n\bar{x} + \bar{x}^2 = \frac{1}{N}\sum_{n=1}^{N} x_n^2 - 2\bar{x}\frac{1}{N}\sum_{n=1}^{N} x_n + \bar{x}^2 = \\
&= \frac{1}{N}\sum_{n=1}^{N} x_n^2 - 2\bar{x}^2 + \bar{x}^2 = \frac{1}{N}\sum_{n=1}^{N} x_n^2 - \bar{x}^2 = \\
&= \frac{1}{N}\sum_{n=1}^{N} x_n^2 - \left(\frac{1}{N}\sum_{n=1}^{N} x_n\right)^2
\end{aligned}$$

Con riferimento all'esempio precedente, nel primo caso la deviazione standard vale circa 0.64, mentre nel secondo circa 15.9.

1.6 Esercizi

1. Calcolare la media delle misure del periodo del pendolo. Verificare che la media degli scarti è nulla.

2. Calcolare la deviazione standard delle misure del periodo del pendolo (senza l'apposita funzione).

Capitolo 2

La probabilità

È molto probabile che la maggior parte della gente non sappia niente di probabilità.

2.1 Il concetto di probabilità

Se lanciamo una moneta in aria, intuitivamente diciamo che la probabilità che esca testa è del 50%. Cosa significa? Potremmo dire che ci possono essere due risultati possibili, testa o croce e che, non avendo motivo di ritenere la moneta truccata, i due eventi ("esce testa", "esce croce") sono equiprobabili, da cui il 50% di probabilità. Cioè, per definire il concetto di "probabilità" si è fatto ricorso al concetto di "equiprobabilità": un circolo vizioso. Convince di più questa spiegazione: se lanciamo la moneta un numero molto grande di volte, il numero di volte che uscirà testa sarà *all'incirca* uguale alla metà del numero totale. Allo stesso modo, lanciando un dado molte volte, il numero 5 uscirà all'incirca un sesto delle volte e dunque la sua probabilità di uscita è 1/6. Per convenzione, si dà probabilità 1 (o 100%) all'evento *certo*. Se la probabilità di un evento è p, la probabilità che non accada è $1 - p$.

2.1.1 Cosa non dice la probabilità

Supponiamo di lanciare 100 volte una moneta e di ottenere 52 volte testa. Il rapporto 52/100 vale 0.52, numero ragionevolmente vicino alla probabilità di uscita di testa (0.5). Supponiamo ora di proseguire con il lancio fino a 10000 volte e di ottenere 5020 volte testa. Il rapporto 5020/10000 è ancora più vicino (come ci si aspetta all'aumentare dei lanci) a 0.5. Supponiamo ora di giocare d'azzardo e di ricevere 100 Euro se esce croce e di perdere 100 Euro se esce testa. Ebbene, con l'esempio precedente, dopo i primi 100 lanci si sarebbe in perdita di 400 Euro. Continuando a giocare, con l'idea (sbagliata!) di rifarsi perché il numero di uscite di croci dovrebbe essere circa uguale al numero di uscite di teste, si arriverebbe a 10000 lanci in perdita di 4000 Euro. E nulla vieta che la perdita possa continuare ad aumentare.

Altra miscredenza piuttosto comune è la seguente: supponiamo di sapere che una moneta non truccata è stata lanciata 20 volte e ha sempre dato testa. Ci viene chiesto di scommettere 200 Euro che al lancio successivo esce croce: voi scommettereste? Ora supponiamo di non sapere cosa è uscito nei 20 lanci precedenti: scommettereste 200 Euro sull'uscita della croce? Cosa cambia, per la moneta, nei due casi? Niente. Perché allora tenere nota dei numeri usciti al Lotto e scommettere sui "ritardatari"?

2.2 Somme e prodotti di probabilità

Supponiamo di lanciare due dadi. Qual è la probabilità di ottenere 8 come somma? Ci sono varie possibilità (2 e 6, 3 e 5, ...) e, poiché ognuna di queste *esclude* le altre, la probabilità complessiva è data dalla *somma* delle probabilità di ogni possibilità (che vale $1/36$). Dunque, la probabilità di ottenere 8 è $5/36$.

Chiediamoci ora qual è la probabilità di ottenere 8 come somma da due dadi e testa dal lancio di una moneta. Poiché il risultato del lancio della moneta è *indipendente* dal lancio dei dati e poiché appare equamente probabile la somma 8 e l'uscita di testa o la somma 8 e l'uscita di croce, la probabilità cercata è data dal prodotto $5/36 \cdot 1/2$. È evidente che in presenza di una moneta truccata (o, semplicemente sbilanciata) che dia testa il $2/3$ delle volte, la probabilità cercata è data dal prodotto $5/36 \cdot 2/3$.

A volte si cerca la probabilità che accadano eventi non indipendenti tra loro. Per esempio, qual è la probabilità di ottenere 8 dalla somma di due dati *e* aver ottenuto 5 dal primo dado? Nel caso particolare, è ovvio che il secondo dado deve aver dato 3. Dunque la probabilità cercata è $1/6 \cdot 1/6$. Nel caso generale, si può ragionare così: è la probabilità di ottenere 8 *sapendo* che il primo dado ha dato 5 ($1/6$) moltiplicato per la probabilità che il primo dado abbia dato 5 ($1/6$).

