CAPITOLO

15

Variabili aleatorie discrete e distribuzioni

Il caso non ha memoria, ma ha una media: è da lì che si comincia a prevederlo. (testo della redazione)

OBIETTIVI

1. Definire una variabile aleatoria discreta e rappresentarne la distribuzione di probabilità e la funzione di ripartizione.
2. Calcolare valore atteso, varianza e deviazione standard e dimostrarne le proprietà per le trasformazioni lineari.
3. Riconoscere e usare le distribuzioni uniforme, binomiale e di Poisson, con il loro valore atteso e la loro varianza.
4. Ottenere la distribuzione di Poisson come limite della binomiale e valutare numericamente l'approssimazione.
5. Applicare la disuguaglianza di Čebyšev e dimostrare con essa la legge debole dei grandi numeri.
6. Scegliere il modello adatto a una situazione e controllarne le ipotesi.

INDICE DEL CAPITOLO

15.1 Variabili aleatorie discrete 2
15.2 Valore atteso, varianza e deviazione standard 6
15.3 Le distribuzioni uniforme e binomiale 9
15.4 La distribuzione di Poisson 13
15.5 La legge dei grandi numeri 17
15.6 Scegliere il modello 19

Matematica 5 · Parte V · Distribuzioni di probabilità · versione 1.0
Capitolo 15: Variabili aleatorie discrete e distribuzioni · pagine 427–452 del volume completo · estratto per lo studio

Un centralino riceve in media due o tre chiamate al minuto; un lotto di pezzi contiene di solito uno o due difettosi; in un test a risposta multipla chi tira a indovinare azzecca, in media, una risposta su quattro. In ciascuna di queste situazioni l'esito è incerto, ma è un **numero**: quante chiamate, quanti difettosi, quante risposte giuste. E di un numero aleatorio vorremmo sapere molto più della probabilità di un singolo evento: quali valori può assumere e con quali probabilità, quanto vale «in media», quanto si allontana di solito da quella media.

In *Matematica 4* abbiamo costruito il calcolo delle probabilità sugli eventi. In questo capitolo spostiamo l'attenzione sulle **variabili aleatorie**, cioè sui numeri prodotti dal caso, e sulle loro **distribuzioni**. Incontreremo tre distribuzioni che descrivono un'enorme varietà di fenomeni (l'uniforme, la binomiale e quella di Poisson), vedremo come la terza nasce dalla seconda con un passaggio al limite, e dimostreremo con la disuguaglianza di Čebyšev la legge dei grandi numeri, che finora avevamo solo osservato nelle simulazioni. L'ultimo paragrafo è dedicato alla scelta del modello: la domanda più importante, e la più trascurata, di ogni problema di probabilità.

15.1 Variabili aleatorie discrete

15.1.1 Che cosa sappiamo già

La Tabella 15.1 riassume i concetti di probabilità studiati negli anni precedenti, che in questo capitolo useremo senza dimostrarli di nuovo.

Tabella 15.1 Richiamo dei concetti di base (vedi *Matematica 4*, capitoli 13–15).

Concetto	Formula o enunciato
Definizione classica	$p(E)$ = casi favorevoli : casi possibili, se i casi sono equiprobabili
Definizione frequentista	$p(E) \approx$ frequenza relativa di E su un gran numero di prove ripetute nelle stesse condizioni
Definizione soggettiva	$p(E)$ = prezzo equo che un individuo coerente è disposto a pagare per ricevere 1 se E si verifica
Assiomi di Kolmogorov	$p(E) \geq 0$; $p(\Omega) = 1$; se $A \cap B = \varnothing$, $p(A \cup B) = p(A) + p(B)$
Evento contrario, unione	$p(\overline{E}) = 1 - p(E)$; $p(A \cup B) = p(A) + p(B) - p(A \cap B)$
Probabilità condizionata	$p(A \mid B) = \dfrac{p(A \cap B)}{p(B)}$, con $p(B) > 0$
Probabilità composta, indipendenza	$p(A \cap B) = p(B) \cdot p(A \mid B)$; A e B indipendenti se $p(A \cap B) = p(A) \cdot p(B)$
Formula di Bayes	$p(H_i \mid E) = \dfrac{p(H_i) \cdot p(E \mid H_i)}{p(E)}$, con $p(E)$ dalle probabilità totali
Prove ripetute (Bernoulli)	n prove indipendenti con successo di probabilità p: $p(k \text{ successi}) = \binom{n}{k} p^k (1-p)^{n-k}$

15.1.2 Un numero che dipende dal caso

Lanciamo due dadi e annotiamo la somma dei punti. Lo spazio campionario Ω è formato dalle 36 coppie ordinate $(a; b)$, con a e b da 1 a 6, tutte equiprobabili. La somma è un modo di associare a ogni coppia un numero: a (2; 5) associa 7, a (6; 6) associa 12. È dunque una **funzione** definita su Ω a valori reali.

Si chiama **variabile aleatoria** (o **variabile casuale**) una funzione $X\colon \Omega \to \mathbb{R}$ che associa un numero reale a ogni esito di un esperimento aleatorio. La variabile si dice **discreta** se assume un numero finito di valori $x_1, x_2, \ldots, x_n$, oppure un'infinità di valori che si possono elencare in una successione $x_1, x_2, x_3, \ldots$ (come i numeri naturali). Le variabili aleatorie si indicano con lettere maiuscole, X, Y, Z; i valori che assumono con le minuscole.

Il nome è un po' ingannevole: una variabile aleatoria non è né una variabile né qualcosa di aleatorio in sé, ma una funzione del tutto determinata. Il caso sta nell'esito ω dell'esperimento; una volta noto l'esito, il valore $X(\omega)$ è fissato. Sullo stesso spazio campionario si possono definire molte variabili diverse: per i due dadi, oltre alla somma, il punteggio massimo, la differenza fra i punti, il numero di sei usciti.

Scrivendo $X = x$ indichiamo l'**evento** formato dagli esiti ai quali X associa il valore x. Per la somma dei dadi, l'evento $X = 4$ è l'insieme $\{(1; 3), (2; 2), (3; 1)\}$, quindi $p(X = 4) = \dfrac{3}{36}$. Allo stes-

so modo $X \le 4$ è l'evento «la somma non supera 4» e $p(X \le 4) = \frac{6}{36}$.

15.1.3 La distribuzione di probabilità

La **distribuzione di probabilità** di una variabile aleatoria discreta è la funzione che a ogni valore x_i associa la probabilità $p_i = p(X = x_i)$. Si può dare con una formula, con una tabella o con un grafico, l'**istogramma** o diagramma a barre, nel quale sopra ogni valore x_i si alza una barra di altezza p_i. Poiché gli eventi $X = x_i$ sono incompatibili a due a due e la loro unione è Ω, per gli assiomi della probabilità valgono sempre due condizioni:

$$p_i \ge 0 \text{ per ogni } i \quad \text{e} \quad p_1 + p_2 + \ldots = 1.$$

Viceversa, ogni elenco di numeri non negativi con somma 1 è la distribuzione di qualche variabile aleatoria. Per la somma di due dadi, contando le coppie che danno ciascun valore, si trova $p(X = k) = \frac{6 - |k - 7|}{36}$ per $k = 2, 3, \ldots, 12$: la distribuzione cresce da $\frac{1}{36}$ a $\frac{6}{36}$ e poi decresce simmetricamente (Figura 15.1, a sinistra). Le probabilità hanno somma 1:

$$\frac{1+2+3+4+5+6+5+4+3+2+1}{36} = \frac{36}{36} = 1.$$

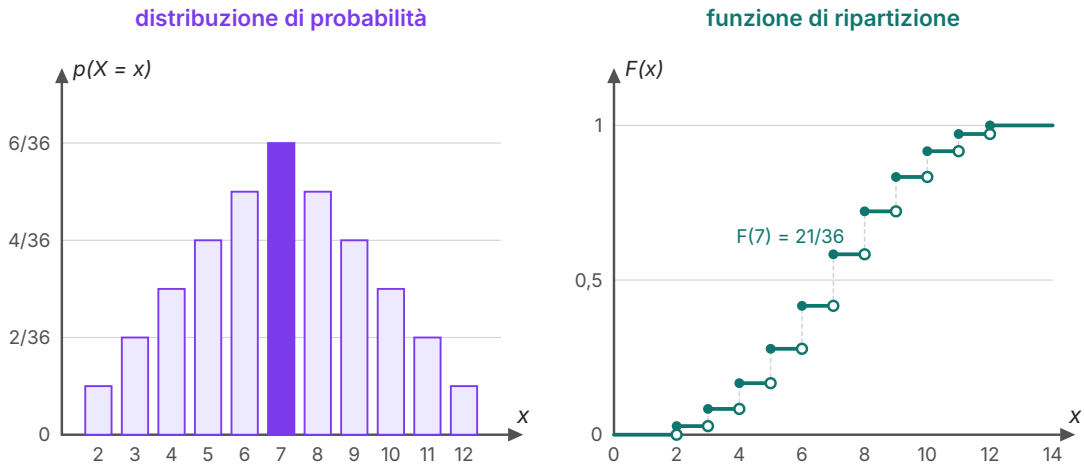


Figura 15.1 La somma X dei punti di due dadi. A sinistra la distribuzione di probabilità $p(X = x) = (6 - |x - 7|) : 36$, simmetrica intorno a 7. A destra la funzione di ripartizione $F(x) = p(X \le x)$: è costante a tratti e salta, in ogni valore x da 2 a 12, di un'altezza uguale alla barra corrispondente; il pallino pieno indica il valore assunto nel punto di salto.

