CAPITOLO

11

Distribuzioni di probabilità

Il singolo lancio è imprevedibile; mille lanci non lo sono quasi più. Le distribuzioni di probabilità descrivono la regolarità che il caso produce. (testo della redazione)

OBIETTIVI

1. Definire una variabile aleatoria discreta e descriverla con la distribuzione di probabilità e la funzione di ripartizione.
2. Calcolare valore atteso, varianza e deviazione standard, anche di una trasformazione lineare $aX + b$.
3. Riconoscere le situazioni descritte dalla distribuzione binomiale e calcolarne probabilità, media e varianza.
4. Usare una densità di probabilità per calcolare probabilità come aree, in particolare per la distribuzione uniforme.
5. Risolvere problemi con la distribuzione normale mediante la standardizzazione, la tabella di Φ e la regola 68–95–99,7.
6. Discutere quando una distribuzione è un buon modello di un fenomeno, e il rapporto tra caso e determinismo.

INDICE DEL CAPITOLO

11.1 Variabili aleatorie discrete 2
11.2 La distribuzione binomiale 6
11.3 Distribuzioni continue 9
11.4 La distribuzione normale 11
11.5 Modelli probabilistici 16

Matematica 5 · Parte V · Probabilità, modelli e fondamenti · versione 1.0
Capitolo 11: Distribuzioni di probabilità · pagine 293–316 del volume completo · estratto per lo studio

Nessuno sa dire se il prossimo lancio di una moneta darà testa, se l'autobus arriverà tra due minuti o tra dieci, quanto sarà alta la prossima persona che entrerà in classe. Eppure su mille lanci, su cento giorni di attesa alla fermata, su migliaia di persone, compaiono regolarità precise: le teste sono circa la metà, le attese si distribuiscono in modo prevedibile, le altezze si addensano attorno a un valore centrale e diventano rare man mano che ci si allontana. Il singolo caso è imprevedibile; l'insieme dei casi ha una *forma*.

Questa forma si chiama *distribuzione di probabilità*, ed è l'oggetto di questo capitolo. Riprendiamo la probabilità studiata negli anni precedenti, la colleghiamo a un numero che dipende dal caso (la *variabile aleatoria*) e incontriamo le due distribuzioni più importanti: la *binomiale*, che conta i successi in prove ripetute, e la *normale*, la curva a campana di Gauss già incontrata nel capitolo 7. Per le grandezze continue la probabilità diventerà un'area, e qui torneranno utili gli integrali dei capitoli 9 e 10.

11.1 Variabili aleatorie discrete

11.1.1 Che cosa sappiamo già

In *Matematica 2*, cap. 8, la probabilità di un evento è stata definita come rapporto tra casi favorevoli e casi possibili (quando sono equiprobabili) e confrontata con la frequenza relativa osservata su molte prove. Lì abbiamo incontrato anche il *valore atteso* di un guadagno, cioè la somma dei guadagni possibili moltiplicati per le loro probabilità, e il *gioco equo*, quello in cui il valore atteso del guadagno è zero. In *Matematica 4*, cap. 10, abbiamo ricavato la formula di Bernoulli per le prove ripetute; useremo per le combinazioni, come in quel volume, il simbolo $\mathrm{C}_{n,k} = \dfrac{n!}{k!\,(n-k)!}$. Infine in *Matematica 4*, cap. 11, abbiamo misurato la dispersione di una serie di dati con la deviazione standard σ, la radice della media dei quadrati degli scarti dalla media. In questo capitolo vedremo che queste idee sono pezzi di un unico disegno.

11.1.2 Un numero per ogni esito

Lanciamo due dadi e guardiamo la somma dei punti. L'esito dell'esperimento è una coppia, per esempio (2; 5), ma a noi interessa un numero, 7. Una **variabile aleatoria** (o *casuale*) è proprio questo: una regola che associa un numero reale a ogni esito di un esperimento casuale. La indichiamo con una lettera maiuscola, X, e i suoi valori con la minuscola, x. La somma di due dadi, il numero di teste in tre lanci di una moneta, il guadagno in un gioco, il numero di pezzi difettosi in una scatola sono variabili aleatorie.

Quando i valori possibili sono in numero finito (o si possono elencare uno dopo l'altro come 0, 1, 2, …) la variabile si dice **discreta**. Una grandezza che può assumere tutti i valori di un inter-

vallo, come un tempo di attesa o un'altezza, si dice invece *continua*: la studieremo più avanti nel capitolo, perché richiede un'idea nuova.

11.1.3 La distribuzione di probabilità

Dei 36 esiti equiprobabili del lancio di due dadi, uno solo dà somma 2, cioè (1; 1); due danno somma 3, (1; 2) e (2; 1); e così via fino a sei esiti per la somma 7, poi di nuovo in calo fino a 12. La tabella che associa a ogni valore x_i la sua probabilità $p_i = \mathrm{P}(X = x_i)$ si chiama **distribuzione di probabilità** di X (Tabella 11.1).

Tabella 11.1 Distribuzione di probabilità della somma X dei punti di due dadi.

x_i	2	3	4	5	6	7	8	9	10	11	12
p_i	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Poiché gli eventi $X = x_i$ sono incompatibili e uno di essi si verifica di sicuro, le probabilità di una distribuzione sono sempre non negative e hanno somma 1: $\Sigma\, p_i = 1$. Qui $1 + 2 + 3 + 4 + 5 + 6 + 5 + 4 + 3 + 2 + 1 = 36$, e infatti $\frac{36}{36} = 1$. La distribuzione si rappresenta con un *diagramma a barre*, con un'asta di altezza p_i sopra ogni valore x_i (Figura 11.1, a sinistra): la forma «a tetto» dice a colpo d'occhio che 7 è la somma più probabile e che i valori estremi sono rari.

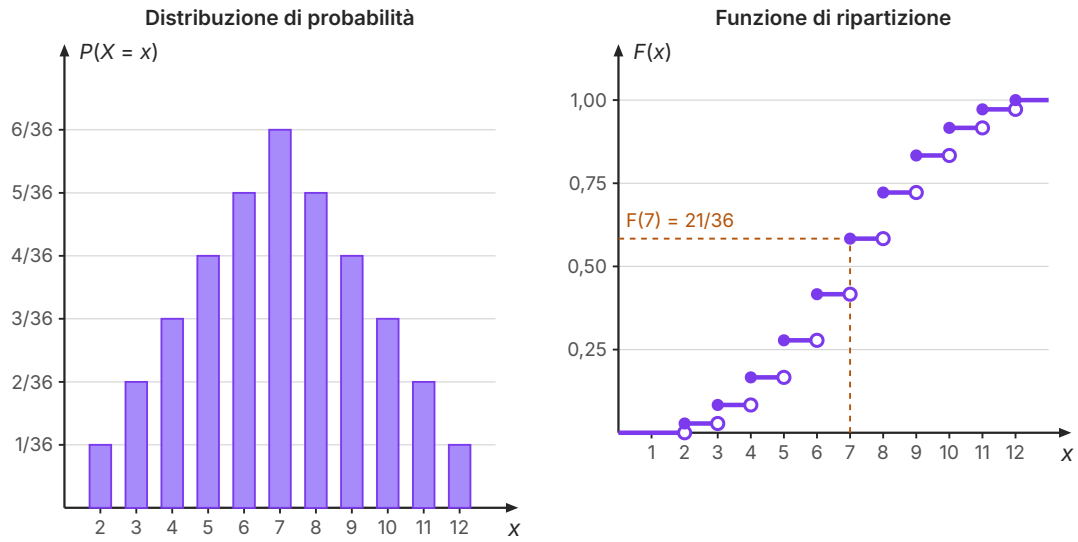

Figura 11.1 La somma X dei punti di due dadi. A sinistra la distribuzione di probabilità; a destra la funzione di ripartizione $F(x) = \mathrm{P}(X \le x)$, costante a tratti, con un salto di altezza p_i in ogni valore x_i (pallino pieno: valore assunto; pallino vuoto: valore escluso). Scale diverse sui due assi.

11.1.4 La funzione di ripartizione

Spesso interessa la probabilità che X non superi un certo valore: «al massimo 7», «non più di due pezzi difettosi». La **funzione di ripartizione** di X è la funzione definita per ogni numero reale x da

$$F(x) = \mathrm{P}(X \le x).$$

Per la somma di due dadi $F(x) = 0$ per $x < 2$, perché nessuna somma è minore di 2; in $x = 2$ sale a $\frac{1}{36}$ e resta costante fino a 3, dove sale a $\frac{3}{36}$; in 7 vale $\frac{21}{36}$, e da 12 in poi vale 1. Il grafico è una *scala* (Figura 11.1, a destra): ogni gradino è alto quanto la probabilità del valore in cui si trova. La funzione di ripartizione è non decrescente, vale 0 a sinistra di tutti i valori e 1 a destra; inoltre $\mathrm{P}(a < X \le b) = F(b) - F(a)$. Per esempio la probabilità che la somma sia 8, 9 o 10 è $F(10) - F(7) = \frac{33}{36} - \frac{21}{36} = \frac{12}{36} = \frac{1}{3}$.

11.1.5 Valore atteso, varianza, deviazione standard

Il **valore atteso** (o *valore medio*, o *speranza matematica*) di una variabile aleatoria discreta è la media dei suoi valori, ciascuno pesato con la sua probabilità:

$$\mu = \mathrm{E}(X) = x_1 p_1 + x_2 p_2 + \ldots + x_n p_n = \Sigma\, x_i p_i.$$

È la stessa formula della media ponderata della statistica, con le probabilità al posto delle frequenze relative; e per la legge empirica del caso, ripetendo l'esperimento molte volte, la media dei valori osservati tende ad avvicinarsi a $\mathrm{E}(X)$. Per i due dadi, grazie alla simmetria della distribuzione attorno a 7, si trova $\mathrm{E}(X) = 7$: provare per credere ($\frac{2\cdot 1 + 3\cdot 2 + \ldots + 12\cdot 1}{36} = \frac{252}{36} = 7$).