2.3 Palline

Supponiamo che dentro un'urna ci siano sei palline bianche e quattro nere. Qual è la probabilità che estraendone due siano una bianca e una nera? Ci sono due possibilità: la prima pallina estratta è bianca e la seconda nera o viceversa. Dunque, $6/10 \cdot 4/9 + 4/10 \cdot 6/9$ (una volta estratta la prima pallina, rimangono nove palline nell'urna). Allo stesso modo si può calcolare la probabilità di fare sei al Superenalotto. Scelti sei numeri tra novanta, possiamo immaginare che le sei palline corrispondenti ai numeri scelti siano nere e le altre bianche. Qual è la probabilità che vengano estratte sei palline nere? Evidentemente è $6/90 \cdot 5/89 \cdot 4/88 \cdot 3/87 \cdot 2/86 \cdot 1/85$ (per inciso, vale $1/622614630$). E la probabilità di fare cinque (avendo giocato sei numeri)? Si tratta di calcolare la probabilità che vengano estratte cinque palline nere e una bianca. Ci sono sei possibilità, a seconda che la pallina bianca venga estratta per prima, per seconda e così via. In conclusione, la probabilità cercata è $6 \cdot 84/90 \cdot 6/89 \cdot 5/88 \cdot 4/87 \cdot 3/86 \cdot 2/85$ (cioè $25/30883662$).

Anche nel gioco del Superenalotto esistono false credenze. Per esempio, è più probabile che escano i numeri 3, 2, 1, 5, 4, 6 oppure i numeri 3, 2, 9, 17, 33, 5? Ovviamente la probabilità è la stessa. Perché allora il primo gruppo di numeri ci sembra molto meno probabile? Perché tendiamo a dare un senso alle cose e il primo gruppo ci appare più sensato del secondo. E poichè sappiamo dare un senso a molti meno gruppi di quelli a cui non sappiamo dare un senso e confondiamo la probabilià che esca un particolare gruppo sensato con la probabilità che esca uno *qualunque* di questi, reputiamo minore la probabilità del primo gruppo. D'altra i numeri del secondo gruppo, se ordinati, sono a distanza 1,2,4,8,16 (e dunque sono "sensati").

2.4 Permutazioni e combinazioni

In quanti modi si possono ordinare quattro palline numerate? Evidentemente

①②③④	②①③④	③①②④	④①②③
①②④③	②①④③	③①④②	④①③②
①③②④	②③①④	③②①④	④②①③
①③④②	②③④①	③②④①	④②③①
①④②③	②④①③	③④①②	④③①②
①④③②	②④③①	③④②①	④③②①

La prima pallina può essere scelta tra quattro, la seconda tra le rimanenti tre e la terza tra le rimanenti due. In totale ci sono $4 \cdot 3 \cdot 2 = 24$ modi di ordinare le palline. In generale, il numero di *permutazioni* di N oggetti vale $N \cdot (N-1) \cdot \ldots \cdot 2$ e si scrive $N!$ (N fattoriale).

Consideriamo ora il problema di scegliere quattro palline da un gruppo di dieci palline. In quanti modi diversi può avvenire la scelta? La prima può essere scelta tra dieci, la seconda tra nove e così via. Si arriva a $10 \cdot 9 \cdot 8 \cdot 7$. Se quelle estratte sono le palline, per esempio, ①, ②, ③, ④ (proprio in quest'ordine), è chiaro che esistono altri 23 modi per scegliere le stesse palline. Quindi, per ogni insieme di quattro palline esistono $4 \cdot 3 \cdot 2$ modi di estrarle. Allora, la risposta alla domanda "in quanti modi diversi posso scegliere quattro palline *senza tener conto dell'ordine* da un gruppo di dieci palline" è $10{\cdot}9{\cdot}8{\cdot}7/(4{\cdot}3{\cdot}2{\cdot}1)$ (si mette $\cdot 1$ per avere lo stesso numero di fattori a numeratore e a denominatore). In generale, il numero di *combinazioni* di N oggetti presi a gruppi di n $(0 \leq n \leq N)$ è

$$\frac{N \cdot (N-1) \cdot \ldots \cdot (N-n+1)}{n!} = \frac{N!}{(N-n)!n!} = \binom{N}{n}$$

Evidentemente, c'è un solo gruppo di N oggetti: dunque $\binom{N}{N} = 1$, da cui, per definizione, $0! = 1$. Il simbolo $\binom{N}{n}$ si chiama coefficiente binomiale di N e n. Il nome deriva dal calcolo della potenza N-esima di un binomio

$$(a+b)^N = \sum_{n=0}^{N} \binom{N}{n} a^n b^{N-n}$$

2.5 Esercizi

1. Qual è la probabilità che in un gruppo di N persone ce ne siano alcune che compiono gli anni nello stesso giorno e mese?

2. Calcolare la probabilità di fare n, $0 \leq n \leq 6$, al Superenalotto avendo giocato m numeri, $6 \leq m \leq 8$.

Capitolo 3

Distribuzione di probabilità

Per una più dettagliata analisi del concetto di probabilità abbiamo bisogno di considerare metodi più efficienti per trattare le probabilità di intere classi di eventi.