15.1.4 La funzione di ripartizione

Spesso interessa la probabilità che X non superi una certa soglia: che i difettosi non siano più di due, che la somma dei dadi sia al massimo 7. Si chiama **funzione di ripartizione** di X la funzione definita per **ogni** numero reale x

$$F(x) = p(X \le x).$$

Per una variabile discreta, $F(x)$ è la somma delle probabilità dei valori $x_i \le x$. Il suo grafico è una **funzione a gradini** (Figura 15.1, a destra): vale 0 a sinistra del valore più piccolo, resta costante fra due valori consecutivi, in ogni x_i salta in su di p_i e, dopo il valore più grande, vale 1. Ne seguono le proprietà generali:

- $0 \le F(x) \le 1$, e F è **non decrescente**: se $x_1 < x_2$, l'evento $X \le x_1$ è contenuto in $X \le x_2$, quindi $F(x_1) \le F(x_2)$;
- $\lim_{x \to -\infty} F(x) = 0$ e $\lim_{x \to +\infty} F(x) = 1$;
- nei punti di salto F assume il valore del gradino **superiore**, perché $X \le x_i$ comprende $X = x_i$: F è continua da destra;
- per $a < b$, $p(a < X \le b) = F(b) - F(a)$.

L'ultima formula è la più utile, ma va letta con attenzione alle disuguaglianze: per una variabile discreta $p(X \le b)$ e $p(X < b)$ possono essere diverse. Con i dadi:

$$p(5 \le X \le 9) = p(4 < X \le 9) = F(9) - F(4) = \frac{30}{36} - \frac{6}{36} = \frac{24}{36} = \frac{2}{3},$$

che si ritrova sommando le barre da 5 a 9: $\dfrac{4+5+6+5+4}{36} = \dfrac{24}{36}$.

ESEMPIO SVOLTO

Il problema. Un'urna contiene 4 palline rosse e 6 bianche. Se ne estraggono 3 senza reimmissione e si indica con X il numero di palline rosse estratte. Determina la distribuzione di probabilità e la funzione di ripartizione di X, e calcola $p(X \geq 2)$.

I casi possibili. Le terne di palline (non ordinate) sono $\binom{10}{3} = 120$, tutte equiprobabili.

La distribuzione. Per avere k rosse occorre scegliere k rosse fra 4 e $3 - k$ bianche fra 6, quindi $p(X = k) = \binom{4}{k}\binom{6}{3-k} : 120$. Si ottiene

$$p(X=0) = \frac{20}{120} = \frac{1}{6}, \quad p(X=1) = \frac{60}{120} = \frac{1}{2}, \quad p(X=2) = \frac{36}{120} = \frac{3}{10},$$
$$p(X=3) = \frac{4}{120} = \frac{1}{30}.$$

Controllo: $20 + 60 + 36 + 4 = 120$, la somma è 1.

La funzione di ripartizione. Sommando: $F(x) = 0$ per $x < 0$; $F(x) = \frac{1}{6}$ per $0 \leq x < 1$; $F(x) = \frac{2}{3}$ per $1 \leq x < 2$; $F(x) = \frac{29}{30}$ per $2 \leq x < 3$; $F(x) = 1$ per $x \geq 3$.

La probabilità richiesta. $p(X \geq 2) = 1 - F(1) = 1 - \frac{2}{3} = \frac{1}{3}$, oppure $\frac{3}{10} + \frac{1}{30} = \frac{10}{30} = \frac{1}{3}$.

Una variabile di questo tipo (successi in estrazioni senza reimmissione da un'urna di composizione nota) si dice **ipergeometrica**; la ritroveremo nell'ultimo paragrafo.

15.2 Valore atteso, varianza e deviazione standard

15.2.1 Il valore atteso

Una distribuzione contiene tutta l'informazione su X, ma spesso vogliamo riassumerla in pochi numeri, come in statistica si riassume una distribuzione di frequenze con la media e la deviazione standard. Si chiama **valore atteso** (o **speranza matematica**, o **media**) di X il numero

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

indicato anche con μ. Se X assume infiniti valori, la somma diventa il limite delle somme parziali e si richiede che esista finito. Il valore atteso è una media ponderata dei valori, con pesi uguali alle probabilità. È la stessa formula della media di dati raggruppati di *Matematica 3*, $\overline{x} = \Sigma\, x_k f_k : n$, con le probabilità p_i al posto delle frequenze relative $f_k : n$; e poiché, per la legge dei grandi numeri, le frequenze relative si avvicinano alle probabilità, $\mathrm{E}(X)$ è il valore intorno al quale si stabilizza la media aritmetica dei risultati di moltissime ripetizioni dell'esperimento.

Per la somma di due dadi $E(X) = (2 \cdot 1 + 3 \cdot 2 + 4 \cdot 3 + \ldots + 12 \cdot 1) : 36 = 252 : 36 = 7$, come suggerisce la simmetria del grafico intorno a 7. Per le palline rosse dell'esempio svolto $E(X) = 0 \cdot \frac{1}{6} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{3}{10} + 3 \cdot \frac{1}{30} = \frac{6}{5} = 1{,}2$: un valore che X non può assumere. In fisica $E(X)$ ha un'interpretazione immediata: se si pongono sulla retta reale masse p_i nei punti x_i, $E(X)$ è l'ascissa del loro **baricentro**.

In *Matematica 2* il valore atteso è servito a giudicare i giochi d'azzardo. Se G è il guadagno netto di un giocatore (vincita meno posta), il gioco si dice **equo** quando $E(G) = 0$: ripetendolo molte volte, nessuno dei due giocatori è avvantaggiato. Il prezzo equo di una scommessa è il valore atteso della vincita.

ESEMPIO SVOLTO

Il problema. In una fiera di beneficenza (dati inventati) si pagano 2 € per lanciare due dadi: se la somma è almeno 11 si vincono 10 €, se è 7 se ne vincono 4, negli altri casi nulla. Calcola il valore atteso e la deviazione standard della vincita V e il guadagno medio dell'organizzatore per giocata. Qual è il prezzo equo del gioco?

La distribuzione di V. La somma è almeno 11 per 3 coppie su 36 ((5; 6), (6; 5), (6; 6)) ed è 7 per 6 coppie. Quindi $p(V = 10) = \frac{3}{36}$, $p(V = 4) = \frac{6}{36}$, $p(V = 0) = \frac{27}{36}$.

Il valore atteso. $E(V) = 10 \cdot \frac{3}{36} + 4 \cdot \frac{6}{36} + 0 = \frac{54}{36} = 1{,}5$ €. Il guadagno netto del giocatore è $G = V - 2$, con $E(G) = -0{,}5$ €: in media l'organizzatore incassa 50 centesimi a giocata. Il gioco sarebbe equo con un prezzo di 1,50 €.

La deviazione standard (con la formula del teorema che segue). $E(V^2) = 100 \cdot \frac{3}{36} + 16 \cdot \frac{6}{36} = \frac{396}{36} = 11$, quindi $\mathrm{Var}(V) = 11 - 1{,}5^2 = 8{,}75$ e $\sigma(V) = \sqrt{8{,}75} \approx 2{,}96$ €. La vincita si discosta tipicamente di circa 3 € dalla sua media: un gioco molto più «rischioso» di quanto dica il valore atteso.

15.2.2 Varianza e deviazione standard

Due variabili con lo stesso valore atteso possono comportarsi in modo molto diverso: una vincita certa di 1,50 € e la vincita dell'esempio svolto hanno la stessa media, ma la seconda è quasi sempre 0 e qualche volta 10. Per misurare quanto X si disperde intorno a $\mu = E(X)$ si usa, come in statistica, la media dei quadrati degli scarti. Si chiama **varianza** di X il valore atteso di $(X - \mu)^2$:

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

e **deviazione standard** (o **scarto quadratico medio**) la sua radice quadrata, $\sigma(X) = \sqrt{\mathrm{Var}(X)}$. La varianza si indica anche con σ^2. La deviazione standard si misura nella stessa unità di X (euro,

pezzi, chiamate), ed è quindi il numero da confrontare con i valori; la varianza ha proprietà algebriche più semplici. Anche qui le formule sono quelle della statistica descrittiva di *Matematica 3*, con le probabilità al posto delle frequenze relative.

IN SINTESI

Teorema (proprietà del valore atteso e della varianza).

Ipotesi: X variabile aleatoria discreta con valore atteso μ e varianza finita; a, b numeri reali.

Tesi:

1. $\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2$;
2. $\mathrm{E}(aX + b) = a\,\mathrm{E}(X) + b$;
3. $\mathrm{Var}(aX + b) = a^2\,\mathrm{Var}(X)$, quindi $\sigma(aX + b) = |a|\,\sigma(X)$.

Dimostrazione. Sviluppiamo il quadrato nella definizione e spezziamo la somma:

$$\mathrm{Var}(X) = \Sigma\,(x_i^2 - 2\mu\,x_i + \mu^2)\,p_i = \Sigma\,x_i^2\,p_i - 2\mu\,\Sigma\,x_i\,p_i + \mu^2\,\Sigma\,p_i = \mathrm{E}(X^2) - 2\mu\cdot\mu + \mu^2\cdot 1 = \mathrm{E}(X^2) - \mu^2,$$

dove $\mathrm{E}(X^2) = \Sigma\,x_i^2\,p_i$ è il valore atteso della variabile X^2, che assume il valore x_i^2 con probabilità p_i. Per la seconda proprietà, la variabile $aX + b$ assume il valore $a\,x_i + b$ con probabilità p_i, quindi $\mathrm{E}(aX + b) = \Sigma\,(a\,x_i + b)\,p_i = a\,\Sigma\,x_i\,p_i + b\,\Sigma\,p_i = a\mu + b$. Infine, lo scarto di $aX + b$ dalla sua media è $(aX + b) - (a\mu + b) = a(X - \mu)$, e per la proprietà appena dimostrata $\mathrm{Var}(aX + b) = E[a^2(X - \mu)^2] = a^2\,E[(X - \mu)^2] = a^2\,\mathrm{Var}(X)$. c.v.d.

La prima formula è di solito la più comoda per i calcoli. Per le palline rosse: $\mathrm{E}(X^2) = 1\cdot\dfrac{1}{2} + 4\cdot\dfrac{3}{10} + 9\cdot\dfrac{1}{30} = 2$, quindi $\mathrm{Var}(X) = 2 - 1{,}2^2 = 0{,}56$ e $\sigma(X) = \sqrt{0{,}56} \approx 0{,}75$. Le proprietà 2 e 3 dicono come cambiano media e dispersione se cambiamo unità di misura o origine: aggiungere una costante b sposta la distribuzione senza allargarla (la varianza non cambia), moltiplicare per a moltiplica la deviazione standard per $|a|$. Così, se una temperatura T in gradi Celsius ha media 20 e deviazione standard 2, in gradi Fahrenheit $1{,}8T + 32$ ha media 68 e deviazione standard 3,6.