Il valore atteso non basta a descrivere una distribuzione: due variabili possono avere lo stesso valore medio ma valori più o meno dispersi. Come in statistica, misuriamo la dispersione con la media dei quadrati degli scarti. La **varianza** di X è

$$\mathrm{Var}(X) = \sigma^2 = \Sigma\,(x_i - \mu)^2 p_i = \mathrm{E}(X^2) - \mu^2,$$

dove $\mathrm{E}(X^2) = \Sigma\, x_i^2\, p_i$; la seconda forma (dimostrata negli esercizi) è di solito più comoda. La **deviazione standard** $\sigma = \sqrt{\mathrm{Var}(X)}$ si misura nella stessa unità di X e indica la distanza «tipica» dei valori dal valore medio. Per la somma dei dadi $\mathrm{E}(X^2) = \frac{1974}{36} = \frac{329}{6}$, quindi $\sigma^2 = \frac{329}{6} - 49 = \frac{35}{6} \approx 5{,}83$ e $\sigma \approx 2{,}42$.

ESEMPIO SVOLTO

Problema. Si lancia tre volte una moneta equilibrata e X è il numero di teste. Trova la distribuzione di X, il valore atteso e la deviazione standard.

Distribuzione. Gli 8 esiti TTT, TTC, ..., CCC sono equiprobabili. Zero teste: 1 esito; una testa: 3 esiti (la testa può essere al primo, al secondo o al terzo lancio); due teste: 3; tre teste: 1. Quindi X vale 0, 1, 2, 3 con probabilità $\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8}$ (somma 1).

Valore atteso. $\mathrm{E}(X) = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{12}{8} = \frac{3}{2}$. Si noti che 1,5 teste non può uscire in nessun esperimento: il valore atteso è una media, non un valore previsto.

Varianza. $\mathrm{E}(X^2) = 0 + 1 \cdot \frac{3}{8} + 4 \cdot \frac{3}{8} + 9 \cdot \frac{1}{8} = \frac{24}{8} = 3$, quindi $\sigma^2 = 3 - \frac{9}{4} = \frac{3}{4}$ e $\sigma = \frac{\sqrt{3}}{2} \approx 0{,}87$.

Che cosa succede se trasformiamo X in modo lineare, per esempio cambiando l'unità di misura o aggiungendo una quota fissa? Supponiamo che a ogni testa si vinca 1 € e a ogni croce si perda 1 €: con tre lanci il guadagno è $G = X - (3 - X) = 2X - 3$. Moltiplicare tutti i valori per a moltiplica la media per a e gli scarti per a; aggiungere b sposta la media di b e non cambia gli scarti. Valgono quindi le regole riassunte qui sotto, che si dimostrano direttamente dalle definizioni (esercizio 19).

IN SINTESI

Variabile aleatoria discreta con valori x_i e probabilità p_i ($p_i \geq 0$, $\Sigma\, p_i = 1$):

- valore atteso $\mu = \mathrm{E}(X) = \Sigma\, x_i\, p_i$;
- varianza $\sigma^2 = \mathrm{Var}(X) = \Sigma\, (x_i - \mu)^2\, p_i = \mathrm{E}(X^2) - \mu^2$; deviazione standard $\sigma = \sqrt{\mathrm{Var}(X)}$;
- funzione di ripartizione $F(x) = \mathrm{P}(X \leq x)$, con $\mathrm{P}(a < X \leq b) = F(b) - F(a)$;
- trasformazioni lineari: $\mathrm{E}(aX + b) = aE(X) + b$ e $\mathrm{Var}(aX + b) = a^2\, \mathrm{Var}(X)$, quindi la deviazione standard di $aX + b$ è $|a|\, \sigma$.

Verifichiamo sul guadagno $G = 2X - 3$. Le regole danno $\mathrm{E}(G) = 2 \cdot \frac{3}{2} - 3 = 0$ (il gioco è equo, come ci aspettavamo) e $\mathrm{Var}(G) = 4 \cdot \frac{3}{4} = 3$. Direttamente: G vale -3, -1, 1, 3 con probabilità $\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8}$; la media è 0 e $\mathrm{E}(G^2) = \frac{9 + 3 + 3 + 9}{8} = 3$, quindi $\mathrm{Var}(G) = 3$.

ERRORE FREQUENTE

Aggiungere una costante non cambia la dispersione: $\mathrm{Var}(X + b) = \mathrm{Var}(X)$, non $\mathrm{Var}(X) + b$. E un fattore a moltiplica la varianza per a^2, non per a: se X è una lunghezza in metri con $\sigma = 0{,}02$ m, la stessa lunghezza in centimetri, $100X$, ha $\sigma = 2$ cm e varianza 10 000 volte più grande. Per non sbagliare, ricorda che la deviazione standard si misura nella stessa unità di X, la varianza nel suo quadrato.

11.2 La distribuzione binomiale

11.2.1 Dalla formula di Bernoulli a una distribuzione

Ripetiamo n volte, in modo indipendente, una prova che ha due soli esiti, *successo* con probabilità p e *insuccesso* con probabilità $1 - p$, e chiamiamo X il numero dei successi. In *Matematica 4*, cap. 10, abbiamo trovato la probabilità di ottenere esattamente k successi; ora quella formula, letta per tutti i valori di k insieme, diventa una distribuzione di probabilità.

IN SINTESI

Distribuzione binomiale $\mathrm{B}(n;\, p)$. Il numero X di successi in n prove indipendenti, ciascuna con probabilità di successo p, ha la distribuzione

$$\mathrm{P}(X = k) = \mathrm{C}_{n,k} p^k (1 - p)^{n-k}, \qquad k = 0, 1, \ldots, n.$$

Valore atteso e varianza: $\mu = np$, $\sigma^2 = np(1 - p)$.

Diciamo che X ha **distribuzione binomiale** di parametri n e p e scriviamo $X \sim \mathrm{B}(n;\, p)$. Le probabilità hanno somma 1: per la formula del binomio di Newton, $\sum \mathrm{C}_{n,k}\, p^k (1 - p)^{n-k} = (p + 1 - p)^n = 1$. Il numero di teste in tre lanci dell'esempio precedente è $\mathrm{B}(3;\, \frac{1}{2})$: la formula dà $\mathrm{C}_{3,k} \cdot \frac{1}{8}$, cioè proprio $\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8}$.

ESEMPIO SVOLTO

Problema. Una macchina produce pezzi difettosi con probabilità 5%, indipendentemente l'uno dall'altro (dati inventati). Si controlla una scatola di 20 pezzi. Qual è la probabilità di trovare a) nessun pezzo difettoso; b) esattamente uno; c) almeno due? Quanti difettosi ci si aspetta in media?

Modello. X = numero di pezzi difettosi, $X \sim B(20; 0{,}05)$ («successo» qui è trovare un difetto).

a) $P(X = 0) = 0{,}95^{20} \approx 0{,}358$.

b) $P(X = 1) = C_{20,1} \cdot 0{,}05 \cdot 0{,}95^{19} = 20 \cdot 0{,}05 \cdot 0{,}95^{19} \approx 0{,}377$.

c) Conviene l'evento contrario: $P(X \geq 2) = 1 - P(X = 0) - P(X = 1) \approx 1 - 0{,}3585 - 0{,}3774 \approx 0{,}264$, circa una scatola su quattro.

Media. $\mu = np = 20 \cdot 0{,}05 = 1$ pezzo difettoso per scatola, con $\sigma = \sqrt{20 \cdot 0{,}05 \cdot 0{,}95} \approx 0{,}97$.

11.2.2 Perché la media è np

Il risultato $\mu = np$ è quello che l'intuizione suggerisce: in 20 pezzi con il 5% di difetti ci si aspetta un difetto. Per giustificarlo, scriviamo X come somma di n variabili più semplici. Sia X_i la variabile che vale 1 se l'i-esima prova è un successo e 0 altrimenti (una **variabile indicatrice**): allora $X = X_1 + X_2 + \dots + X_n$, perché sommando gli 1 si contano i successi. Ciascuna X_i ha $E(X_i) = 1 \cdot p + 0 \cdot (1 - p) = p$. Il valore atteso di una somma è la somma dei valori attesi (lo ammettiamo: è la stessa proprietà della media aritmetica), quindi $E(X) = p + p + \dots + p = np$.

Anche la varianza si ottiene dalle indicatrici: $E(X_i^2) = p$, quindi $\mathrm{Var}(X_i) = p - p^2 = p(1 - p)$. Per variabili *indipendenti* la varianza della somma è la somma delle varianze (anche questo lo enunciamo soltanto), e si trova $\sigma^2 = np(1 - p)$. Controlliamolo per $n = 2$: X vale 0, 1, 2 con probabilità $(1 - p)^2$, $2p(1 - p)$, p^2; $E(X) = 2p(1 - p) + 2p^2 = 2p$ e $E(X^2) = 2p(1 - p) + 4p^2$, quindi $\mathrm{Var}(X) = 2p(1 - p) + 4p^2 - 4p^2 = 2p(1 - p)$. Per tre lanci di una moneta la formula dà $3 \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{4}$, come avevamo calcolato.

11.2.3 La forma della distribuzione

La Figura 11.2 mostra tre distribuzioni binomiali calcolate con la formula. Con $p = 0{,}5$ il diagramma è simmetrico attorno a $\frac{n}{2}$, perché successi e insuccessi si scambiano i ruoli. Con $p = 0{,}3$ e $n = 10$ è asimmetrico: il valore più probabile è 3 (con probabilità circa 0,267), e la «coda» si allunga verso destra. Con $n = 50$ e lo stesso p, la distribuzione è centrata in $\mu = 15$, è più bassa e più larga ($\sigma = \sqrt{10{,}5} \approx 3{,}24$ contro $\sqrt{2{,}1} \approx 1{,}45$ per $n = 10$) e ha già quasi la forma di una campana simmetrica. È il primo indizio della distribuzione normale, che incontreremo nel paragrafo 4.