3.1 Distribuzioni di probabilità discrete

Ritornando all'esempio del Superenalotto, possiamo definire la funzione

$$f(n) = \frac{\binom{6}{n} \cdot \binom{84}{6-n}}{\binom{90}{6}}$$

che fornisce la probabilità di fare n avendo giocato sei numeri. Possiamo fare un istogramma di $f(n)$ che mostra chiaramente quanto sia difficile fare più di due. Una funzione che assume certi valori x_n con una probabilità nota p_n si chiama *variabile aleatoria (discreta)* e la funzione $f(n) = p_n$ è la sua *distribuzione di probabilità*. Una variabile aleatoria può essere indicata con la notazione

$$\begin{cases} x_1 & x_2 & \dots & x_N \\ p_1 & p_2 & \dots & p_N \end{cases}$$

Naturalmente in una distribuzione di probabilità deve succedere che

$$\sum_n p_n = \sum_n f(n) = 1$$

ove la sommatoria si intende su tutti i valori possibili di n. Si può definire anche la *media* (o valore medio o valore atteso) di una variabile aleatoria che ha un significato analogo al concetto di media già visto. Si suppone infatti di ripetere molte volte, diciamo Z, un esperimento che ammetta come risultati i

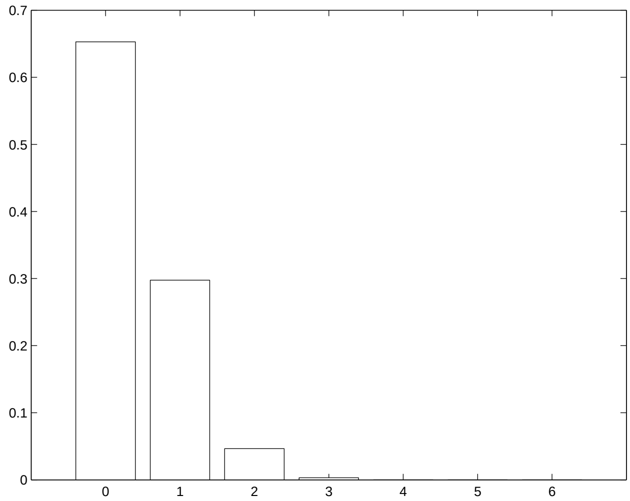

Figura 3.1: Probabilità di fare $0, 1, 2, 3, 4, 5, 6$ giocando sei numeri al Superenalotto.

valori $x_1, x_2, \ldots, x_N$ (per esempio una misura, oppure il lancio di un dado), di sommare tutti i valori che si sono trovati e dividere per Z. Si dovrà eseguire un'operazione del tipo

$$\bar{x} = \frac{\overbrace{x_2 + x_5 + x_4 + x_2 + x_1 + \ldots + x_6}^{Z \text{ valori}}}{Z}$$

Si tratta ora di calcolare il numeratore di tale espressione. Se la probabilità che esca x_n è p_n, per la definizione che abbiamo dato di probabilità significa che ripetendo molte volte un esperimento il numero di volte che esce x_n rispetto al totale Z è p_n. Dunque il numero di volte che esce x_n è Zp_n. Quindi, raccogliendo i vari x_n,

$$\bar{x} = \frac{x_1 \cdot Zp_1 + x_2 \cdot Zp_2 + \ldots + x_N \cdot Zp_N}{Z} = \frac{Z \cdot \sum_n x_n p_n}{Z} = \sum_n x_n p_n$$

È da notare che $\bar{x}$ può anche non essere un valore possibile della variabile aleatoria: per esempio, il valore medio del lancio di un dado è 3.5. Si può anche calcolare il valor medio di $(x_n - \bar{x})^2$, cioè la varianza dei valori x_n, dato da

$$\sigma^2 = \sum_n (x_n - \bar{x})^2 p_n$$

3.2 La distribuzione binomiale

Un caso molto frequente è quello di ripetere più volte (ma non necessariamente tante) uno stesso esperimento (ogni volta *indipendentemente* dall'altra, come misurare una quantità, lanciare dieci volte una moneta) e chiedersi con quale probabilità si hanno i vari risultati possibili. Cominciamo proprio con il lancio di una moneta dieci volte. Con quale probabilità testa esce zero volte? Evidentemente deve uscire croce dieci volte. Poiché ogni lancio è indipendente dagli altri, tale probabilità vale $(1/2)^{10}$. Con quale probabilità esce testa una sola volta? Ci sono varie possibilità: testa esce solo la prima volta, testa esce solo la seconda e così via. Ognuno di questi eventi ha probabilità $(1/2) \cdot (1/2)^9$ (cioè la probabilità che esca testa una volta e croce nove volte), ognuno di questi eventi esclude gli altri e il numero di questi eventi è 10: dunque, la probabilità cercata è $10 \cdot (1/2)^{10}$. Con quale probabilità esce testa n volte ($0 \le n \le 10$)? Ognuno di questi eventi ha probabilità $(1/2)^n \cdot (1/2)^{10-n}$, ognuno di questi eventi esclude gli altri e il numero di questi eventi è $\binom{10}{n}$. Dunque la distribuzione cercata è

$$f(n) = \binom{10}{n} \left(\frac{1}{2}\right)^n \left(\frac{1}{2}\right)^{10-n}$$

Più in generale, un esperimento

$$\begin{cases} 0 & 1 \\ 1-p & p \end{cases}$$

che ha successo con probabilità p ripetuto N volte, fornirà un numero di successi distribuito come

$$f(n) = \binom{N}{n} p^n (1-p)^{N-n}$$

Una variabile aleatoria con tale distribuzione si chiama variabile aleatoria binomiale.