15.2.3 Somme di variabili aleatorie

Spesso una variabile è la somma di altre più semplici: la somma dei due dadi è $X_1 + X_2$, con X_1 e X_2 i punti dei singoli dadi. Due variabili X e Y definite sullo stesso spazio si dicono **indipendenti** se, per ogni valore x di X e ogni valore y di Y, gli eventi $X = x$ e $Y = y$ sono indipendenti: $p(X = x,\, Y = y) = p(X = x)\cdot p(Y = y)$, dove a sinistra c'è la probabilità che i due eventi si verifichino insieme. I punti di due dadi diversi sono indipendenti; il punto di un dado e la somma dei due non lo sono. Enunciamo senza dimostrazione due risultati fondamentali:

- per **ogni** coppia di variabili, $E(X + Y) = E(X) + E(Y)$;
- se X e Y sono **indipendenti**, $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Entrambi si estendono a somme di un numero qualsiasi di variabili (indipendenti a due a due, per la varianza). La differenza fra i due enunciati è essenziale: il valore atteso è sempre additivo, la varianza solo sotto un'ipotesi.

ERRORE FREQUENTE

Il valore atteso non è il valore più probabile. Per le palline rosse $E(X) = 1{,}2$, un valore impossibile; il valore più probabile, la **moda**, è 1. Per la vincita dell'esempio svolto il valore atteso è 1,5 €, mentre il valore più probabile è 0. Il valore atteso è una media su molte ripetizioni, non una previsione del singolo esito.

La varianza di una somma non è sempre la somma delle varianze. Se $Y = X$ (due «copie» dello stesso dado), $X + Y = 2X$ e $\mathrm{Var}(2X) = 4\,\mathrm{Var}(X)$, non $2\,\mathrm{Var}(X)$. L'additività vale per variabili indipendenti. E per la differenza di variabili indipendenti le varianze si **sommano**: $\mathrm{Var}(X - Y) = \mathrm{Var}(X) + (-1)^2\,\mathrm{Var}(Y)$.

15.2.4 La variabile standardizzata

Le proprietà del teorema permettono di riportare ogni variabile a una scala comune. Se X ha media μ e deviazione standard $\sigma > 0$, la variabile

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

si chiama **variabile standardizzata** associata a X. Applicando il teorema con $a = 1/\sigma$ e $b = -\mu/\sigma$: $E(Z) = \frac{\mu}{\sigma} - \frac{\mu}{\sigma} = 0$ e $\mathrm{Var}(Z) = \frac{1}{\sigma^2}\,\mathrm{Var}(X) = 1$. Il valore $z = (x - \mu)/\sigma$ dice di quante deviazioni standard x si discosta dalla media, esattamente come il punteggio standardizzato dei dati in *Matematica 3*. Nell'esempio della fiera, vincere 10 € corrisponde a $z = (10 - 1{,}5) : 2{,}96 \approx 2{,}9$: un risultato lontano quasi tre deviazioni standard dalla media. La standardizzazione sarà lo strumento centrale del prossimo capitolo, per la distribuzione normale.

15.3 Le distribuzioni uniforme e binomiale

15.3.1 La distribuzione uniforme discreta

Una variabile che assume n valori, tutti con probabilità $1/n$, ha **distribuzione uniforme discreta**. È il modello del lancio di un dado equilibrato, dell'estrazione di un numero della tombola, della scelta a caso di uno studente da un elenco. Se i valori sono 1, 2, …, n, usando le somme

$1 + 2 + \dots + n = \dfrac{n(n+1)}{2}$ e $1^2 + 2^2 + \dots + n^2 = \dfrac{n(n+1)(2n+1)}{6}$, dimostrate per induzione in *Matematica 3*, si trova

$$\mathrm{E}(X) = \frac{1}{n} \cdot \frac{n(n+1)}{2} = \frac{n+1}{2}, \quad \mathrm{E}(X^2) = \frac{(n+1)(2n+1)}{6},$$
$$\mathrm{Var}(X) = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{n^2-1}{12}.$$

Per un dado ($n = 6$): $\mathrm{E}(X) = \dfrac{7}{2} = 3{,}5$, $\mathrm{Var}(X) = \dfrac{35}{12} \approx 2{,}92$, $\sigma(X) \approx 1{,}71$. Ora possiamo controllare i risultati sulla somma di due dadi con i teoremi sulle somme: $\mathrm{E}(X_1 + X_2) = 3{,}5 + 3{,}5 = 7$ e, poiché i dadi sono indipendenti, $\mathrm{Var}(X_1 + X_2) = \dfrac{35}{12} + \dfrac{35}{12} = \dfrac{35}{6}$, come si verifica anche direttamente dalla tabella delle 11 probabilità. La deviazione standard della somma è $\sqrt{35/6} \approx 2{,}42$: raddoppiando il numero dei dadi la media raddoppia, ma la deviazione standard cresce solo del fattore $\sqrt{2}$.

15.3.2 La distribuzione binomiale

La variabile più semplice di tutte ha due soli valori: 1 se un evento si verifica (un «successo»), 0 altrimenti. Se il successo ha probabilità p, questa **variabile di Bernoulli** X_i ha

$$\mathrm{E}(X_i) = 1 \cdot p + 0 \cdot (1-p) = p, \quad \mathrm{Var}(X_i) = \mathrm{E}(X_i^2) - p^2 = p - p^2 = p(1-p),$$

perché $X_i^2 = X_i$ (i valori sono 0 e 1). Ripetiamo ora n prove indipendenti, ciascuna con probabilità di successo p, e contiamo i successi: $X = X_1 + X_2 + \dots + X_n$. La formula di Bernoulli di *Matematica 4* ci dà la distribuzione di X, che ora guardiamo come un oggetto unico. Una variabile con questa distribuzione si dice **binomiale** di parametri n e p, e si scrive $X \sim \mathrm{B}(n; p)$.

IN SINTESI

Teorema (distribuzione binomiale).

Ipotesi: X è il numero di successi in n prove indipendenti, ciascuna con probabilità di successo p.

Tesi:

$$p(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n; \quad \mathrm{E}(X) = np, \quad \mathrm{Var}(X) = np(1-p).$$

Dimostrazione (del valore atteso e della varianza; la distribuzione è la formula di Bernoulli). Scriviamo $X = X_1 + \dots + X_n$, dove X_i vale 1 se la prova i-esima è un successo e 0 altrimenti. Per l'additività del valore atteso, $\mathrm{E}(X) = \mathrm{E}(X_1) + \dots + \mathrm{E}(X_n) = p + \dots + p = np$. Poiché le prove sono indipendenti, lo sono anche le variabili X_i, e le varianze si sommano: $\mathrm{Var}(X) = n \cdot p(1-p)$. c.v.d.

Il calcolo diretto, con la somma $\Sigma\, k\, \binom{n}{k} p^k(1-p)^{n-k}$, è possibile ma laborioso: scomporre una variabile complicata in una somma di variabili semplici è un'idea che conviene ricordare. Il risultato $\mathrm{E}(X) = np$, già incontrato in un box di *Matematica 4*, ha un significato intuitivo: lanciando 600 volte un dado ci aspettiamo in media 100 sei. La deviazione standard $\sqrt{np(1-p)}$ quantifica la parola «in media»: per 600 lanci è $\sqrt{600 \cdot \frac{1}{6} \cdot \frac{5}{6}} \approx 9{,}1$, quindi un numero di sei fra 90 e 110 non deve sorprendere.

15.3.3 La forma della distribuzione e la moda

La Figura 15.2 mostra tre distribuzioni binomiali con $n = 10$. Per $p = 0{,}5$ l'istogramma è simmetrico intorno a $np = 5$, perché $\binom{n}{k} = \binom{n}{n-k}$. Per $p = 0{,}2$ le probabilità si concentrano sui valori piccoli e la distribuzione ha una «coda» verso destra; per $p = 0{,}8$ il grafico è il simmetrico del precedente, perché contare i successi con probabilità p equivale a contare gli insuccessi con probabilità $1 - p$: se $X \sim \mathrm{B}(n; p)$, allora $n - X \sim \mathrm{B}(n; 1-p)$.

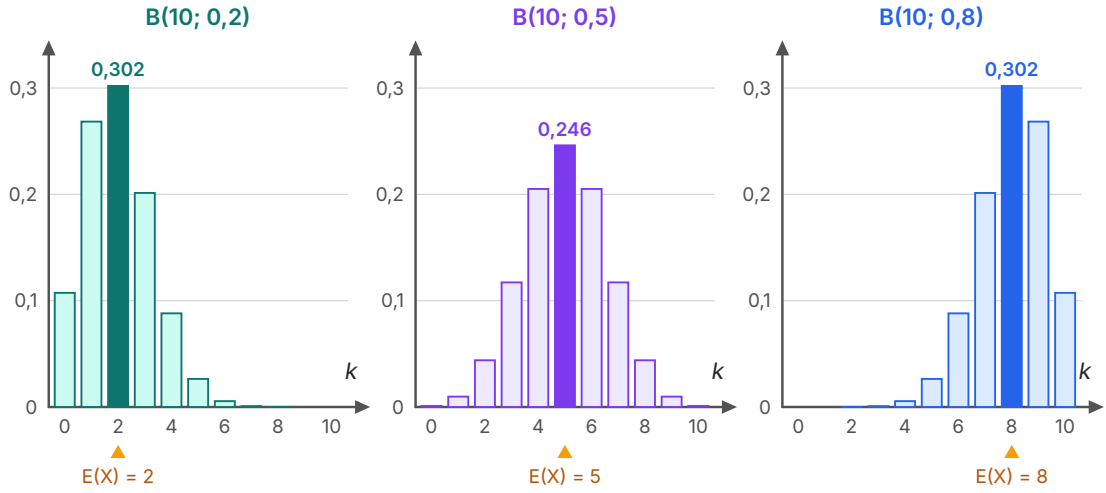


Figura 15.2 Distribuzioni binomiali B(10; p) per $p = 0{,}2$, 0,5 e 0,8 (stessa scala). La barra piena è la moda, con la sua probabilità; il triangolo sotto l'asse indica il valore atteso np. La prima e la terza distribuzione sono l'una la simmetrica dell'altra.