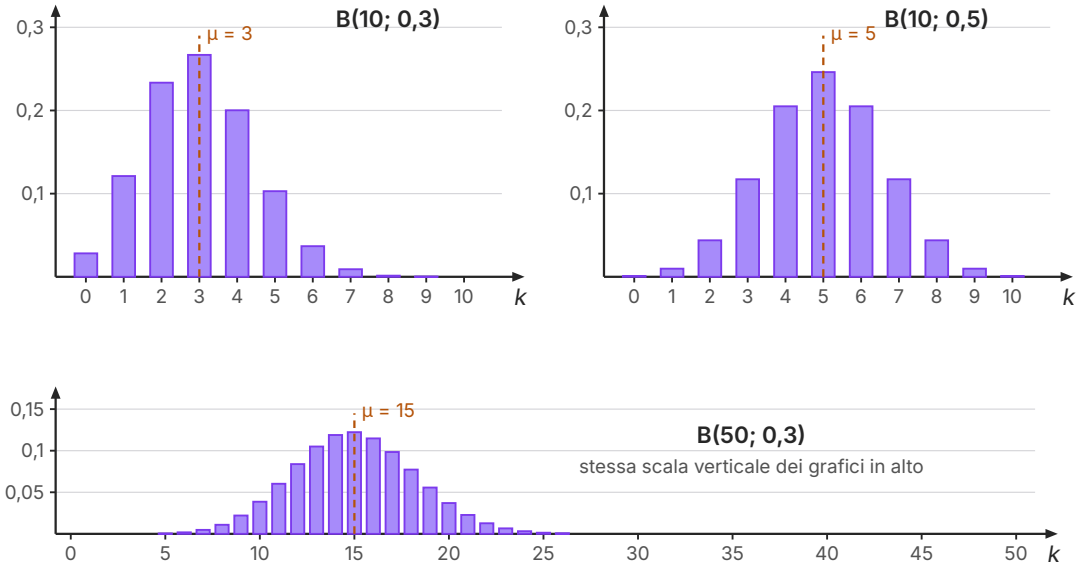

Figura 11.2 Diagrammi a barre di tre distribuzioni binomiali, calcolati da $P(X = k) = C_{n,k}\, p^k\, (1 - p)^{n-k}$. La linea tratteggiata segna il valore atteso $\mu = np$. Per B(50; 0,3) sono disegnate le barre visibili (per k da 5 a 26); le altre probabilità sono minori di 0,0004.

11.2.4 La legge dei grandi numeri

Se $X \sim B(n; p)$, la *frequenza relativa* dei successi è $\frac{X}{n}$. Per le regole delle trasformazioni lineari ha valore atteso $\frac{np}{n} = p$ e deviazione standard $\frac{\sqrt{np(1-p)}}{n} = \sqrt{\frac{p(1-p)}{n}}$, che diventa sempre più piccola al crescere di n. Per una moneta equilibrata la deviazione standard della frequenza delle teste è 0,05 con 100 lanci e 0,005 con 10 000 lanci: la frequenza si stringe attorno a $\frac{1}{2}$. È la **legge dei grandi numeri**, che Jakob Bernoulli dimostrò nell'*Ars Conjectandi*, pubblicata postuma nel 1713 (l'abbiamo incontrata in *Matematica 2*, cap. 8): al crescere del numero delle prove, la probabilità che la frequenza relativa si discosti da p più di un margine fissato, piccolo quanto si vuole, tende a zero.

Attenzione a ciò che la legge *non* dice. Non dice che il numero di successi si avvicina a np: lo scarto assoluto $X - np$ tipicamente *cresce*, come $\sqrt{n}$; è lo scarto relativo che diventa trascurabile. E non dice che il caso «compensa» gli squilibri.

ERRORE FREQUENTE

«Al lotto il 47 non esce da cento estrazioni: ora è più probabile.» È la **fallacia del giocatore**. Le prove sono indipendenti: la moneta, la ruota o l'urna non hanno memoria, e la probabilità alla prossima prova è la stessa di sempre. La legge dei grandi numeri agisce per *diluizione*, non per compensazione: uno squilibrio iniziale resta, ma diventa una frazione sempre più piccola del totale.

11.3 Distribuzioni continue

11.3.1 Quando i valori sono infiniti

Un autobus passa ogni 12 minuti e arriviamo alla fermata senza conoscere l'orario: il tempo di attesa X può essere un qualsiasi numero tra 0 e 12. Qual è la probabilità di aspettare *esattamente* 5 minuti, cioè 5,000 000… minuti? I valori possibili sono infiniti e nessuno è privilegiato: se ciascuno avesse una probabilità positiva, per quanto piccola, sommandone abbastanza si supererebbe 1. Quindi per una variabile **continua** $P(X = c) = 0$ per ogni valore c. Ha senso invece chiedersi qual è la probabilità di aspettare *tra* 3 e 7 minuti: questi sono gli eventi che si osservano davvero, perché ogni misura ha una precisione finita.

Al posto della tabella delle probabilità serve allora una funzione, la **densità di probabilità** $f(x)$, il cui grafico racchiude con l'asse x, tra a e b, un'area uguale alla probabilità che X cada in $[a; b]$. Nel capitolo 9 abbiamo visto che un'area di questo tipo è un integrale definito.

Densità di probabilità. Una variabile aleatoria continua X ha densità f se

- $f(x) \geq 0$ per ogni x;
- l'area totale sotto il grafico di f è 1;
- $P(a \leq X \leq b) = \int_a^b f(x)\,dx$ per ogni intervallo $[a; b]$.

Di conseguenza $P(X = c) = 0$ e $P(a \leq X \leq b) = P(a < X < b)$. Valore atteso e varianza si definiscono come nel caso discreto, con l'integrale al posto della somma: $\mu = \int x\, f(x)\,dx$ e $\sigma^2 = \int (x - \mu)^2 f(x)\,dx$, estesi a tutto l'intervallo in cui f non è nulla.

Il confronto con la statistica aiuta. In un istogramma delle frequenze di una grandezza continua (le altezze, per esempio), l'area di ogni rettangolo è la frequenza relativa della sua classe. Con moltissimi dati e classi sempre più strette, il profilo dell'istogramma si avvicina a una curva: la densità è il limite ideale di quel profilo.

11.3.2 La distribuzione uniforme

Torniamo all'autobus. Se nessun istante è privilegiato, la densità deve essere costante in $[0; 12]$ e nulla fuori; perché l'area totale, un rettangolo di base 12, sia 1, l'altezza deve essere $\frac{1}{12}$. La

probabilità di aspettare tra 3 e 7 minuti è l'area di un rettangolo di base 4 (Figura 11.3):

$$P(3 \le X \le 7) = \int_3^7 \frac{1}{12}\,dx = \frac{7-3}{12} = \frac{1}{3}.$$

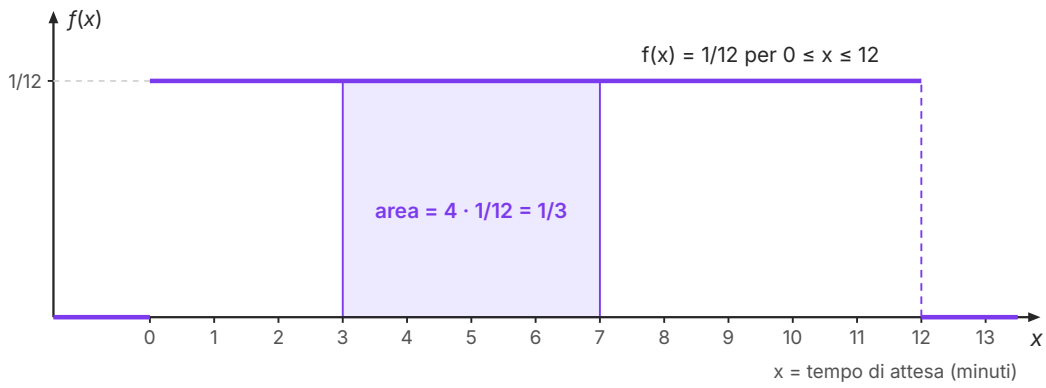

Figura 11.3 Densità uniforme del tempo di attesa X (in minuti) per un autobus che passa ogni 12 minuti. L'area colorata è $P(3 \le X \le 7) = \frac{1}{3}$. Scale diverse sui due assi.

In generale una variabile ha **distribuzione uniforme** nell'intervallo $[a; b]$ se la sua densità vale $f(x) = \frac{1}{b-a}$ in $[a; b]$ e 0 fuori. La probabilità di un sottointervallo è proporzionale alla sua ampiezza: è la versione continua dei «casi equiprobabili». Il valore atteso è il punto medio,

$$\mu = \int_a^b \frac{x}{b-a}\,dx = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2},$$

come suggerisce la simmetria. Per la varianza calcoliamo prima $E(X^2) = \int_a^b \frac{x^2}{b-a}\,dx = \frac{b^3-a^3}{3(b-a)} = \frac{a^2+ab+b^2}{3}$; allora

$$\sigma^2 = \frac{a^2+ab+b^2}{3} - \frac{(a+b)^2}{4} = \frac{4a^2+4ab+4b^2-3a^2-6ab-3b^2}{12} = \frac{(b-a)^2}{12}.$$

Un esempio. Per il tempo di attesa dell'autobus, uniforme in $[0; 12]$, calcola il valore atteso e la deviazione standard, e la probabilità di aspettare più di 10 minuti.

Valore atteso. $\mu = \frac{0+12}{2} = 6$ minuti.

Varianza, direttamente dalla definizione: $\sigma^2 = \int_0^{12} (x-6)^2 \cdot \frac{1}{12}\,dx = \left[\frac{(x-6)^3}{36}\right]_0^{12} = \frac{216-(-216)}{36} = 12$, come dà la formula $\frac{12^2}{12}$. Quindi $\sigma = \sqrt{12} \approx 3{,}46$ minuti.

Attesa lunga. $P(X > 10) = \dfrac{12 - 10}{12} = \dfrac{1}{6}$: circa un giorno su sei.

ERRORE FREQUENTE

Il valore $f(x)$ di una densità *non* è una probabilità: è l'area a esserlo. Per questo una densità può superare 1. Se il tempo di attesa è uniforme in [0; 0,5] (in minuti), la densità vale 2, e l'area totale è $0{,}5 \cdot 2 = 1$. E per lo stesso motivo, con una variabile continua, scrivere ≤ o < non cambia nulla, mentre con una discreta può cambiare molto: per la somma di due dadi $P(X \le 7) = \dfrac{21}{36}$ ma $P(X < 7) = \dfrac{15}{36}$.