Abbiamo detto che deve necessariamente essere $\sum_n f(n) = 1$. È vero per la variabile aleatoria binomiale? Si ha

$$\sum_{n=0}^{N} f(n) = \sum_{n=0}^{N} \binom{N}{n} p^n (1-p)^{N-n} = [p + (1-p)]^N = 1$$

Dunque è vero. Qual è il numero medio di successi? Dobbiamo pensare di eseguire tante volte, diciamo Z, la ripetizione di N esperimenti, sommare il

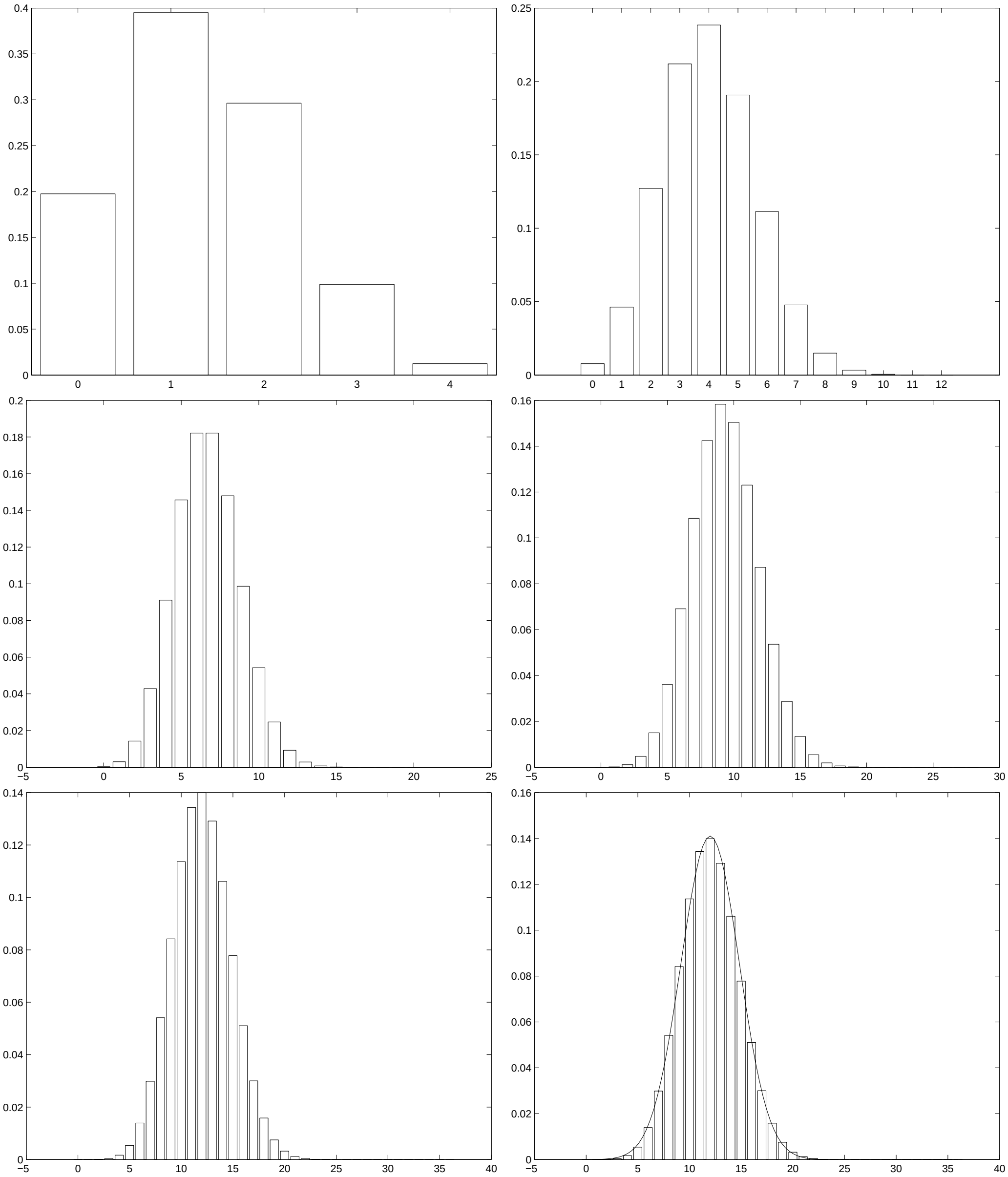

Figura 3.2: Distribuzione binomiale con $p = 1/2$ e $N = 4, 12, 20, 28, 36$ e campana di Gauss.