Il valore più probabile, la **moda**, si trova confrontando due probabilità consecutive. Semplificando i coefficienti binomiali si ottiene

$$\frac{p(X = k+1)}{p(X = k)} = \frac{(n-k)\,p}{(k+1)(1-p)},$$

che è maggiore di 1 se e solo se $(n-k)p > (k+1)(1-p)$, cioè $k < (n+1)p - 1$. Le probabilità quindi crescono finché k è minore di $(n+1)p - 1$ e poi decrescono: la moda è l'intero k compreso fra $(n+1)p - 1$ e $(n+1)p$. Se $(n+1)p$ è intero ci sono due mode, $(n+1)p - 1$ e $(n+1)p$, con la stessa probabilità. Nella figura $(n+1)p$ vale 2,2, 5,5 e 8,8, e le mode sono 2, 5 e 8. La moda è sempre uno dei due interi più vicini al valore atteso np, ma non coincide necessariamente con esso.

ESEMPIO SVOLTO

Il problema. Una macchina produce pezzi difettosi con probabilità del 10%, indipendentemente l'uno dall'altro (dati inventati). I pezzi sono confezionati in scatole da 20. Calcola la probabilità che una scatola contenga nessun pezzo difettoso, almeno 2 pezzi difettosi, più di 4 pezzi difettosi; determina il valore atteso, la deviazione standard e la moda del numero di difettosi.

Il modello. Il numero X di difettosi in una scatola è binomiale, $X \sim B(20; 0{,}1)$: 20 prove indipendenti, «successo» = pezzo difettoso, $p = 0{,}1$.

Le probabilità. $p(X = 0) = 0{,}9^{20} \approx 0{,}1216$. Per «almeno 2» conviene l'evento contrario: $p(X = 1) = 20 \cdot 0{,}1 \cdot 0{,}9^{19} \approx 0{,}2702$, quindi $p(X \geq 2) = 1 - 0{,}1216 - 0{,}2702 = 0{,}6082$ (0,6083 con le probabilità non arrotondate). Allo stesso modo $p(X > 4) = 1 - p(X \leq 4) \approx 1 - 0{,}9568 = 0{,}0432$.

I parametri. $E(X) = 20 \cdot 0{,}1 = 2$, $\mathrm{Var}(X) = 20 \cdot 0{,}1 \cdot 0{,}9 = 1{,}8$, $\sigma(X) = \sqrt{1{,}8} \approx 1{,}34$. Poiché $(n+1)p = 2{,}1$, la moda è 2, che qui coincide con il valore atteso.

Il commento. Circa una scatola su otto è perfetta, e poco più di quattro su cento contengono 5 o più pezzi difettosi. Se un cliente trovasse spesso scatole con 5 difettosi, il modello (e forse la macchina) andrebbe rivisto.

APPROFONDIMENTO · LA DISTRIBUZIONE GEOMETRICA

Lanciamo un dado finché esce il sei. Quanti lanci servono? Il numero T di prove necessarie per ottenere il **primo** successo, in prove indipendenti con probabilità di successo p, assume i valori 1, 2, 3, ...: l'evento $T = k$ significa $k - 1$ insuccessi seguiti da un successo, quindi

$$p(T = k) = (1-p)^{k-1}\, p, \quad k = 1, 2, 3, \ldots$$

Si tratta di una **distribuzione geometrica**: le probabilità formano una progressione geometrica di ragione $q = 1 - p$, e la loro somma è $p\,(1 + q + q^2 + \ldots) = p \cdot \dfrac{1}{1-q} = 1$ (serie geometrica con $|q| < 1$). Per il valore atteso c'è un ragionamento elegante: con probabilità p basta un lancio; con probabilità q il primo lancio fallisce e, poiché le prove sono indipendenti, la situazione ricomincia da capo, con un lancio già speso. Quindi $E(T) = p \cdot 1 + q \cdot (1 + E(T))$, da cui $p\,E(T) = 1$ e

$$E(T) = \frac{1}{p}.$$

Per il sei servono in media 6 lanci. Il ragionamento presuppone che $E(T)$ sia finito, cosa che si può dimostrare; il calcolo diretto richiede la somma $\Sigma\, k\, q^{k-1}$, che va oltre il programma.

15.4 La distribuzione di Poisson

15.4.1 Eventi rari: dalla binomiale al limite

Il centralino di un servizio di assistenza riceve in media λ chiamate al minuto. Quante ne arrivano in un minuto preciso? Le chiamate provengono da un gran numero di utenti, ciascuno dei quali ha una probabilità molto piccola di telefonare proprio in quel minuto, indipendentemente dagli altri. Possiamo costruire un modello così: dividiamo il minuto in n intervallini uguali, tanto brevi che in ciascuno arrivi al massimo una chiamata, con probabilità $p = \lambda/n$ (in modo che la media totale sia $np = \lambda$). Il numero di chiamate è allora binomiale, $\mathrm{B}(n;\, \lambda/n)$. Il modello è tanto più realistico quanto più gli intervallini sono brevi: facciamo tendere n a $+\infty$.

IN SINTESI

Teorema (la binomiale tende alla distribuzione di Poisson).

Ipotesi: $\lambda > 0$ fissato, k intero naturale fissato; per $n > \lambda$, $X_n \sim \mathrm{B}(n;\, \lambda/n)$.

Tesi:

$$\lim_{n \to +\infty} p(X_n = k) = \frac{\lambda^k}{k!}\, e^{-\lambda}.$$

Dimostrazione. Per $n \geq k$ scriviamo la probabilità binomiale con $p = \lambda/n$ e separiamo i fattori:

$$p(X_n = k) = \frac{n(n-1)\cdots(n-k+1)}{k!} \cdot \frac{\lambda^k}{n^k} \cdot \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{\lambda^k}{k!} \cdot \left[\frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-k+1}{n}\right] \cdot \left(1 - \frac{\lambda}{n}\right)^{n} \cdot \left(1 - \frac{\lambda}{n}\right)^{-k}.$$

Studiamo il limite di ciascun fattore per $n \to +\infty$. Il primo, $\lambda^k/k!$, non dipende da n. La parentesi quadra è il prodotto di k fattori (un numero fisso) ciascuno dei quali tende a 1, come $(n-2)/n = 1 - 2/n$; per il teorema sul limite di un prodotto tende a 1. Il terzo fattore è il limite notevole $\lim (1 + a/n)^n = e^a$, con $a = -\lambda$: tende a $e^{-\lambda}$. L'ultimo è una potenza con esponente fisso $-k$ di una base che tende a 1, e tende a 1. Il prodotto dei limiti è $\lambda^k\, e^{-\lambda}/k!$. c.v.d.

15.4.2 La distribuzione e i suoi parametri

Una variabile aleatoria X che assume i valori 0, 1, 2, … (tutti i numeri naturali) con probabilità

$$p(X = k) = \frac{\lambda^k}{k!}\, e^{-\lambda}, \quad k = 0, 1, 2, \ldots, \quad \lambda > 0$$

si dice **variabile di Poisson** di parametro λ. A differenza della binomiale, i valori possibili non hanno un massimo. Che le probabilità abbiano somma 1 dipende da un risultato di analisi che accettiamo senza dimostrazione: per ogni reale λ la somma $1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \dots$ (limite delle somme parziali) vale e^{λ}. Allora $\Sigma\, p(X=k) = e^{-\lambda} \cdot e^{\lambda} = 1$.

Lo stesso risultato dà il valore atteso. Il termine con $k = 0$ è nullo, e per $k \geq 1$ si semplifica $k/k! = 1/(k-1)!$:

$$\mathrm{E}(X) = \Sigma_{k\geq 1}\, k\, \frac{\lambda^k}{k!}\, e^{-\lambda} = \lambda\, e^{-\lambda}\, \Sigma_{k\geq 1}\, \frac{\lambda^{k-1}}{(k-1)!} = \lambda\, e^{-\lambda} \cdot e^{\lambda} = \lambda.$$

Anche la varianza vale λ. Lo suggerisce il passaggio al limite: la varianza di $\mathrm{B}(n;\, \lambda/n)$ è $n \cdot \frac{\lambda}{n}\,(1 - \frac{\lambda}{n}) = \lambda\,(1 - \frac{\lambda}{n})$, che tende a λ; il calcolo esatto, con lo stesso metodo usato per $\mathrm{E}(X)$, è proposto negli esercizi. In sintesi:

$$\mathrm{E}(X) = \mathrm{Var}(X) = \lambda, \quad \sigma(X) = \sqrt{\lambda}.$$

La Figura 15.3 mostra la distribuzione per tre valori di λ. Al crescere di λ la distribuzione si sposta verso destra, si allarga (la deviazione standard è $\sqrt{\lambda}$) e diventa più simmetrica. Confrontando $p(X=k)$ e $p(X=k-1)$, il loro rapporto vale λ/k: le probabilità crescono finché $k < \lambda$ e poi decrescono, quindi la moda è la parte intera di λ; quando λ è intero ci sono due mode, $\lambda - 1$ e λ, come si vede nei tre grafici.