APPROFONDIMENTO · IL TEMPO DI ATTESA ESPONENZIALE

Non tutte le attese sono uniformi. Il tempo che passa prima che un nucleo radioattivo decada, o prima che arrivi una telefonata a un centralino, si descrive bene con la densità *esponenziale* $f(x) = \lambda\, e^{-\lambda x}$ per $x \ge 0$ (e 0 per $x < 0$), dove $\lambda > 0$ è una costante. L'area da 0 a t è $\int_0^t \lambda\, e^{-\lambda x}\, dx = [-e^{-\lambda x}]_0^t = 1 - e^{-\lambda t}$, che tende a 1 per $t \to +\infty$: l'area totale è 1, anche se si estende su una semiretta. Quindi $P(X > t) = e^{-\lambda t}$: la probabilità di aspettare più di t decresce esponenzialmente, come il numero dei nuclei non ancora decaduti nel decadimento radioattivo di *Matematica 4*, cap. 7. Il tempo medio di attesa è $\mu = \dfrac{1}{\lambda}$.

Per esempio, se le chiamate arrivano in media ogni 2 minuti ($\lambda = 0{,}5$ al minuto), la probabilità di aspettare più di 4 minuti è $e^{-2} \approx 0{,}135$. Curiosamente, aver già aspettato non accorcia l'attesa che resta: la distribuzione esponenziale «non ha memoria», come i lanci di una moneta.

11.4 La distribuzione normale

11.4.1 La curva a campana

Misuriamo l'altezza di migliaia di persone dello stesso sesso e della stessa età, o ripetiamo molte volte la stessa misura in laboratorio, o pesiamo i pacchi di pasta usciti da una macchina: in tutti questi casi l'istogramma ha una forma a campana, simmetrica attorno al valore medio. Il modello matematico di questa forma è la **distribuzione normale** (o *gaussiana*), con densità

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

Non occorre ricordare la formula, ma conviene saperla leggere. Dipende da due parametri: μ, il valore atteso, e $\sigma > 0$, la deviazione standard. Scriviamo $X \sim N(\mu; \sigma)$: in questo libro **il secondo parametro è sempre la deviazione standard σ**, non la varianza σ^2 (altri testi usano l'altra con-

venzione: controlla sempre). Il fattore $\frac{1}{\sigma\sqrt{2\pi}}$ serve a rendere uguale a 1 l'area totale; che lo faccia davvero è un risultato che non dimostriamo, perché la funzione e^{-x^2} non ha una primitiva esprimibile con le funzioni elementari.

Le proprietà del grafico (Figura 11.4) si ricavano con lo studio di funzione del capitolo 7, dove abbiamo studiato $y = e^{-x^2}$. Prendiamo prima $\mu = 0$ e $\sigma = 1$, cioè $g(x) = e^{-x^2/2}$ a meno del fattore costante. Si ha $g'(x) = -x\, e^{-x^2/2}$, positiva per $x < 0$ e negativa per $x > 0$: massimo in 0. Inoltre $g''(x) = (x^2 - 1)\, e^{-x^2/2}$, che cambia segno in $x = \pm 1$: i flessi sono a distanza 1 dal massimo. Sostituendo x con $\frac{x-\mu}{\sigma}$ il grafico si sposta di μ e si dilata orizzontalmente di un fattore σ, quindi:

- il grafico è simmetrico rispetto alla retta $x = \mu$, dove ha il massimo $\frac{1}{\sigma\sqrt{2\pi}}$;
- i **flessi** sono in $x = \mu - \sigma$ e $x = \mu + \sigma$: la deviazione standard si vede sul grafico come la distanza dei flessi dall'asse di simmetria;
- l'asse x è asintoto orizzontale per $x \to -\infty$ e per $x \to +\infty$; la densità è positiva ovunque, ma oltre 3 o 4 deviazioni standard dal centro è praticamente nulla;
- a σ più grande corrisponde una campana più bassa e più larga (l'area resta 1); cambiare μ la trasla.

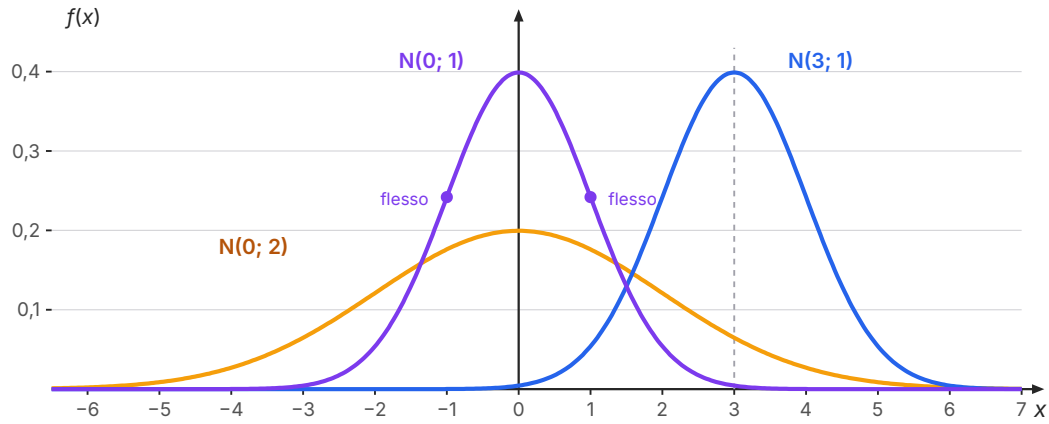

Figura 11.4 Densità normali N(0; 1), N(3; 1) e N(0; 2), calcolate dalla formula. Cambiare μ trasla la curva; raddoppiare σ la allarga e ne dimezza l'altezza. I pallini segnano i flessi di N(0; 1), in $x = \pm 1$. Scale diverse sui due assi.

11.4.2 La standardizzazione

Le aree sotto una curva normale non si calcolano con una primitiva elementare: si leggono su una tabella. Basta però una sola tabella per tutte le normali. Se $X \sim \mathrm{N}(\mu; \sigma)$, la variabile

$$Z = \frac{X - \mu}{\sigma}$$

misura la distanza di X dal valore medio *in deviazioni standard*, e per le regole delle trasformazioni lineari ha media 0 e deviazione standard 1. Si dimostra che è ancora normale: $Z \sim N(0; 1)$, la **normale standard**. Il passaggio da X a Z si chiama **standardizzazione**, e $z = \frac{x - \mu}{\sigma}$ si dice *punteggio standard* del valore x.

La funzione di ripartizione della normale standard si indica con $\Phi(z) = P(Z \le z)$, ed è tabulata nel Formulario in fondo al volume per z da 0 a 3,2. Per la simmetria della campana bastano i valori positivi, perché le probabilità per z negativo si ricavano per simmetria (seconda regola qui sotto).

Calcolo con la normale. Se $X \sim N(\mu; \sigma)$, si standardizza con $z = \frac{x - \mu}{\sigma}$ e si usa la tabella di Φ:

- $P(X \le x) = \Phi(z)$ e $P(X > x) = 1 - \Phi(z)$;
- $\Phi(-z) = 1 - \Phi(z)$ (simmetria); in particolare $\Phi(0) = 0{,}5$;
- $P(x_1 \le X \le x_2) = \Phi(z_2) - \Phi(z_1)$;
- $P(|X - \mu| < k\sigma) = 2\Phi(k) - 1$.

11.4.3 La regola 68–95–99,7

L'ultima formula, con $k = 1, 2, 3$ e i valori della tabella, dà tre numeri da ricordare (Figura 11.5):

$$P(|X - \mu| < \sigma) \approx 68{,}3\%, \quad P(|X - \mu| < 2\sigma) \approx 95{,}4\%, \quad P(|X - \mu| < 3\sigma) \approx 99{,}7\%.$$

Per esempio $2\Phi(1) - 1 = 2 \cdot 0{,}8413 - 1 = 0{,}6826$. In parole: in una distribuzione normale circa due valori su tre cadono entro una deviazione standard dalla media, circa 19 su 20 entro due, quasi tutti entro tre. Un valore a più di tre deviazioni standard dalla media capita circa 3 volte su 1000: quando lo si osserva, è lecito sospettare che qualcosa non vada, nella misura o nel modello.

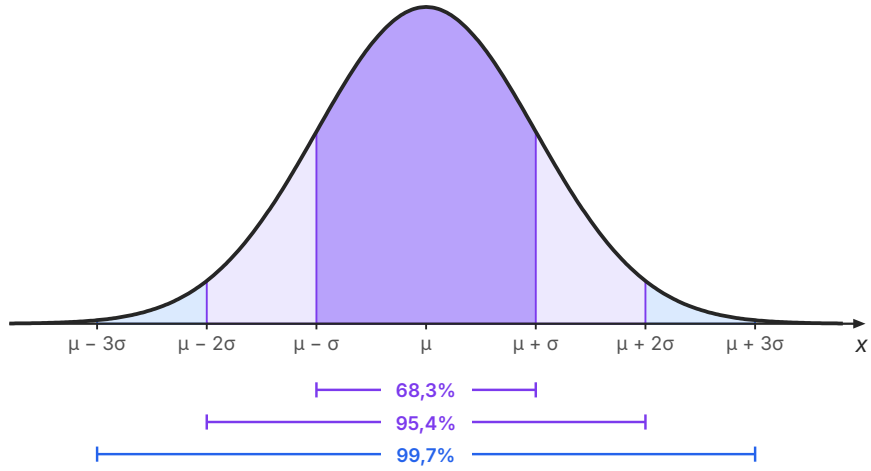

Figura 11.5 La regola 68–95–99,7 per una distribuzione normale N(μ; σ): le aree comprese entro una, due e tre deviazioni standard dal valore atteso, calcolate con Φ.