numero di successi ogni volta ottenuto e dividere per Z. Come visto sopra, si ottiene

$$\bar{n} = \sum_{n=0}^{N} n \binom{N}{n} p^n (1-p)^{N-n}$$

Si può dimostrare che

$$\sum_{n=0}^{N} n \binom{N}{n} p^n (1-p)^{N-n} = Np$$

Tale risultato non dovrebbe sorprendere. Consideriamo ancora il lancio di monete. La probabilità che esca testa è $p = 1/2$ e questo significa che se lanciamo la moneta tante volte, diciamo $Z \cdot N$, il numero di teste che ci aspettiamo è $Z \cdot N \cdot p$. Associamo $x_2 = 1$ a questo evento e $x_1 = 0$ all'uscita di croce. Possiamo pensare di aver lanciato le monete a gruppi di N lanci. E poiché questi gruppi sono Z, la media dei risultati di ogni gruppo è $(x_1 \cdot ZN(1-p) + x_2 \cdot ZNp)/Z = Np$. Questa non è una vera dimostrazione, ma una giustificazione della sensatezza del risultato.

3.3 La distribuzione di Gauss o "legge normale degli errori"

Abbiamo già visto che quando si misura una certa quantità, si commettono degli errori. Sembrerebbe sensato allora calcolare la distribuzione di tali risultati, mettendo in un grafico i valori trovati nelle ascisse e il numero di volte che tale valore compare (rispetto al totale) nelle ordinate. In generale, il risultato di una misura non è un numero naturale (cioè $0, 1, 2, \ldots$) e quindi è molto probabile che ogni valore misurato compaia solamente una volta nella lista dei valori (e ciò è tanto più vero quanto più la misura è accurata, in termini per esempio di cifre decimali riportate). Conviene allora fare in questo modo: nell'insieme delle misure $\{x_n\}$, si identificano il valore minimo $x_{\min}$ e il valore massimo $x_{\max}$, si divide l'intervallo $[x_{\min}, x_{\max}]$ in n parti (solitamente n dispari) di lunghezza uguale $h = (x_{\max} - x_{\min})/n$ e si contano quanti valori rientrano in ogni intervallo $[x_{\min}+(i-1)h, x_{\min}+ih]$, $i = 1, \ldots, n$. A questo punto, si possono mettere in un grafico i centri degli intervalli nelle ascisse e il numero di valori che cadono in quell'intervallo rispetto al totale, divisi per h (per questioni di normalizzazione). Il risultato è simile a quello in Figura 3.3, in cui si è misurato (molto grossolanamente) fino a 1000 volte una quantità di misura teorica 10 e si sono divisi i risultati in 21 intervalli. La curva disegnata nel grafico si chiama *curva (o campana) di Gauss*. Essa può essere riguardata come un risultato analitico derivato da elementari considerazioni matematiche, o come una formula empirica che si è trovato essere in accordo con la distribuzione degli errori casuali che realmente intervengono in una data misura. Qualcuno ha osservato a questo proposito che tutti sono convinti che la curva di Gauss descriva fedelmente il