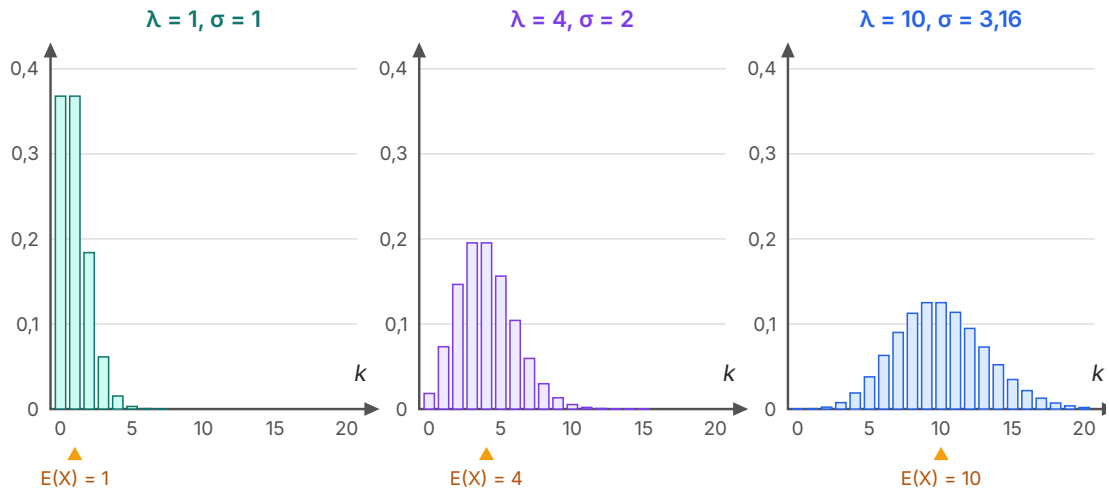


Figura 15.3 Distribuzioni di Poisson con $\lambda = 1$, 4 e 10, per k da 0 a 20 (stessa scala). Il triangolo sotto l'asse indica il valore atteso λ. Con λ intero le due barre in $\lambda - 1$ e λ hanno la stessa altezza; al crescere di λ la distribuzione si allarga e diventa più simmetrica.

15.4.3 Quanto è buona l'approssimazione

Il teorema giustifica l'uso della distribuzione di Poisson al posto della binomiale quando n è grande e p è piccolo, con $\lambda = np$. La Tabella 15.2 e la Figura 15.4 confrontano B(100; 0,02) con la distribuzione di Poisson di parametro $\lambda = 100 \cdot 0{,}02 = 2$: le differenze non superano tre millesimi. I calcoli con la Poisson sono molto più semplici: non servono coefficienti binomiali con $n = 100$, e un solo parametro, λ, descrive la situazione.

Tabella 15.2 Confronto fra B(100; 0,02) e la distribuzione di Poisson con $\lambda = 2$ (valori arrotondati alla quarta cifra decimale; le differenze sono calcolate sui valori non arrotondati).

k	B(100; 0,02)	Poisson, $\lambda = 2$	Differenza
0	0,1326	0,1353	−0,0027
1	0,2707	0,2707	0,0000
2	0,2734	0,2707	+0,0027
3	0,1823	0,1804	+0,0018
4	0,0902	0,0902	0,0000
5	0,0353	0,0361	−0,0007
6	0,0114	0,0120	−0,0006
7	0,0031	0,0034	−0,0003
8	0,0007	0,0009	−0,0001

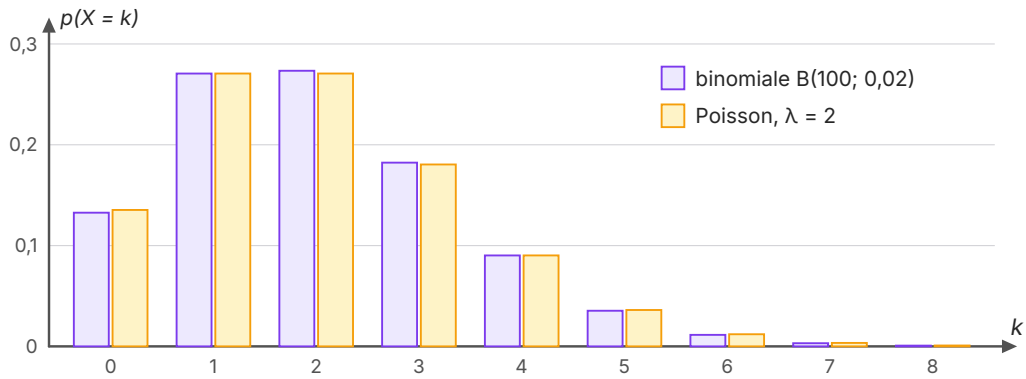


Figura 15.4 Le probabilità $p(X = k)$ per k da 0 a 8 secondo la binomiale B(100; 0,02) (barre viola) e secondo la distribuzione di Poisson con $\lambda = 2$ (barre ambra): le coppie di barre sono quasi uguali. Le probabilità per $k > 8$ sono minori di un millesimo in entrambe.

In pratica la si usa quando n è grande (qualche decina o più) e p è piccolo (qualche centesimo), così che np resti moderato. Ma il suo interesse va oltre l'approssimazione: la distribuzione di Poisson è il modello naturale per il **numero di eventi rari** che si verificano in un intervallo di tempo (o di spazio) quando gli eventi sono indipendenti fra loro, avvengono uno alla volta e con un ritmo medio costante. Chiamate a un centralino, clienti che arrivano a uno sportello in un quarto d'ora, refusi in una pagina, guasti di un impianto in un anno: in tutti questi casi non conosciamo né n né p, ma solo il numero medio λ, e basta quello. Se il ritmo medio è costante, in un intervallo lungo t volte il numero medio è λt.

Torniamo al centralino, supponendo (dati inventati) che riceva in media 2,4 chiamate al minuto. La probabilità che in un dato minuto non arrivi alcuna chiamata è $e^{-2,4} \approx 0,0907$; la probabilità che ne arrivino almeno 3 è

$$p(X \geq 3) = 1 - e^{-2,4}\left(1 + 2,4 + \frac{2,4^2}{2}\right) = 1 - 6,28\, e^{-2,4} \approx 1 - 0,5697 = 0,4303.$$

In due minuti il numero medio diventa $\lambda = 4,8$ e la probabilità di nessuna chiamata scende a $e^{-4,8} \approx 0,0082$: meno di un caso su cento.

ERRORE FREQUENTE

Usare la distribuzione di Poisson con *p* grande. L'approssimazione richiede che p sia piccolo, non solo che n sia grande. Il numero di teste in 10 lanci di una moneta è B(10; 0,5), e la probabilità di 5 teste è ≈ 0,2461; la Poisson con $\lambda = 5$ darebbe ≈ 0,1755, un errore di oltre il 28%. Inoltre la Poisson assegna probabilità positiva (circa 0,014) a più di 10 teste, che sono impossibili. Un segnale d'allarme: per la binomiale $\mathrm{Var}(X) = np(1-p)$ è minore di $\mathrm{E}(X)$, per la Poisson le due sono uguali; se $1 - p$ è lontano da 1, la Poisson sovrastima la dispersione.

COLLEGAMENTI · FISICA E BIOLOGIA

Un campione di una sostanza radioattiva contiene un numero enorme di nuclei, ciascuno dei quali ha una probabilità piccolissima di decadere in un dato secondo, indipendentemente dagli altri. È la situazione del teorema: il numero di decadimenti registrati da un contatore in un secondo segue con ottima approssimazione una distribuzione di Poisson, purché la sostanza decada lentamente rispetto alla durata della misura (altrimenti il ritmo medio non è costante). Una conseguenza pratica: se in media si contano λ impulsi, la fluttuazione tipica è $\sqrt{\lambda}$, quindi l'errore relativo $\sqrt{\lambda}/\lambda = 1/\sqrt{\lambda}$ diminuisce contando più a lungo; per dimezzarlo occorre contare per un tempo quattro volte maggiore.

In biologia, lo stesso modello descrive il numero di mutazioni in un tratto di DNA o il numero di cellule in un quadratino di un vetrino contacellule, quando le cellule sono ben disperse nella sospensione.

APPROFONDIMENTO · POISSON E I CALCI DI CAVALLO

Siméon-Denis Poisson (1781–1840) ricavò la distribuzione che porta il suo nome come limite della binomiale nelle *Recherches sur la probabilité des jugements en matière criminelle et en matière civile* (1837), un'opera dedicata alla probabilità che una giuria arrivi a un verdetto corretto. Per decenni il risultato rimase una curiosità; lo rese celebre lo statistico Ladislaus von Bortkiewicz, che nel 1898 mostrò che il numero annuo di soldati uccisi da un calcio di cavallo nei corpi d'armata dell'esercito prussiano seguiva molto bene una distribuzione di Poisson: il modello degli eventi rari e indipendenti funzionava anche per un fenomeno che nessuno avrebbe definito «matematico».

15.5 La legge dei grandi numeri

15.5.1 La disuguaglianza di Čebyšev

La deviazione standard misura la dispersione «tipica», ma quanto è probabile che X si allontani dalla media più di un certo numero di deviazioni standard? Sorprendentemente, una risposta vale per **ogni** distribuzione, purché la varianza sia finita. La dimostrò il matematico russo Pafnutij L'vovič Čebyšev (1821–1894).

IN SINTESI

Teorema (disuguaglianza di Čebyšev).

Ipotesi: X variabile aleatoria con valore atteso μ e varianza σ^2 finita; $\varepsilon > 0$.

Tesi:

$$p(|X-\mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2} \quad \text{e, posto } \varepsilon = k\sigma \text{ con } k>0, \quad p(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

Equivalentemente, passando all'evento contrario, $p(|X-\mu| < k\sigma) \geq 1 - 1/k^2$: qualunque sia la distribuzione, almeno il 75% della probabilità sta entro 2 deviazioni standard dalla media, almeno l'89% ($1 - 1/9$) entro 3. Per $k \leq 1$ la disuguaglianza non dice nulla, perché ogni probabilità è minore o uguale a 1.

Controlliamola sulla somma di due dadi, con $\mu = 7$ e $\sigma = \sqrt{35/6} \approx 2{,}415$ (Figura 15.5). Con $k = 2$ la fascia $\mu \pm 2\sigma$ va da circa 2,17 a circa 11,83: ne restano fuori soltanto i valori 2 e 12, con probabilità complessiva $\frac{2}{36} \approx 0{,}056$, molto meno del limite $\frac{1}{4}$ garantito da Čebyšev. La disuguaglianza è grossolana proprio perché è universale: deve valere anche per le distribuzioni più

sfavorevoli. Il suo valore è teorico, e il risultato più importante che se ne ricava è la legge dei grandi numeri.