ESEMPIO SVOLTO

Problema. L'altezza delle studentesse di 18 anni di una grande scuola è distribuita approssimativamente come N(165; 6), in centimetri (dati inventati). Quale percentuale delle studentesse a) è alta al massimo 171 cm; b) supera 177 cm; c) è alta tra 156 e 174 cm? d) Quale altezza è superata solo dal 10% delle studentesse?

a) $z = \dfrac{171 - 165}{6} = 1$, quindi $P(X \le 171) = \Phi(1) = 0{,}8413$: circa l'84%.

b) $z = \dfrac{177 - 165}{6} = 2$, quindi $P(X > 177) = 1 - \Phi(2) = 1 - 0{,}9772 = 0{,}0228$: circa il 2,3%.

c) I punteggi standard sono −1,5 e 1,5: $P = \Phi(1{,}5) - \Phi(-1{,}5) = 2\Phi(1{,}5) - 1 = 2 \cdot 0{,}9332 - 1 = 0{,}8664$, circa l'87%.

d) Si cerca z con $\Phi(z) = 0{,}90$. Nella tabella $\Phi(1{,}2) = 0{,}8849$ e $\Phi(1{,}3) = 0{,}9032$: interpolando, $z \approx 1{,}28$. L'altezza è $x = \mu + z\sigma \approx 165 + 1{,}28 \cdot 6 \approx 172{,}7$ cm.

Un secondo problema. Una macchina riempie pacchi di pasta da 500 g; il peso effettivo è distribuito come N(505; 5), in grammi (dati inventati). Quale frazione dei pacchi pesa meno di 500 g? Su quale valore medio bisognerebbe regolare la macchina perché i pacchi sottopeso fossero solo il 2,3% circa?

Sottopeso. $z = \dfrac{500 - 505}{5} = -1$, quindi $\mathrm{P}(X < 500) = \Phi(-1) = 1 - 0{,}8413 = 0{,}1587$: circa un pacco su sei.

Nuova regolazione. $1 - \Phi(2) = 0{,}0228$, quindi serve che 500 g sia due deviazioni standard sotto la media: $\mu = 500 + 2 \cdot 5 = 510$ g. Lo stesso risultato si otterrebbe a parità di media rendendo la macchina più precisa, cioè con $\sigma = 2{,}5$ g: è il dilemma pratico tra «riempire di più» e «riempire meglio».

11.4.4 Dalla binomiale alla normale

Nella Figura 11.2 la binomiale B(50; 0,3) aveva già quasi la forma di una campana. Non è un caso: quando n è grande, la distribuzione B(n; p) è ben approssimata dalla normale con la stessa media e la stessa deviazione standard, N(np; $\sqrt{np(1-p)}$). La Figura 11.6 sovrappone le barre di B(50; 0,5) alla densità di N(25; 3,54): le due descrizioni quasi coincidono. Una regola pratica diffusa considera buona l'approssimazione quando np e $n(1-p)$ sono entrambi almeno 5 (o 10, per chi è più prudente).

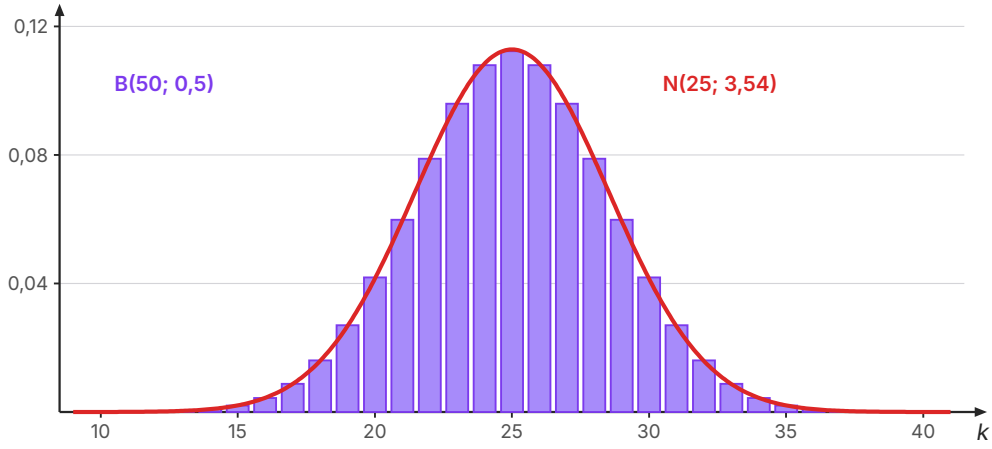

Figura 11.6 Le barre della distribuzione binomiale B(50; 0,5) e la densità normale con la stessa media e la stessa deviazione standard, $\mu = 25$ e $\sigma = \sqrt{12{,}5} \approx 3{,}54$. Sono disegnate le barre da $k = 10$ a $k = 40$; fuori da questo intervallo le probabilità sono minori di 0,0001.

Esempio: con 100 lanci di una moneta equilibrata, qual è la probabilità di ottenere almeno 60 teste? Il calcolo esatto richiederebbe di sommare 41 termini binomiali. Con l'approssimazione normale, $\mu = 50$ e $\sigma = \sqrt{100 \cdot 0{,}5 \cdot 0{,}5} = 5$: 60 teste sono due deviazioni standard sopra la media, e $\mathrm{P}(X \geq 60) \approx 1 - \Phi(2) \approx 0{,}023$. Il valore esatto, calcolato al computer, è circa 0,028: l'ordine di grandezza è giusto, e un piccolo accorgimento (l'Approfondimento qui sotto) migliora molto la stima.

APPROFONDIMENTO · LA CORREZIONE DI CONTINUITÀ

La binomiale assegna la probabilità al valore intero 60 con una barra di base 1 che, come quelle della Figura 11.6, occupa l'intervallo da 59,5 a 60,5. Approssimando $P(X \geq 60)$ con l'area della normale a destra di 60 si perde metà di quella barra. Conviene allora partire da 59,5: $z = \dfrac{59{,}5 - 50}{5} = 1{,}9$ e $P(X \geq 60) \approx 1 - \Phi(1{,}9) = 1 - 0{,}9713 = 0{,}0287$, molto vicino al valore esatto 0,0284. Questo accorgimento si chiama *correzione di continuità*.

La scoperta di questo legame risale al Settecento: Abraham de Moivre, un matematico francese vissuto a Londra, ottenne la curva a campana come forma limite delle probabilità binomiali, a partire dal caso della moneta. All'inizio dell'Ottocento Carl Friedrich Gauss ritrovò la stessa curva studiando gli errori delle misure astronomiche, e da lui prende il nome più diffuso.

COLLEGAMENTI · FISICA: GLI ERRORI DI MISURA

In laboratorio, ripetendo più volte la stessa misura (il periodo di un pendolo con un cronometro, per esempio) si ottengono valori diversi, sparsi attorno a un valore centrale. Se gli errori sono *casuali*, cioè dovuti a molte piccole cause indipendenti che spostano la misura ora in un senso ora nell'altro, la loro distribuzione è approssimativamente normale: è la «teoria degli errori» di Gauss. Per questo si dà come risultato il valore medio con la sua incertezza, e la deviazione standard misura la precisione dello strumento e del metodo. Gli errori *sistematici* (un cronometro che ritarda, una bilancia mal tarata) spostano invece tutte le misure nello stesso senso: nessuna media li elimina, e la dispersione delle misure non li rivela: si scoprono soltanto confrontando lo strumento con un riferimento.

11.5 Modelli probabilistici

11.5.1 Una distribuzione è un modello

Quando diciamo che il numero di pezzi difettosi «è» binomiale o che le altezze «sono» normali, stiamo scegliendo un **modello probabilistico**: una descrizione matematica semplificata del fenomeno, che poggia su ipotesi. La binomiale richiede prove *indipendenti* e una probabilità di successo *costante*: se i difetti nascono da un guasto che colpisce intere serie di pezzi consecutivi, l'indipendenza cade e il modello sottovaluta la probabilità di scatole con molti difetti. La distribuzione uniforme dell'attesa richiede che l'orario di arrivo alla fermata non sia legato a quello dell'autobus: chi conosce l'orario aspetta molto meno. Un modello si giudica confrontando le sue previsioni con i dati, e lo stesso fenomeno può ammettere modelli diversi, più o meno adatti a seconda della domanda che si pone.

11.5.2 Quando la normale non va bene

La distribuzione normale è così diffusa che si rischia di vederla dappertutto. Ma molte grandezze non sono simmetriche. I redditi delle famiglie di un paese, per esempio, hanno una distribuzione asimmetrica, con molti redditi medio-bassi e una lunga coda di redditi molto alti: la media è maggiore della mediana, e un modello normale darebbe probabilità positive a redditi negativi e sottovaluterebbe gravemente i redditi molto alti. Anche i tempi di attesa, che non possono essere negativi e hanno spesso una coda lunga a destra, si descrivono meglio con altre distribuzioni, come l'esponenziale. Due controlli semplici: la media e la mediana dei dati sono vicine? Le percentuali di dati entro una, due e tre deviazioni standard sono vicine al 68, 95 e 99,7 per cento?

11.5.3 Le medie dei campioni

C'è però una ragione profonda per cui la normale compare così spesso. In *Matematica 4*, cap. 11, abbiamo visto che da un campione si stima una caratteristica della popolazione. Se si estraggono a caso molti campioni di n elementi e di ciascuno si calcola la media, le medie dei campioni variano da un campione all'altro; ma si addensano attorno alla media della popolazione, con una deviazione standard $\frac{\sigma}{\sqrt{n}}$ che diminuisce all'aumentare di n (è la stessa radice incontrata nella legge dei grandi numeri), e, per n abbastanza grande, la loro distribuzione è approssimativamente normale *anche quando la popolazione non lo è*. È l'idea del **teorema del limite centrale**, che si dimostra all'università: una grandezza che nasce dalla somma di molti piccoli contributi indipendenti tende ad avere una distribuzione normale. La binomiale, somma di n indicatrici, ne è un caso particolare, e gli errori casuali di misura un altro.

COLLEGAMENTI · FILOSOFIA: CASO E DETERMINISMO

All'inizio dell'Ottocento Pierre-Simon Laplace, che contribuì in modo decisivo alla teoria delle probabilità e ai risultati sulla curva normale, era anche uno dei più convinti sostenitori del determinismo (lo abbiamo ricordato nel capitolo 6): un'intelligenza che conoscesse tutte le forze e le posizioni dei corpi in un istante vedrebbe il futuro come il passato. Per Laplace non c'è contraddizione: la probabilità non misura un caso presente nella natura, ma la *nostra ignoranza* delle cause. Il lancio di una moneta è governato dalle leggi della meccanica; lo diciamo casuale perché non conosciamo con precisione la spinta, la rotazione, l'aria.