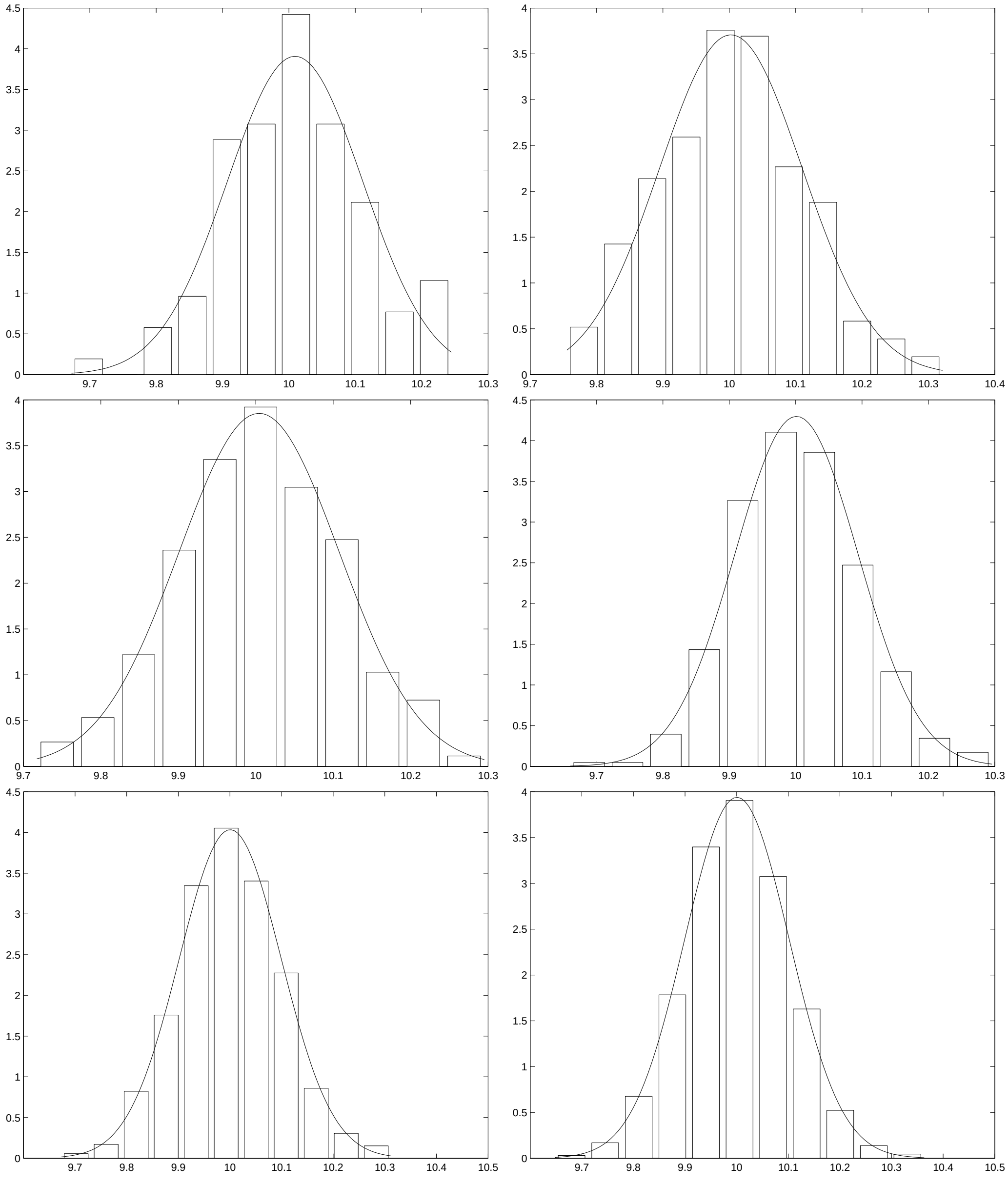

Figura 3.3: Distribuzione di $N = 100, 300, 500, 700, 900, 1000$ misure e corrispondente campana di Gauss.

comportamento degli errori casuali: i matematici perché credono che i fisici l'abbiano verificata sperimentalmente, i fisici perché credono che i matematici l'abbiano dimostrata su rigorose basi teoriche.

Da un punto di vista teorico si può fare la ragionevole affermazione che ogni errore accidentale si può pensare come il risultato di un gran numero di errori elementari, tutti di uguale entità, e ciascuno con uguale probabilità di produrre una variazione in eccesso o in difetto. La curva di Gauss può così essere considerata come una forma limite della distribuzione binomiale, quando in questa ultima il numero N degli eventi indipendenti (corrispondenti agli errori elementari) diventa elevato, mentre la probabilità p di successo di ogni prova (la probabilità per ogni errore elementare di essere positivo) è $1/2$. La formula della curva di Gauss è

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

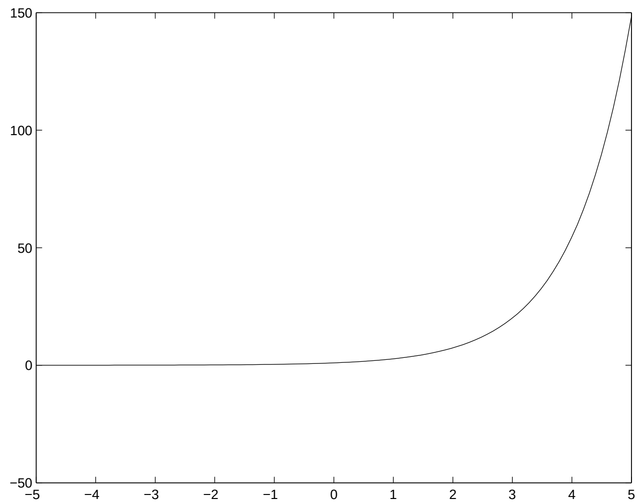

Figura 3.4: Grafico di $y = \exp(x)$.

Il grafico della funzione $y = \exp(x)$ è riportato in Figura 3.4 e in particolare, $\exp(0) = 1$. Ad ogni modo, la funzione exp è implementata in qualunque sistema di calcolo. Pertanto, per disegnare la curva gaussiana, è sufficiente conoscere il valori μ e σ, cioè la sua media e la sua deviazione standard. Riassumendo, quando si effettuano diverse misure indipendenti di una quantità, si ottengono valori che si distribuiscono come una curva di Gauss. Tante più misure si fanno e tanto meglio sarà la corrispondenza con la curva. La curva dipende però dai due parametri μ e σ che sono la media e la deviazione standard di un altissimo numero di misure. In pratica, però, si conoscono solo i valori x_n del *campione* di misure (diciamo di lunghezza N) che si sono effettuate. Come calcolare allora μ e σ? Per quanto riguarda μ, una buona stima viene dalla media del campione

$$\mu \approx \bar{x} = \sum_{n=1}^{N} x_n$$

Per quanto riguarda la σ, si potrebbe usare la deviazione standard del campione. Ma risulta più accurata la quantità

$$\sigma \approx \sqrt{\frac{1}{N-1}\sum_{n=1}^{N} d_n^2}$$

tanto che in alcuni sistemi di calcolo la deviazione standard di un insieme di valori è definita proprio in questo modo (cioè con il denominatore $N-1$ invece di N). Una giustificazione potrebbe essere la seguente: supponiamo di effettuare una sola misura x_1 (e dunque $N = 1$). Ha senso parlare di dispersione dei dati? Certamente no. Ma se si usa la definizione con N a denominatore si ottiene $\sigma = 0$ (quasi a dire che i dati non sono dispersi), mentre con $N-1$ a denominatore si ottiene una forma indeterminata $0/0$, quasi a dire, appunto, che non ha senso parlare di dispersione.