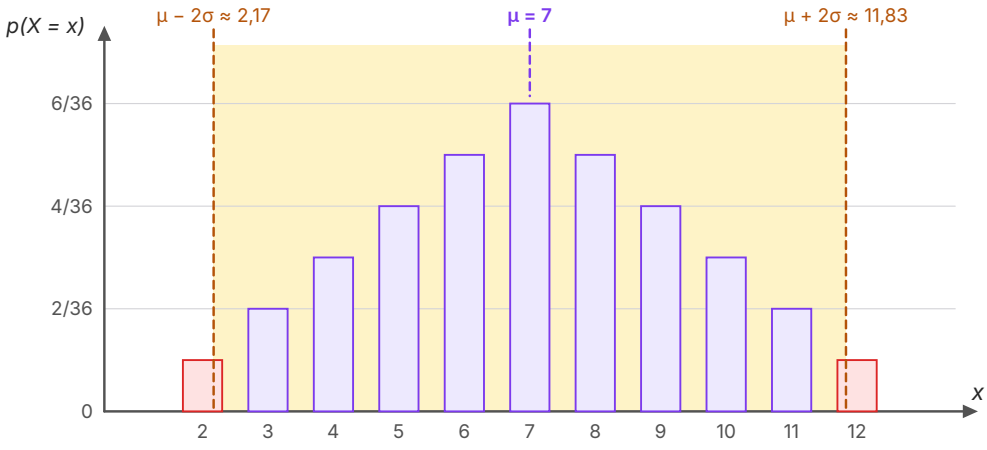


Figura 15.5 La disuguaglianza di Čebyšev per la somma di due dadi. La fascia colorata è l'intervallo $\mu \pm 2\sigma$, con $\mu = 7$ e $\sigma = \sqrt{35/6}$. La probabilità dei valori fuori dalla fascia (barre rosse, 2 e 12) è $2/36 \approx 0{,}056$, ben al di sotto del limite $1/k^2 = 1/4$.

APPROFONDIMENTO · PERCHÉ VALE LA DISUGUAGLIANZA DI ČEBYŠEV

La dimostrazione, per una variabile discreta, richiede due righe. Nella somma che definisce la varianza tutti i termini sono non negativi; se ne trascuriamo alcuni, la somma non aumenta. Teniamo solo i valori x_i lontani dalla media almeno ε, per i quali $(x_i - \mu)^2 \geq \varepsilon^2$:

$$\sigma^2 = \sum_{\text{tutti}} (x_i - \mu)^2 p_i \geq \sum_{|x_i - \mu| \geq \varepsilon} (x_i - \mu)^2 p_i \geq \varepsilon^2 \sum_{|x_i - \mu| \geq \varepsilon} p_i = \varepsilon^2 \cdot p(|X - \mu| \geq \varepsilon).$$

Dividendo per $\varepsilon^2 > 0$ si ottiene la tesi. c.v.d. L'idea, «una varianza piccola non lascia molta probabilità lontano dalla media», è tutta qui.

15.5.2 La legge debole dei grandi numeri

In *Matematica 4* abbiamo visto, con le simulazioni, che la frequenza relativa di un evento si stabilizza intorno alla sua probabilità quando il numero delle prove cresce. Ora possiamo dimostrarlo. Eseguiamo n prove indipendenti di un evento di probabilità p; il numero X dei successi è $B(n; p)$ e la **frequenza relativa** dei successi è $F_n = X/n$. Per le proprietà del valore atteso e della varianza

$$\mathrm{E}(F_n) = \frac{np}{n} = p, \quad \mathrm{Var}(F_n) = \frac{1}{n^2} \cdot np(1-p) = \frac{p(1-p)}{n}.$$

La frequenza relativa ha dunque valore atteso p e una varianza che tende a 0 come $1/n$. Applicando la disuguaglianza di Čebyšev a F_n, per ogni $\varepsilon > 0$:

$$p(|F_n - p| \geq \varepsilon) \leq \frac{p(1-p)}{n\,\varepsilon^2} \leq \frac{1}{4n\,\varepsilon^2},$$

dove l'ultimo passaggio usa $p(1-p) \leq 1/4$ (la parabola $p - p^2$ ha il massimo in $p = 1/2$). Il secondo membro tende a 0 per $n \to +\infty$. Abbiamo così dimostrato la **legge debole dei grandi numeri** (o teorema di Bernoulli): **per ogni $\varepsilon > 0$, la probabilità che la frequenza relativa dei successi differisca dalla probabilità p di almeno ε tende a 0 al crescere del numero delle prove**:

$$\lim_{n \to +\infty} p(|F_n - p| \geq \varepsilon) = 0.$$

La stima di Čebyšev dà anche un'indicazione quantitativa, pur prudente. Quanti lanci di una moneta equilibrata servono per essere sicuri almeno al 90% che la frequenza delle teste sia compresa fra 0,45 e 0,55? Basta che $1/(4n \cdot 0{,}05^2) \leq 0{,}1$, cioè $n \geq 1000$. (Il numero davvero necessario è molto minore, poco meno di 300: la stima di Čebyšev vale per ogni distribuzione e per questo è prudente. Con la distribuzione normale del prossimo capitolo si ottengono stime molto più precise.)

La legge dei grandi numeri riguarda le **frequenze relative**, non i numeri assoluti, e non dice nulla sulla singola prova. Dopo dieci teste consecutive la probabilità di testa al lancio successivo è ancora 1/2: la moneta non «ricorda» e non «compensa». Lo scarto fra numero di teste e $n/2$ anzi tende a crescere, come $\sqrt{n}$; è il suo rapporto con n a tendere a 0. Credere il contrario è la cosiddetta «fallacia del giocatore», che alimenta le scommesse sui numeri «ritardatari» del lotto.

15.6 Scegliere il modello

15.6.1 Tre situazioni, tre distribuzioni

Le formule di questo capitolo servono solo se il modello è quello giusto. Le distribuzioni discrete che abbiamo incontrato nascono da situazioni diverse, ciascuna con le sue ipotesi:

- la **binomiale** $\mathrm{B}(n; p)$ conta i successi in un numero **fissato** n di prove **indipendenti**, con probabilità di successo **costante** p: lanci di dadi o monete, estrazioni **con** reimmissione, pezzi prodotti da una macchina stabile;
- l'**ipergeometrica** conta i successi in n estrazioni **senza** reimmissione da una popolazione finita di N elementi, K dei quali «favorevoli»: $p(X = k) = \binom{K}{k}\binom{N-K}{n-k} : \binom{N}{n}$. Le estrazioni non sono

indipendenti, perché ogni estrazione cambia la composizione dell'urna; il valore atteso è comunque $n\,K/N$ (per l'urna dell'esempio svolto, $3 \cdot 4/10 = 1{,}2$);

- la distribuzione di **Poisson** conta eventi rari e indipendenti in un intervallo di tempo o di spazio, con un ritmo medio costante, quando si conosce solo il numero medio λ; approssima $B(n; p)$ con n grande e p piccolo.

La Tabella 15.3 riassume le distribuzioni del capitolo.

Tabella 15.3 Le distribuzioni discrete del capitolo. Le varianze dell'ipergeometrica e della geometrica sono riportate senza dimostrazione.

Distribuzione	Valori	$p(X = k)$	$E(X)$	$\mathrm{Var}(X)$
Uniforme su 1, ..., n	1, 2, ..., n	$1/n$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$
Di Bernoulli	0, 1	$p(X = 1) = p$	p	$p(1-p)$
Binomiale $B(n; p)$	0, 1, ..., n	$\binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
Ipergeometrica	0, 1, ..., n (al più)	$\binom{K}{k}\binom{N-K}{n-k} : \binom{N}{n}$	$n\,K/N$	$n\,\frac{K}{N}\left(1-\frac{K}{N}\right)\frac{N-n}{N-1}$
Geometrica (approfondimento)	1, 2, 3, ...	$(1-p)^{k-1} p$	$1/p$	$(1-p)/p^2$
Di Poisson	0, 1, 2, ...	$\lambda^k e^{-\lambda}/k!$	λ	λ

Quando la popolazione è molto grande rispetto al numero di estrazioni, togliere un elemento non cambia quasi la composizione, e l'ipergeometrica si può approssimare con la binomiale. Un controllo su 5 pezzi presi senza reimmissione da un lotto di 1000, di cui 50 difettosi, trova nessun difettoso con probabilità $\binom{950}{5} : \binom{1000}{5} \approx 0{,}7734$; la binomiale $B(5; 0{,}05)$ dà $0{,}95^5 \approx 0{,}7738$. Con un lotto di 50 pezzi, invece, la differenza sarebbe sensibile (un esercizio lo mostra).

15.6.2 Quando le ipotesi cadono

Ogni modello è una scommessa sulle ipotesi, e conviene chiedersi sempre quale ipotesi potrebbe cadere. La binomiale richiede **indipendenza**: se i pezzi difettosi nascono da un guasto che dura qualche minuto, arrivano a gruppi, e il numero di scatole con molti difettosi è più alto di quanto prevede $B(n; p)$. Richiede poi una probabilità **costante**: la probabilità che un elettore voti un certo partito non è la stessa in tutte le regioni, e un campione preso in una sola città non è un insieme di prove con lo stesso p. La Poisson richiede un **ritmo costante**: le chiamate a un centralino non arrivano con lo stesso ritmo alle 3 di notte e alle 10 di mattina, quindi il modello va applicato a intervalli in cui il ritmo è circa costante.

Un controllo semplice viene dai parametri. Se un conteggio osservato molte volte ha una varianza molto maggiore della media, la Poisson non è adatta (per la Poisson le due sono uguali);

se si osserva una frequenza di successo che cambia nel tempo, p non è costante. È questo il senso del **modello matematico** di cui parlano le Indicazioni: una descrizione semplificata che si sceglie, si usa e si mette alla prova con i dati, pronti a cambiarla.

ESEMPIO SVOLTO · VERSO L'ESAME

Il quesito. Una variabile aleatoria X assume i valori 1, 2, 3, 4 con probabilità $p(X = k) = ak$. Determina a, rappresenta la funzione di ripartizione e calcola $p(X > 2)$, il valore atteso e la deviazione standard di X. Verifica infine che $Y = 2X - 6$ ha valore atteso 0 e calcolane la deviazione standard.