Tra la fine dell'Ottocento e il Novecento il quadro è cambiato in due modi. Lo studio dei sistemi sensibili alle condizioni iniziali, avviato da Poincaré negli ultimi anni dell'Ottocento, ha mostrato che anche leggi deterministiche possono produrre comportamenti di fatto imprevedibili. E la meccanica quantistica descrive alcuni fenomeni, come l'istante in cui decade un singolo nucleo radioattivo, con leggi che danno soltanto probabilità; per l'interpretazione più diffusa, non per un difetto delle nostre conoscenze. Che il caso sia nelle cose o soltanto nella nostra mente è una domanda filosofica che la matematica delle distribuzioni non risolve: funziona allo stesso modo in entrambi i casi.

Infine, un esempio dall'economia. Due investimenti (dati inventati) promettono lo stesso rendimento atteso del 3% annuo; il primo ha una deviazione standard del rendimento dell'1%, il secondo dell'8%. Il valore atteso è lo stesso, ma con il secondo un anno con rendimento negativo è tutt'altro che raro, mentre con il primo è un evento a tre deviazioni standard dalla media. In finanza la deviazione standard è una prima misura del **rischio**: due distribuzioni con la stessa media possono descrivere scelte molto diverse, e la scelta tra loro dipende da quanto rischio si è disposti a correre.

SINTESI

- Una variabile aleatoria associa un numero a ogni esito di un esperimento casuale; è discreta se i suoi valori si possono elencare, continua se riempiono un intervallo.
- La distribuzione di una variabile discreta assegna a ogni valore x_i la probabilità p_i, con $p_i \geq 0$ e $\Sigma\, p_i = 1$; la funzione di ripartizione $F(x) = P(X \leq x)$ è una scala non decrescente da 0 a 1.
- Valore atteso $\mu = \Sigma\, x_i\, p_i$; varianza $\sigma^2 = \Sigma\, (x_i - \mu)^2\, p_i = E(X^2) - \mu^2$; deviazione standard σ. Inoltre $E(aX + b) = aE(X) + b$ e $\mathrm{Var}(aX + b) = a^2\, \mathrm{Var}(X)$.
- Il numero di successi in n prove indipendenti con probabilità p ha distribuzione binomiale $B(n; p)$: $P(X = k) = C_{n,k}\, p^k\, (1 - p)^{n-k}$, con $\mu = np$ e $\sigma^2 = np(1 - p)$.
- Legge dei grandi numeri: la frequenza relativa dei successi si stringe attorno a p, con deviazione standard $\sqrt{\dfrac{p(1-p)}{n}}$. Le prove indipendenti non hanno memoria: il caso non compensa.
- Una variabile continua ha una densità $f \geq 0$ con area totale 1: $P(a \leq X \leq b) = \int_a^b f(x)\, dx$ e $P(X = c) = 0$. Il valore $f(x)$ non è una probabilità.
- Uniforme in $[a; b]$: $f = \dfrac{1}{b-a}$, $\mu = \dfrac{a+b}{2}$, $\sigma^2 = \dfrac{(b-a)^2}{12}$.
- Normale $N(\mu; \sigma)$ (il secondo parametro è la deviazione standard): campana simmetrica, massimo in μ, flessi in $\mu \pm \sigma$, area totale 1.
- Standardizzazione $z = \dfrac{x - \mu}{\sigma}$ e tabella di $\Phi(z) = P(Z \leq z)$, con $\Phi(-z) = 1 - \Phi(z)$. Regola 68–95–99,7: entro σ, 2σ, 3σ dalla media cade il 68,3%, il 95,4%, il 99,7% dei valori.
- Per n grande la binomiale $B(n; p)$ è approssimata dalla normale $N(np; \sqrt{np(1-p)})$ (de Moivre, Settecento); le medie dei campioni tendono a essere normali (idea del limite centrale).
- Una distribuzione è un modello: poggia su ipotesi (indipendenza, probabilità costante, simmetria) da controllare sui dati. La normale non descrive bene grandezze asimmetriche come i redditi.

ESERCIZI

Comprendere

1. ● Una variabile aleatoria X assume i valori 1, 2, 3 con probabilità 0,2, 0,5 e p. Quanto vale p?

   A. 0,2

   B. 0,3

   C. 0,5

   D. Non si può stabilire

2. ● Se $X \sim B(20; 0{,}25)$, il valore atteso e la varianza di X sono

   A. 4 e 3
   B. 5 e 1,94
   C. 5 e 15
   D. 5 e 3,75

3. ● X è una variabile aleatoria continua con densità f. La probabilità $P(X = 2)$ vale

   A. 0
   B. $f(2)$
   C. $F(2)$, il valore della funzione di ripartizione
   D. dipende dalla densità

4. ● La densità della distribuzione normale $N(\mu; \sigma)$ ha i flessi in

   A. $x = \mu \pm 2\sigma$
   B. $x = \mu \pm \frac{\sigma}{2}$
   C. $x = \mu \pm \sigma$
   D. $x = \mu$ e $x = \mu + \sigma$

5. ● $X \sim N(50; 10)$. La probabilità $P(40 < X < 60)$ è circa

   A. 34,1%
   B. 68,3%
   C. 95,4%
   D. 50%

6. ● Se $\text{Var}(X) = 4$, quanto vale $\text{Var}(3X + 5)$?

   A. 17
   B. 12
   C. 41
   D. 36

7. ● Vero o falso? Correggi le affermazioni false. a) La funzione di ripartizione di una variabile discreta è costante a tratti e non decrescente. b) Il valore atteso di una variabile aleatoria è sempre uno dei suoi valori possibili. c) Una densità di probabilità non può mai superare 1. d) La distribuzione $B(n; 0{,}5)$ è simmetrica rispetto a $\frac{n}{2}$.

8. ● Completa. La densità di probabilità di una variabile continua è una funzione f tale che $f(x)$ _____ 0 per ogni x e l'area totale sotto il suo grafico è uguale a _____ ; la probabilità che X cada nell'intervallo $[a; b]$ è l' _____ $\int_a^b f(x)\,dx$. Nella distribuzione normale standard, la media è _____ e la deviazione standard è _____ .

Calcolare

9. ● Si lancia due volte una moneta equilibrata; X è il numero di teste. Scrivi la distribuzione di X, calcola $E(X)$ e $\text{Var}(X)$ e confronta con le formule della binomiale.

10. ●● La variabile X assume i valori 0, 1, 2, 3 con probabilità 0,1; 0,3; 0,4; 0,2. a) Scrivi la funzione di ripartizione e disegnane il grafico. b) Calcola $P(1 < X \le 3)$. c) Calcola $E(X)$, $\text{Var}(X)$ e σ.

11. ●● Per partecipare a un gioco si pagano 3 €; si lancia un dado e si riceve in euro il numero uscito. a) Calcola il valore atteso del guadagno netto G. b) Il gioco è equo? Quale prezzo lo renderebbe equo? c) Calcola la deviazione standard di G.

12. ● Si lancia 5 volte un dado; X è il numero di volte in cui esce 6. Calcola a) $P(X = 2)$; b) $P(X \ge 1)$; c) valore atteso e deviazione standard di X.

13. ●● Una cestista realizza un tiro libero con probabilità 0,8, indipendentemente dagli altri tiri. Su 10 tiri calcola a) la probabilità di realizzarne esattamente 8; b) la probabilità di realizzarne almeno 9; c) il valore atteso e la deviazione standard del numero di canestri.

14. ●● X ha distribuzione uniforme in [2; 10]. Scrivi la densità e calcola $P(3 \le X \le 6)$, $P(X > 9)$, $E(X)$ e σ.

15. ●● Verifica che $f(x) = \frac{3}{8}x^2$ per $0 \le x \le 2$ (e $f(x) = 0$ altrove) è una densità di probabilità. Calcola poi $P(X \le 1)$, $E(X)$ e $\text{Var}(X)$.

16. ●● Trova la costante k per cui $f(x) = kx$ in [0; 4] (e 0 altrove) è una densità. Calcola $P(1 \le X \le 3)$ e il valore m tale che $P(X \le m) = \frac{1}{2}$ (la *mediana*).

17. ● Z è una variabile normale standard. Con la tabella di Φ calcola a) $P(Z \le 1{,}4)$; b) $P(Z > 0{,}8)$; c) $P(Z \le -1{,}5)$; d) $P(-0{,}5 \le Z \le 1)$.

18. ●● $X \sim N(80; 5)$. Calcola a) $P(X \le 90)$; b) $P(X < 72)$; c) $P(75 \le X \le 88)$; d) il valore x per cui $P(X \le x) = 0{,}9332$.

Dimostrare e argomentare

19. ●● Dimostra che, per una variabile discreta X con valore atteso μ e per due numeri a, b, si ha $E(aX + b) = aE(X) + b$ e $\text{Var}(aX + b) = a^2\,\text{Var}(X)$.

20. ●● Dimostra che $\text{Var}(X) = E(X^2) - \mu^2$.

21. ●● Trova l'errore. «In una serie di lanci di una moneta equilibrata è uscita croce otto volte di fila. Al prossimo lancio testa è più probabile di croce, perché per la legge dei grandi numeri le teste devono recuperare.»

22. ●● Sia $f(x) = c \cdot e^{-(x-\mu)^2/(2\sigma^2)}$, con $c = \frac{1}{\sigma\sqrt{2\pi}}$. Calcola $f'(x)$ e $f''(x)$ e dimostra che f ha il massimo in $x = \mu$ e i flessi in $x = \mu \pm \sigma$.

Problemi

23. ●● Un test ha 10 domande a scelta multipla, ciascuna con 4 risposte di cui una sola corretta. Uno studente risponde a caso a tutte. a) Qual è la distribuzione del numero X di risposte corrette? b) Calcola valore atteso e deviazione standard. c) Il test si supera con almeno 6 risposte corrette: qual è la probabilità di superarlo rispondendo a caso?

24. ●● Una compagnia aerea vende 52 biglietti per un volo con 50 posti, perché ciascun passeggero, indipendentemente dagli altri, si presenta all'imbarco con probabilità 0,9 (dati inventati). a) Quanti passeggeri si presentano in media? b) Qual è la probabilità che qualcuno resti a terra, cioè che si presentino più di 50 passeggeri? c) Quale ipotesi del modello potrebbe essere falsa nella realtà?