3.4 Esercizi

1. Utilizzando i dati raccolti nell'esperimento del pendolo e reperibili qui `http://profs.sci.univr.it/~zuccher/downloads/datipendolo.xls`, suddividere le misure tra il minimo ed il massimo valore in 5 intervalli ciascuno di ampiezza h e determinare l'istogramma delle frequenze (ossia del numero di volte che i valori compaiono all'interno di uno dei 5 range).

2. Dall'istogramma delle frequenze si ottenga quello della distribuzione di probabilità (*suggerimento: si divida la frequenza per l'ampiezza dell'intervallo h e per il numero totale di misure*). L'istogramma è come ci si aspettava? In ogni caso si giustifichi adeguatamente la risposta. Se si aumenta il numero di intervalli da 5 a 7 o 9 (o in generale a $N > 5$), come cambiano le cose (se cambiano)?

3. Si calcoli la probabilità che una misura "cada" nel secondo e nel terzo intervallo (*suggerimento: la probabilità che una misura stia tra x_1 e x_2 è l'area che sta sotto la curva di distribuzione di probabilità tra x_1 e x_2*).

Capitolo 4

Esperimento in doppio cieco

4.1 Protocollo

Lo studente più anziano è il giudice, un docente il suo segretario. Il giudice mette il PowerBalance e l'altro braccialetto dentro due sacchetti di stoffa indistinguibili, davanti a tutti gli studenti. Lo studente più giovane esce dalla stanza con i due sacchetti e un sacchetto di carta. Chiude la porta, mette i sacchetti di stoffa dentro il sacchetto di carta e lo chiude piegandone la sommità. A questo punto bussa alla porta. Il giudice apre la porta, prende il sacchetto di carta e richiude la porta. Il giudice estrae i sacchetti di stoffa e li posiziona sulla sua postazione, dentro usca scatola di scarpe, sopra due fogli di carta recanti le scritte "A" e "B", rispettivamente. Il giudice richiama lo studente fuori dall'aula. A questo punto, nessuna delle persone presenti in aula conosce il contenuto dei sacchetti di stoffa. Comincia l'esperimento: lo studente esegue una torsione con il braccio destro teso e il palmo rivolto verso l'alto. Il giudice procede alla generazione casuale di una lettera "A" o "B" tramite l'apposito software. Dopo l'uscita della lettera, il segretario prende il sacchetto di stoffa corrispondente e lo pone sul palmo aperto dello studente di turno, senza che lo studente sappia quale sacchetto è stato sorteggiato. Lo studente esegue la prova di torsione e il giudice e il segretario annotano il risultato. Il sacchetto viene cambiato e lo studente esegue nuovamente la prova e il giudice e il segretario annotano il risultato. A questo punto aggiornano il foglio dei risultati (vedi allegato) mettendo il nome dello studente, la lettera uscita, e segnando la casella corrispondente al risultato ottenuto dalle due torsioni ($\theta_A \leq \theta_B$ oppure $\theta_A > \theta_B$). Alla fine dell'esperimento, il segretario apre i sacchetti di stoffa, svelando così in quale si trova il PowerBalance. Il giudice e il segretario firmano il foglio dei risultati *senza* calcolare i totali. I totali saranno calcolati successivamente.

4.2 E se PowerBalance non funziona?

Proviamo a ripetere mentalmente l'esperimento fatto, *supponendo* di avere usato un PowerBalance difettoso che non produce effetti (o, meglio, che produce gli stessi effetti di un qualunque altro braccialetto). Una prima variabile aleatoria genera le sequenze AB o BA in maniera casuale con probabilità 1/2 ciascuna, cioè

$$\begin{cases} \text{AB} & \text{BA} \\ \frac{1}{2} & \frac{1}{2} \end{cases}$$

e una seconda variabile aleatoria genera $\theta_A \leq \theta_B$, $\theta_A > \theta_B$ con probabilità $1 - pp$ e p, cioè

$$\begin{cases} \theta_A \leq \theta_B & \theta_A > \theta_B \\ 1-p & p \end{cases}$$

Se PowerBalance (contenuto, per esempio, nel sacchetto A) non funziona, le due variabili sono indipendenti e dunque è facile calcolare la probabilità che un singolo esperimento abbia dato successo: è data dalla probabilità che sia uscita la sequenza AB e $\theta_A > \theta_B$ $(1/2 \cdot p)$ più la probabilità che sia uscita la sequenza BA e $\theta_A \leq \theta_B$ $(1/2 \cdot (1-p))$. In tutto, $1/2 \cdot (p + 1 - p) = 1/2$, indipendentemente dal valore di p che sono ignoti (dipendono da come ognuno reagisce alla "non" influenza di due sacchetti).