Il parametro. La somma delle probabilità è 1: $a(1 + 2 + 3 + 4) = 10a = 1$, quindi $a = \frac{1}{10}$. Le probabilità, 0,1, 0,2, 0,3 e 0,4, sono positive.

La funzione di ripartizione. $F(x) = 0$ per $x < 1$; 0,1 per $1 \le x < 2$; 0,3 per $2 \le x < 3$; 0,6 per $3 \le x < 4$; 1 per $x \ge 4$: una scala di quattro gradini, con il valore superiore in ogni punto di salto. Allora $p(X > 2) = 1 - F(2) = 0{,}7$.

Valore atteso e deviazione standard. $\mathrm{E}(X) = \frac{1 + 4 + 9 + 16}{10} = 3$; $\mathrm{E}(X^2) = \frac{1 + 8 + 27 + 64}{10} = 10$; $\mathrm{Var}(X) = 10 - 9 = 1$, $\sigma(X) = 1$.

La variabile Y. $\mathrm{E}(Y) = 2\,\mathrm{E}(X) - 6 = 0$ e $\sigma(Y) = |2| \cdot \sigma(X) = 2$, per le proprietà dimostrate nel secondo paragrafo; non serve ricalcolare la distribuzione di Y.

Il commento. Nella seconda prova un quesito di questo tipo vale quanto gli altri: va scritto ogni passaggio, citando le proprietà usate (somma delle probabilità uguale a 1, linearità del valore atteso, $\mathrm{Var}(aX + b) = a^2\,\mathrm{Var}(X)$).

15.6.3 Con il computer

I fogli di calcolo hanno funzioni che restituiscono le probabilità della distribuzione binomiale e di quella di Poisson, sia nella forma $p(X = k)$ sia in quella cumulativa $F(k)$: basta indicare k, i parametri e un valore logico che sceglie fra le due forme (i nomi esatti cambiano da un programma all'altro e si trovano nella guida). Con una colonna di valori di k si costruiscono in pochi secondi tabelle come la Tabella 15.2 e i relativi istogrammi. Il generatore di numeri casuali permette invece di **simulare** un esperimento: un numero casuale fra 0 e 1 minore di p rappresenta un successo, e ripetendo la simulazione migliaia di volte si confronta la media dei risultati con il valore atteso teorico. È un modo concreto di vedere al lavoro la legge dei grandi numeri.

SINTESI

- Una **variabile aleatoria** è una funzione $X: \Omega \to \mathbb{R}$; è **discreta** se i suoi valori sono finiti o si possono elencare in una successione. La sua **distribuzione di probabilità** associa 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 definita su tutto $\mathbb{R}$, non decrescente, a gradini per le variabili discrete, da 0 a 1; $p(a < X \leq b) = F(b) - F(a)$.
- **Valore atteso** $\mathrm{E}(X) = \Sigma\, x_i\, p_i$; **varianza** $\mathrm{Var}(X) = \Sigma\, (x_i - \mu)^2\, p_i = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2$; **deviazione standard** $\sigma = \sqrt{\mathrm{Var}(X)}$. Il valore atteso non è il valore più probabile (la moda).
- $\mathrm{E}(aX + b) = a\mathrm{E}(X) + b$, $\mathrm{Var}(aX + b) = a^2\,\mathrm{Var}(X)$; $\mathrm{E}(X + Y) = \mathrm{E}(X) + \mathrm{E}(Y)$ sempre, $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$ se X e Y sono indipendenti. La **variabile standardizzata** $Z = (X - \mu)/\sigma$ ha media 0 e varianza 1.
- **Uniforme** su $1, \ldots, n$: $E = (n + 1)/2$, $\mathrm{Var} = (n^2 - 1)/12$; per un dado 3,5 e $\frac{35}{12}$.
- **Binomiale** $\mathrm{B}(n; p)$, numero di successi in n prove indipendenti: $p(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$, $E = np$, $\mathrm{Var} = np(1 - p)$, dimostrati scrivendo X come somma di n variabili di Bernoulli. La moda è l'intero fra $(n + 1)p - 1$ e $(n + 1)p$.
- **Poisson** di parametro λ: $p(X = k) = \lambda^k e^{-\lambda}/k!$, $k = 0, 1, 2, \ldots$; $E = \mathrm{Var} = \lambda$. È il limite di $\mathrm{B}(n; \lambda/n)$ per $n \to +\infty$ e il modello degli eventi rari, indipendenti e con ritmo medio costante.
- La Poisson approssima la binomiale per n grande e p piccolo, con $\lambda = np$ (per B(100; 0,02) le differenze sono di millesimi); non va usata con p grande.
- **Disuguaglianza di Čebyšev**: $p(|X - \mu| \geq k\sigma) \leq 1/k^2$ per ogni distribuzione con varianza finita; è universale ma grossolana.
- **Legge debole dei grandi numeri**: la frequenza relativa F_n dei successi ha media p e varianza $p(1 - p)/n$, e $p(|F_n - p| \geq \varepsilon) \to 0$. Non dice nulla sulla singola prova.
- Scegliere il modello: binomiale (prove indipendenti, p costante, n fissato), ipergeometrica (estrazioni senza reimmissione), Poisson (eventi rari in un intervallo). Le ipotesi vanno sempre controllate.

ESERCIZI

Comprendere

1. ● Una variabile aleatoria X assume i valori 1, 2, 3. Quale terna di probabilità $p(X = 1)$, $p(X = 2)$, $p(X = 3)$ definisce una distribuzione di probabilità?

   A. 0,2; 0,5; 0,4

   B. 0,3; 0,3; 0,4

   C. 0,5; −0,1; 0,6

   D. $\frac{1}{2}; \frac{1}{3}; \frac{1}{4}$

2. ● Se $\mathrm{E}(X) = 3$ e $\mathrm{Var}(X) = 4$, la varianza di $1 - 2X$ è

   A. 16

   B. −8

   C. 9

   D. 17

3. ● Per una variabile $X \sim \mathrm{B}(n; p)$ con n fissato, la varianza è massima quando

   A. $p = 1$

   B. $p = 1/n$

   C. $p = 0$

   D. $p = 1/2$

4. ● Una variabile di Poisson ha valore atteso 9. La sua deviazione standard è

   A. 9

   B. 81

   C. 3

   D. 4,5

5. ● La variabile X assume i valori 0, 1, 2 con probabilità 0,2, 0,5 e 0,3. La sua funzione di ripartizione in $x = 1{,}5$ vale

   A. 0,7

   B. 0,5

   C. 0,3

   D. 0,8

6. ● Vero o falso? Correggi le affermazioni false. **a)** Il valore atteso di una variabile aleatoria è sempre uno dei valori che essa può assumere. **b)** Per due variabili qualsiasi $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$. **c)** Per una variabile di Poisson il valore atteso e la varianza coincidono. **d)** In una distribuzione binomiale il valore più probabile è sempre np.

7. ● Completa. **a)** Var(X) = E(_____) − _____. **b)** Se $X \sim B(n; p)$, E(X) = _____ e Var(X) = _____. **c)** La distribuzione di Poisson è il limite di B(n; p) per $n \to$ _____ con il prodotto _____ costante e uguale a λ. **d)** Per la disuguaglianza di Čebyšev, $p(|X - \mu| \geq 3\sigma) \leq$ _____.

8. ● Vero o falso? Correggi le affermazioni false. **a)** La funzione di ripartizione può assumere valori maggiori di 1. **b)** Per una variabile discreta, $p(a < X \leq b) = F(b) - F(a)$. **c)** Una variabile standardizzata ha deviazione standard 1. **d)** La moda di una variabile di Poisson con $\lambda = 2{,}5$ è 2.

Calcolare

9. ● Si lancia tre volte una moneta equilibrata e si indica con X il numero di teste. Scrivi la distribuzione di probabilità e la funzione di ripartizione di X e calcola E(X) e Var(X).

10. ●● Si lanciano due dadi e si indica con M il punteggio più alto (se i punteggi sono uguali, M è il loro valore comune). **a)** Dimostra che $p(M = k) = (2k - 1)/36$ per $k = 1, \ldots, 6$. **b)** Calcola E(M). **c)** Calcola $p(M \leq 3)$.

11. ● La variabile X assume i valori −1, 0, 2, 5 con probabilità 0,2, 0,3, 0,4, 0,1. Calcola E(X), Var(X), $\sigma(X)$, $F(1)$ e $p(X > 0)$.

12. ●● La funzione di ripartizione di una variabile discreta X è $F(x) = 0$ per $x < 1$, $F(x) = 0{,}25$ per $1 \leq x < 2$, $F(x) = 0{,}6$ per $2 \leq x < 4$, $F(x) = 1$ per $x \geq 4$. Ricava la distribuzione di probabilità di X, E(X) e $p(1 < X \leq 4)$.

13. ●● Determina k in modo che $p(X = x) = kx^2$, per $x = 1, 2, 3$, sia una distribuzione di probabilità; calcola poi E(X), Var(X) e $p(X \geq 2)$.

14. ● Una variabile X ha E(X) = 4 e Var(X) = 2,5. Calcola E($3X - 2$), Var($3X - 2$), $\sigma(7 - X)$ ed E(X^2).

15. ●● Un test ha 8 domande, ciascuna con 4 risposte di cui una sola esatta; uno studente risponde a caso a tutte. Detto X il numero di risposte esatte, calcola $p(X = 2)$, $p(X \geq 1)$, E(X), $\sigma(X)$ e la moda di X.

16. ●● Una variabile binomiale ha valore atteso 6 e varianza 4,2. Determina n e p, e calcola $p(X = 6)$.

17. ● X è una variabile di Poisson con $\lambda = 3$. Calcola $p(X = 0)$, $p(X = 2)$, $p(X \leq 2)$ e $p(X \geq 2)$.

18. ●● Calcola $p(X = k)$ per $k = 0, 1, 2$ con $X \sim B(200; 0{,}01)$ e con la distribuzione di Poisson che la approssima, e confronta i risultati.