25. ●● Uno strumento misura una lunghezza con errori casuali distribuiti come N(0; 0,2), in millimetri. a) Qual è la probabilità che una misura sbagli di più di 0,4 mm, in un senso o nell'altro? b) E di più di 0,3 mm? c) Che cosa cambierebbe se lo strumento avesse anche un errore sistematico di +0,1 mm?

26. ●● La durata delle batterie di un certo tipo è distribuita come N(400; 40), in ore (dati inventati). a) Quale frazione delle batterie dura meno di 340 ore? b) Su un lotto di 1000 batterie, quante ci si aspetta che durino meno di 340 ore? c) Il produttore vuole garantire una durata minima tale che solo il 2,3% circa delle batterie non la raggiunga: quale durata deve garantire?

Con il computer

27. ●● Con un foglio di calcolo simula 1000 lanci di un dado (una funzione che genera un intero casuale da 1 a 6 in ogni cella di una colonna). a) In una seconda colonna calcola la frequenza relativa del 6 dopo i primi n lanci e rappresentala in un grafico al variare di n: che cosa osservi? b) Simula ora 1000 lanci di *due* dadi, costruisci il diagramma delle frequenze relative della somma e confrontalo con la Tabella 11.1. c) Calcola la media e la deviazione standard delle 1000 somme e confrontale con 7 e 2,42.

Verso l'esame

28. ●● Nello stile delle prove nazionali. Il numero X di clienti che entrano in un negozio in 5 minuti ha la seguente distribuzione (dati inventati): $P(X = 0) = 0{,}15$; $P(X = 1) = 0{,}30$; $P(X = 2) = 0{,}25$; $P(X = 3) = 0{,}20$; $P(X = 4) = 0{,}10$. Quale affermazione è corretta?

    A. Il valore atteso di X è 2.

    B. Il valore più probabile di X è 2.

    C. $P(X \geq 2) = 0{,}55$.

    D. $F(1) = 0{,}30$.

29. ●● Nello stile delle prove nazionali. Il punteggio (in centesimi) ottenuto da 2000 candidati in una prova di ammissione è distribuito approssimativamente come N(60; 12) (dati simulati). a) Circa quanti candidati hanno ottenuto più di 84? b) Quale punteggio minimo serve, circa, per essere nel 16% migliore? c) Perché il modello normale non può essere esatto per questi punteggi? È comunque ragionevole?

30. ●● Spunto per il colloquio. Laplace, all'inizio dell'Ottocento, sostenne in sostanza che la probabilità dipende in parte dalla nostra ignoranza e in parte dalle nostre conoscenze (parafrasi). Spiega che cosa intendeva, mostra con un esempio come una distribuzione di probabilità renda prevedibile l'insieme di molti eventi singolarmente imprevedibili, e collega il tema alla fisica, alla filosofia e alla letteratura.

31. ●●● Sfida. Sia $X \sim \mathrm{B}(n; p)$ con $0 < p < 1$. a) Dimostra che $\dfrac{\mathrm{P}(X = k+1)}{\mathrm{P}(X = k)} = \dfrac{(n-k)p}{(k+1)(1-p)}$. b) Deduci che $\mathrm{P}(X = k+1) > \mathrm{P}(X = k)$ se e solo se $k < (n+1)p - 1$, e quindi che, quando $(n+1)p$ non è intero, il valore più probabile di X è la parte intera di $(n+1)p$. c) Verifica il risultato sulle tre distribuzioni della Figura 11.2.

PER APPROFONDIRE

- [FEL] Feller, *An Introduction to Probability Theory and Its Applications*, vol. I: le prove di Bernoulli, la distribuzione binomiale, l'approssimazione normale e la legge dei grandi numeri, con moltissimi esempi.
- [BOY] Boyer, *Storia della matematica*: le pagine sul calcolo delle probabilità da Jakob Bernoulli a de Moivre e Laplace, e su Gauss.

Glossario del capitolo

Densità di probabilità
Funzione $f \geq 0$ con area totale 1 tale che $P(a \leq X \leq b) = \int_a^b f(x)\,dx$.

Deviazione standard
Radice quadrata della varianza; misura, nella stessa unità di X, la distanza tipica dei valori dalla media.

Distribuzione binomiale
Distribuzione $B(n; p)$ del numero di successi in n prove indipendenti con probabilità di successo p; $\mu = np$, $\sigma^2 = np(1-p)$.

Distribuzione di probabilità
Corrispondenza che associa a ogni valore x_i di una variabile discreta la probabilità $p_i = P(X = x_i)$; le probabilità hanno somma 1.

Distribuzione normale
Distribuzione continua $N(\mu; \sigma)$ con densità a campana, simmetrica attorno a μ, con flessi in $\mu \pm \sigma$.

Distribuzione uniforme
Distribuzione continua con densità costante $\frac{1}{b-a}$ in $[a; b]$; $\mu = \frac{a+b}{2}$, $\sigma^2 = \frac{(b-a)^2}{12}$.

Fallacia del giocatore
L'errore di credere che in prove indipendenti un esito «in ritardo» diventi più probabile.

Funzione di ripartizione
La funzione $F(x) = P(X \leq x)$; per una variabile discreta è una scala non decrescente da 0 a 1.

Legge dei grandi numeri
Al crescere del numero di prove, la probabilità che la frequenza relativa di un evento si discosti dalla sua probabilità più di un margine fissato tende a zero (Jakob Bernoulli, 1713).

Modello probabilistico
Descrizione di un fenomeno casuale mediante una distribuzione, basata su ipotesi da confrontare con i dati.

Normale standard
La distribuzione $N(0; 1)$; la sua funzione di ripartizione $\Phi(z)$ è tabulata.

Rischio
In finanza, la variabilità del rendimento di un investimento, misurata in prima approssimazione dalla deviazione standard.

Standardizzazione
Il passaggio da $X \sim N(\mu; \sigma)$ a $Z = \frac{X - \mu}{\sigma} \sim N(0; 1)$.

Teorema del limite centrale
Risultato secondo cui la somma (o la media) di molte variabili indipendenti ha distribuzione approssimativamente normale.

Valore atteso
Media dei valori di una variabile aleatoria pesati con le loro probabilità: $E(X) = \Sigma\, x_i\, p_i$, o $\int x\, f(x)\,dx$ nel caso continuo.

Variabile aleatoria
Regola che associa un numero reale a ogni esito di un esperimento casuale; si indica con una lettera maiuscola.

Variabile aleatoria continua
Variabile aleatoria che può assumere tutti i valori di un intervallo; ogni singolo valore ha probabilità 0.

Variabile aleatoria discreta
Variabile aleatoria i cui valori sono in numero finito o si possono elencare uno dopo l'altro.

Variabile indicatrice
Variabile che vale 1 se un evento si verifica e 0 altrimenti; il suo valore atteso è la probabilità dell'evento.

Varianza
Valore atteso del quadrato dello scarto dalla media: $\mathrm{Var}(X) = E((X-\mu)^2) = E(X^2) - \mu^2$.

Soluzioni

1. B. Le probabilità hanno somma 1: $p = 1 - 0{,}2 - 0{,}5 = 0{,}3$.

2. D. $\mu = np = 20 \cdot 0{,}25 = 5$; $\sigma^2 = np(1-p) = 5 \cdot 0{,}75 = 3{,}75$ (1,94 è circa la deviazione standard).

3. A. Per una variabile continua la probabilità di un singolo valore è un'area di base nulla, quindi è 0.

4. C. La derivata seconda cambia segno in $x = \mu \pm \sigma$ (esercizio 22).

5. B. L'intervallo è $\mu \pm \sigma$: la regola 68–95–99,7 dà circa 68,3%.

6. D. $\text{Var}(3X + 5) = 3^2 \cdot \text{Var}(X) = 9 \cdot 4 = 36$: la costante 5 non cambia la dispersione.

7. a) Vero. b) Falso: il numero di teste in tre lanci ha valore atteso 1,5, che non è un valore possibile. c) Falso: la densità uniforme in [0; 0,5] vale 2; è l'area, non il valore di f, a essere una probabilità. d) Vero: $C_{n,k} = C_{n,n-k}$ e con $p = 1 - p = 0{,}5$ si ha $P(X = k) = P(X = n - k)$.

8. ≥; 1; integrale (l'area sotto il grafico tra a e b); 0; 1.

9. X vale 0, 1, 2 con probabilità $\frac{1}{4}$, $\frac{1}{2}$, $\frac{1}{4}$.

$E(X) = \frac{1}{2} + 2 \cdot \frac{1}{4} = 1$; $E(X^2) = \frac{1}{2} + 4 \cdot \frac{1}{4} = \frac{3}{2}$, $\text{Var}(X) = \frac{3}{2} - 1 = \frac{1}{2}$. Con $B(2; \frac{1}{2})$: $np = 1$ e $np(1-p) = \frac{1}{2}$.

10. a) $F(x) = 0$ per $x < 0$; 0,1 per $0 \le x < 1$; 0,4 per $1 \le x < 2$; 0,8 per $2 \le x < 3$; 1 per $x \ge 3$ (una scala con gradini di altezza 0,1, 0,3, 0,4, 0,2). b) $F(3) - F(1) = 1 - 0{,}4 = 0{,}6$. c) $E(X) = 0{,}3 + 0{,}8 + 0{,}6 = 1{,}7$; $E(X^2) = 0{,}3 + 1{,}6 + 1{,}8 = 3{,}7$; $\text{Var}(X) = 3{,}7 - 1{,}7^2 = 0{,}81$; $\sigma = 0{,}9$.

11. a) Detto X il punteggio, $E(X) = \frac{21}{6} = 3{,}5$ e $G = X - 3$, quindi $E(G) = 0{,}5$ €. b) No, è favorevole al giocatore; è equo con un prezzo di 3,50 €. c) $E(X^2) = \frac{91}{6}$, $\text{Var}(X) = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$; $\text{Var}(G) = \text{Var}(X)$, quindi $\sigma = \sqrt{\frac{35}{12}} \approx 1{,}71$ €.