Se ripetiamo l'esperimento N volte, sappiamo quanto vale la probabilità di avere esattamente n successi,

$$f(n) = \binom{N}{n} \left(\frac{1}{2}\right)^N$$

che è la distribuzione di una variabile aleatoria binomiale. Siamo pronti per verificare se PowerBalance funziona oppure no.

4.3 Verifica di ipotesi

Nel test di verifica di ipotesi unilaterale, si formula un'ipotesi e una sua alternativa:

H_0 : il numero di successi è meta del numero di esperimenti
H_1 : il numero di successi è maggiore della metà del numero di esperimenti

H_0 è chiamata ipotesi *nulla:* è quella che mantiene lo stato attuale delle cose e che non crea sconvolgimenti nel modo corrente di intendere le stesse. Ovviamente, se si accetta H_0 (che equivale a dire che l'esperimento che abbiamo condotto si comporta come una variabile aleatoria binomiale) si può concludere che il PowerBalance non ha alcun effetto. Si conduce l'esperimento N volte. Possono verificarsi tre casi.

1. Il numero di successi è inferiore alla media $N \cdot 1/2$. In tal caso si conclude che PowerBalance non ha alcun effetto. Anzi, per qualche motivo (estraneo allo scopo del test) il PowerBalance funziona solo su una minoranza del campione.

2. Il numero di successi è uguale alla media. In tal caso si conclude che il PowerBalance si comporta esattamente come l'altro braccialetto.

3. Il numero di successi è maggiore della media. È del tutto evidente che se è di poco maggiore, l'efficacia del PowerBalance non è dimostrata. Si tratta di quantificare uno scarto di riferimento Δ dalla media che occorre osservare per rifiutare H_0.

Da un lato, non si vorrebbe incorrere nell'errore (detto di prima specie) di rifiutare l'ipotesi H_0 qualora essa fosse vera. E siccome, in linea di principio, se H_0 è vera può succedere che lo scarto dalla media del numero di successi sia molto alto (al massimo quando si verificano proprio N successi, la cui probabilità non è nulla ma vale $\binom{N}{N}\left(\frac{1}{2}\right)^N = \left(\frac{1}{2}\right)^N$), bisognerebbe accettare sempre l'ipotesi H_0, qualunque sia l'esito dell'esperimento. Ma così facendo si rischia di incorrere nell'errore detto di seconda specie, cioè di accettare H_0 qualora essa sia falsa. Bisogna fare un compromesso e correre il rischio di commettere un errore, o di prima specie o di seconda. Universalmente, si preferisce limitare la possibilità di incorrere nell'errore di prima specie ed evitare lo sforzo di tentare di capire un nuovo fenomeno che invece non esiste. Si fissa allora un livello α di *significatività*, cioè la probabilità massima di fare un errore di prima specie.

Quindi, la strategia da adottare sarà quella di rifiutare l'ipotesi nulla se la probabilità che una variabile binomiale dia un valore maggiore o uguale a $N \cdot 1/2 + \Delta$ vale α, ove α è generalmente un numero preso tra $0.05, 0.01, 0.001$. Se l'esperimento fornisce proprio uno scarto maggiore o uguale di Δ e l'ipotesi H_0 era veramente sbagliata, si è fatto bene a rifiutarla. Se invece era giusta, si è fatto male: ma lo scarto di riferimento Δ era stato scelto in modo che fosse poco probabile (precisamente, α) che venisse fuori uno scarto maggiore o uguale. Quindi, è poco probabile aver sbagliato. Resta la domanda: come si calcola lo scarto di riferimento Δ? Semplicemente con la formula della distribuzione della variabile binomiale. Infatti, dobbiamo considerare, al variare di i, le quantità

$$\sum_{n=N/2+i}^{N} f(n) = \sum_{n=N/2+i}^{N} \binom{N}{n} \left(\frac{1}{2}\right)^{N}, \quad i = 0, \ldots, N/2$$

(assumiamo N pari) che esprimono la probabilità di ottenere $N/2 + i$ o più successi. Se $i = 0$, tale quantità vale poco più di 1/2. Al crescere di i, tali quantità decrescono. Ad un certo punto, per un valore particolare di i tale quantità vale α (o un po' di meno, visto che non variano con continuità). Quello sarà il valore dello scarto di riferimento Δ.

	0.05	0.01	0.001
20	5	6	8
40	6	8	11
60	7	10	13
80	8	11	15

Tabella 4.1: Valori degli scarti di riferimento Δ per vari $\alpha = 0.05, 0.01, 0.001$ e $N = 20, 40, 60, 80$.

Nella Tabella 4.1 è possibile vedere alcuni valori dello scarto di riferimento per diversi valori di N e di α.

4.4 Esercizi

1. Verificare la correttezza della Tabella 4.1.

2. L'esperimento in doppio cieco condotto il 28 Febbraio 2011 al "Medi", su 60 prove, ha dato i seguenti risultati: 35 A, 25 B, 32 casi in cui il Powerbalance funziona. Quali conclusioni ne puoi trarre?

Bibliografia

[1] Hugh D. Young, Elaborazione statistica dei dati sperimentali, Veschi Editore, Roma.