12. $X \sim B(5; \frac{1}{6})$. a) $C_{5,2} (\frac{1}{6})^2 (\frac{5}{6})^3 = \frac{10 \cdot 125}{7776} = \frac{625}{3888} \approx 0{,}161$. b) $1 - (\frac{5}{6})^5 = 1 - \frac{3125}{7776} = \frac{4651}{7776} \approx 0{,}598$. c) $\mu = \frac{5}{6} \approx 0{,}83$; $\sigma^2 = 5 \cdot \frac{1}{6} \cdot \frac{5}{6} = \frac{25}{36}$, $\sigma = \frac{5}{6} \approx 0{,}83$.

13. $X \sim B(10; 0{,}8)$. a) $C_{10,8} \cdot 0{,}8^8 \cdot 0{,}2^2 = 45 \cdot 0{,}8^8 \cdot 0{,}04 \approx 0{,}302$. b) $10 \cdot 0{,}8^9 \cdot 0{,}2 + 0{,}8^{10} \approx 0{,}2684 + 0{,}1074 \approx 0{,}376$. c) $\mu = 8$, $\sigma = \sqrt{10 \cdot 0{,}8 \cdot 0{,}2} = \sqrt{1{,}6} \approx 1{,}26$.

14. $f(x) = \frac{1}{8}$ in [2; 10], 0 fuori. $P(3 \le X \le 6) = \frac{3}{8}$; $P(X > 9) = \frac{1}{8}$; $E(X) = 6$; $\sigma^2 = \frac{8^2}{12} = \frac{16}{3}$, $\sigma \approx 2{,}31$.

15. $f \ge 0$ e $\int_0^2 \frac{3}{8} x^2 \, dx = [\frac{x^3}{8}]_0^2 = 1$. $P(X \le 1) = \frac{1}{8}$. $E(X) = \int_0^2 \frac{3}{8} x^3 \, dx = [\frac{3x^4}{32}]_0^2 = \frac{3}{2}$; $E(X^2) = \int_0^2 \frac{3}{8} x^4 \, dx = [\frac{3x^5}{40}]_0^2 = \frac{12}{5}$; $\text{Var}(X) = \frac{12}{5} - \frac{9}{4} = \frac{3}{20} = 0{,}15$.

16. $\int_0^4 kx \, dx = 8k = 1$, quindi $k = \frac{1}{8}$. $P(1 \le X \le 3) = [\frac{x^2}{16}]_1^3 = \frac{9-1}{16} = \frac{1}{2}$. $\frac{m^2}{16} = \frac{1}{2}$ dà $m = 2\sqrt{2} \approx 2{,}83$.

17. a) 0,9192. b) $1 - 0{,}7881 = 0{,}2119$. c) $1 - 0{,}9332 = 0{,}0668$. d) $\Phi(1) - \Phi(-0{,}5) = 0{,}8413 - (1 - 0{,}6915) = 0{,}5328$.

18. a) $z = 2$: 0,9772. b) $z = -1{,}6$: $1 - 0{,}9452 = 0{,}0548$. c) $\Phi(1{,}6) - \Phi(-1) = 0{,}9452 - 0{,}1587 = 0{,}7865$. d) $\Phi(1{,}5) = 0{,}9332$, quindi $x = 80 + 1{,}5 \cdot 5 = 87{,}5$.

19. $E(aX + b) = \Sigma (a x_i + b) p_i = a \Sigma x_i p_i + b \Sigma p_i = a\mu + b$, perché $\Sigma p_i = 1$. Gli scarti di $aX + b$ dalla sua media sono $a x_i + b - (a\mu + b) = a(x_i - \mu)$, quindi $\text{Var}(aX + b) = \Sigma a^2 (x_i - \mu)^2 p_i = a^2 \text{Var}(X)$. c.v.d.

20. Sviluppando il quadrato: $\Sigma\,(x_i - \mu)^2\,p_i = \Sigma\,x_i^2\,p_i - 2\mu\,\Sigma\,x_i\,p_i + \mu^2\,\Sigma\,p_i = \mathrm{E}(X^2) - 2\mu\cdot\mu + \mu^2 = \mathrm{E}(X^2) - \mu^2$. c.v.d.

21. I lanci sono indipendenti: la probabilità di testa al nono lancio è ancora $\frac{1}{2}$ (fallacia del giocatore). La legge dei grandi numeri riguarda la *frequenza relativa* su moltissime prove: lo squilibrio di 8 croci non viene compensato, ma diventa trascurabile rispetto al numero totale dei lanci.

22. $f'(x) = -c\,\frac{x-\mu}{\sigma^2}\,e^{-(x-\mu)^2/(2\sigma^2)}$: positiva per $x < \mu$, negativa per $x > \mu$, quindi massimo in μ. $f''(x) = \frac{c}{\sigma^2}\left(\frac{(x-\mu)^2}{\sigma^2} - 1\right)e^{-(x-\mu)^2/(2\sigma^2)}$: è negativa per $|x-\mu| < \sigma$ e positiva per $|x-\mu| > \sigma$, quindi cambia segno (flessi) in $x = \mu \pm \sigma$. c.v.d.

23. a) $X \sim \mathrm{B}(10;\,0{,}25)$. b) $\mu = 2{,}5$, $\sigma = \sqrt{1{,}875} \approx 1{,}37$. c) $\mathrm{P}(X \geq 6) = \Sigma_{k=6}^{10}\,\mathrm{C}_{10,k}\,0{,}25^k\,0{,}75^{10-k} \approx 0{,}0162 + 0{,}0031 + 0{,}0004 + \ldots$ (gli ultimi due termini sono minori di 0,0001), cioè circa 0,020: circa il 2%.

24. $X \sim \mathrm{B}(52;\,0{,}9)$. a) $\mu = 46{,}8$. b) $\mathrm{P}(X = 51) + \mathrm{P}(X = 52) = 52\cdot 0{,}9^{51}\cdot 0{,}1 + 0{,}9^{52} \approx 0{,}0241 + 0{,}0042 \approx 0{,}028$, circa il 2,8%. c) L'indipendenza: famiglie o gruppi che viaggiano insieme si presentano (o mancano) tutti insieme.

25. a) $0{,}4 = 2\sigma$: $1 - 0{,}954 = 0{,}046$, circa il 4,6%. b) $z = 1{,}5$: $2(1 - \Phi(1{,}5)) = 2\cdot 0{,}0668 = 0{,}1336$, circa il 13%. c) Gli errori diventerebbero $\mathrm{N}(0{,}1;\,0{,}2)$: la campana si sposta di 0,1 mm, tutte le misure tendono a eccedere, e la media di molte misure non elimina l'errore.

26. a) $z = -1{,}5$: $\Phi(-1{,}5) = 0{,}0668$, circa il 6,7%. b) $1000\cdot 0{,}0668 \approx 67$ batterie. c) $\Phi(-2) = 0{,}0228$: $400 - 2\cdot 40 = 320$ ore.

27. Traccia: a) la frequenza oscilla molto per n piccolo e poi si stabilizza attorno a $\frac{1}{6} \approx 0{,}167$ (legge dei grandi numeri); lo scarto tipico è $\sqrt{\frac{1}{6}\cdot\frac{5}{6}\cdot\frac{1}{n}}$, circa 0,012 per $n = 1000$. b) Il diagramma ha la forma «a tetto» con il massimo in 7; le frequenze differiscono dalle probabilità di qualche millesimo o centesimo. c) La media è vicina a 7 e la deviazione standard a $\sqrt{\frac{35}{6}} \approx 2{,}42$; ripetendo la simulazione i valori cambiano leggermente.

28. C. $\mathrm{P}(X \geq 2) = 0{,}25 + 0{,}20 + 0{,}10 = 0{,}55$. A è falsa: $\mathrm{E}(X) = 0{,}30 + 0{,}50 + 0{,}60 + 0{,}40 = 1{,}8$. B è falsa: il valore più probabile è 1. D è falsa: $F(1) = 0{,}15 + 0{,}30 = 0{,}45$.

29. a) $z = 2$: $1 - 0{,}9772 = 0{,}0228$, e $0{,}0228\cdot 2000 \approx 46$ candidati. b) $1 - \Phi(1) = 0{,}1587 \approx 16\%$: serve $60 + 12 = 72$. c) I punteggi sono interi e compresi tra 0 e 100, mentre la normale dà probabilità a ogni numero reale; ma 0 e 100 distano 5 e 3,3 deviazioni standard dalla media, quindi la probabilità assegnata ai valori impossibili è trascurabile e il modello è ragionevole.

30. Traccia: 1) il determinismo di Laplace (capitolo 6): la probabilità come misura della nostra ignoranza delle cause, non del caso nella natura; 2) l'esempio della moneta o della binomiale: il singolo lancio è imprevedibile, la frequenza su molti lanci si stringe attorno a p (legge dei grandi numeri, Jakob Bernoulli, 1713); 3) la curva normale e la teoria degli errori di Gauss: la regolarità che nasce da molte piccole cause indipendenti; 4) fisica: errori di misura; nel Novecento la meccanica quantistica e il decadimento radioattivo come leggi intrinsecamente probabilistiche, i sistemi caotici studiati a partire da Poincaré; 5) filosofia: il positivismo e la fiducia nella previsione scientifica, la crisi del determinismo nel Novecento; 6) letteratura: il caso e la fortuna nel romanzo, per esempio la vincita alla roulette di Montecarlo che cambia la vita del protagonista in *Il fu Mattia Pascal* di Pirandello.

31. a) $\frac{\mathrm{C}_{n,k+1}}{\mathrm{C}_{n,k}} = \frac{n!}{(k+1)!\,(n-k-1)!}\cdot\frac{k!\,(n-k)!}{n!} = \frac{n-k}{k+1}$, e il rapporto delle potenze è $\frac{p}{1-p}$. b) Il rapporto supera 1 se e solo se $(n-k)p > (k+1)(1-p)$, cioè $np - kp > k + 1 - kp - p$, cioè $k < (n+1)p - 1$. Le probabilità crescono finché $k + 1 < (n+1)p$ e poi decrescono: il massimo è nel più grande intero minore di $(n+1)p$, cioè la sua parte intera. c) $\mathrm{B}(10;\,0{,}3)$: $(n+1)p = 3{,}3$, moda 3; $\mathrm{B}(10;\,0{,}5)$: 5,5, moda 5; $\mathrm{B}(50;\,0{,}3)$: 15,3, moda 15. Coincidono con le barre più alte della figura.