19. ●● Di una variabile X si sa soltanto che $\mu = 50$ e $\sigma = 4$. **a)** Con la disuguaglianza di Čebyšev, trova un limite inferiore per $p(40 < X < 60)$. **b)** Determina il più piccolo ε per il quale la disuguaglianza garantisce $p(|X - 50| < \varepsilon) \geq 0{,}96$.

Dimostrare e argomentare

20. ●● Dimostra che Var(X) ≥ 0 per ogni variabile aleatoria e deducine che E(X^2) ≥ [E(X)]^2. In quale caso vale l'uguaglianza?

21. ●●● Sia X una variabile di Poisson di parametro λ. Usando $\Sigma_{j\ge 0}\, \lambda^j/j! = e^{\lambda}$, dimostra che $E[X(X-1)] = \lambda^2$ e deducine che $\mathrm{Var}(X) = \lambda$.

22. ●●● Per $X \sim \mathrm{B}(n; p)$ con $0 < p < 1$, verifica che $p(X = k+1)/p(X = k) = (n-k)p/((k+1)(1-p))$ e dimostra che le probabilità crescono finché $k < (n+1)p - 1$. Trova la moda di B(15; 0,3).

23. ● Trova l'errore. «X e Y sono indipendenti, con $\mathrm{Var}(X) = 3$ e $\mathrm{Var}(Y) = 2$; quindi $\mathrm{Var}(X - Y) = 3 - 2 = 1$.»

24. ●● Un controesempio. Sia X il punteggio di un dado e $Y = X$. Calcola $\mathrm{E}(XY)$ ed $\mathrm{E}(X) \cdot \mathrm{E}(Y)$, $\mathrm{Var}(X + Y)$ e $\mathrm{Var}(X) + \mathrm{Var}(Y)$, e spiega perché i risultati non contraddicono i teoremi del capitolo.

Problemi

25. ●● In una lotteria di beneficenza (dati inventati) si vendono 1000 biglietti a 5 € l'uno. I premi sono: uno da 1000 €, cinque da 200 €, venti da 50 €. Detta V la vincita di chi compra un biglietto, calcola $\mathrm{E}(V)$ e $\sigma(V)$, il guadagno medio di chi compra un biglietto e il prezzo che renderebbe equa la lotteria.

26. ●● In un libro (dati inventati) i refusi sono in media 0,5 per pagina, distribuiti secondo il modello di Poisson. **a)** Qual è la probabilità che una pagina non contenga refusi? **b)** Qual è la probabilità che un capitolo di 4 pagine ne contenga al massimo uno? **c)** Quante pagine senza refusi ci si aspetta in un libro di 200 pagine?

27. ●●● Una compagnia aerea vende 52 biglietti per un volo con 50 posti, sapendo dall'esperienza (dati inventati) che ogni passeggero si presenta all'imbarco con probabilità 0,9, indipendentemente dagli altri. Calcola il numero atteso di passeggeri che si presentano, la deviazione standard e la probabilità che qualcuno resti a terra.

28. ●●● Un lotto di 50 pezzi contiene 4 pezzi difettosi. Se ne controllano 5, presi senza reimmissione. Calcola la probabilità che fra i pezzi controllati non ci siano difettosi, prima con la distribuzione ipergeometrica, poi approssimandola con la binomiale B(5; 0,08). L'approssimazione è accettabile? Che cosa cambierebbe con un lotto di 5000 pezzi, 400 dei quali difettosi?

Con il computer

29. ●● Con un foglio di calcolo costruisci, per k da 0 a 20, le probabilità della binomiale B(20; 0,15) e della distribuzione di Poisson con lo stesso valore atteso, usando le funzioni del programma per le due distribuzioni. Rappresentale con un istogramma doppio e individua il valore di k in cui la differenza è massima. L'approssimazione è buona come quella della Tabella 15.2? Perché?

30. ●● Simula 1000 volte l'esperimento «10 prove con probabilità di successo 0,3»: in ogni riga genera 10 numeri casuali fra 0 e 1, conta quanti sono minori di 0,3 e scrivi il conteggio. Calcola la media e la varianza dei 1000 conteggi e confrontale con np e $np(1-p)$. Ripeti la simulazione qualche volta: di quanto varia la media da una simulazione all'altra?

Verso l'esame

31. ●● Una variabile aleatoria X assume i valori 1, 2, 3 con probabilità a, b, $\frac{1}{4}$. Sapendo che $\mathrm{E}(X) = 2$, determina a e b e calcola la varianza di X.

32. ●● Il numero di clienti che entrano in un negozio in 5 minuti è una variabile di Poisson di parametro λ. Determina λ sapendo che la probabilità che entri esattamente un cliente è uguale alla probabilità che ne entrino esattamente due, e calcola la probabilità che ne entrino almeno tre.

33. ●● Quante volte occorre lanciare un dado equilibrato perché la probabilità di ottenere almeno un sei superi 0,99? Con quel numero di lanci, qual è il valore atteso del numero di sei?

34. ●●● Sia $X \sim \mathrm{B}(n; p)$ con $n \geq 2$ fissato. **a)** Determina il valore di $p \in [0; 1]$ che rende massima la probabilità $g(p) = p(X = 1)$. **b)** Calcola il valore massimo e il suo limite per $n \to +\infty$.

35. ●●● **Problema.** Un'azienda (dati inventati) produce microchip; ogni chip è difettoso con probabilità p, indipendentemente dagli altri, e i chip sono venduti in confezioni da 50. Sia X il numero di chip difettosi in una confezione. **a)** Con $p = 0{,}02$, calcola $p(X = 0)$, $p(X \leq 1)$, il valore atteso e la deviazione standard di X. **b)** Sempre con $p = 0{,}02$ e confezioni da 50 chip, ricalcola $p(X = 0)$ e $p(X \leq 1)$ con l'approssimazione di Poisson e valuta l'errore commesso. **c)** L'azienda sostituisce gratuitamente ogni confezione che contiene almeno 3 chip difettosi, con un costo di 40 € per sostituzione. Con $p = 0{,}02$ e $X \sim \mathrm{B}(50; 0{,}02)$, qual è il costo medio delle sostituzioni per confezione venduta? **d)** Considera la funzione $g(p) = p(X = 1) = 50p(1 - p)^{49}$ per $p \in [0; 1]$: studiane la monotonia, trova il massimo e interpreta il risultato. Qual è il massimo valore di p per cui la probabilità di una confezione senza difetti è almeno 0,5?

PER APPROFONDIRE

- FEL — il manuale classico della probabilità discreta: i capitoli sulle variabili aleatorie, sulla distribuzione binomiale e sulla sua approssimazione con quella di Poisson contengono molti esempi tratti dalla fisica e dalla biologia.
- BOY — notizie su Poisson e sulle sue *Recherches* del 1837, nel quadro della matematica francese dell'Ottocento.
- QDR — i quadri di riferimento della seconda prova, con l'elenco delle conoscenze di probabilità richieste all'esame.

Glossario del capitolo

Deviazione standard
Radice quadrata della varianza; misura la dispersione di una variabile aleatoria nella stessa unità dei suoi valori.

Distribuzione binomiale
Distribuzione $B(n; p)$ del numero di successi in n prove indipendenti con probabilità di successo p; media np, varianza $np(1 - p)$.

Distribuzione di Poisson
Distribuzione $p(X = k) = \lambda^k e^{-\lambda}/k!$ sui numeri naturali, limite della binomiale $B(n; \lambda/n)$; media e varianza uguali a λ. Modello degli eventi rari.

Distribuzione di probabilità
Funzione che associa a ogni valore x_i di una variabile discreta la probabilità $p(X = x_i)$; le probabilità sono non negative e hanno somma 1.

Distribuzione geometrica
Distribuzione del numero di prove necessarie per ottenere il primo successo; ha valore atteso $1/p$.

Distribuzione ipergeometrica
Distribuzione del numero di elementi favorevoli in n estrazioni senza reimmissione da una popolazione finita.

Distribuzione uniforme discreta
Distribuzione di una variabile che assume n valori, tutti con probabilità $1/n$.

Disuguaglianza di Čebyšev
$p(|X - \mu| \geq k\sigma) \leq 1/k^2$, valida per ogni variabile con varianza finita.

Funzione di ripartizione
La funzione $F(x) = p(X \leq x)$, definita per ogni reale x: non decrescente, compresa fra 0 e 1, a gradini per le variabili discrete.

Gioco equo
Gioco in cui il valore atteso del guadagno netto di ciascun giocatore è zero.

Legge debole dei grandi numeri
La probabilità che la frequenza relativa dei successi differisca dalla probabilità p di almeno ε tende a 0 al crescere del numero delle prove.

Moda
Valore (o valori) di una variabile discreta con la probabilità più alta.

Valore atteso
Media dei valori di una variabile aleatoria pesati con le loro probabilità, $E(X) = \Sigma\, x_i\, p_i$; detto anche speranza matematica.

Variabile aleatoria
Funzione che associa un numero reale a ogni esito di un esperimento aleatorio; si dice discreta se i suoi valori sono finiti o si possono elencare in una successione.

Variabile di Bernoulli
Variabile che vale 1 (successo) con probabilità p e 0 con probabilità $1 - p$; ha media p e varianza $p(1 - p)$.

Variabile standardizzata
La variabile $Z = (X - \mu)/\sigma$, che ha valore atteso 0 e varianza 1.

Variabili aleatorie indipendenti
Variabili tali che $p(X = x, Y = y) = p(X = x) \cdot p(Y = y)$ per ogni coppia di valori.

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 sono non negative e $0{,}3 + 0{,}3 + 0{,}4 = 1$. In A la somma è 1,1, in C c'è un valore negativo, in D la somma è $\frac{13}{12}$.

2. A. $\mathrm{Var}(1 - 2X) = (-2)^2\,\mathrm{Var}(X) = 4 \cdot 4 = 16$; la costante 1 non conta.

3. D. $\mathrm{Var}(X) = np(1-p)$ è, per n fissato, una parabola in p con il vertice in $p = 1/2$.

4. C. Per la Poisson $\mathrm{Var}(X) = \mathrm{E}(X) = 9$, quindi $\sigma = 3$.

5. A. $F(1{,}5) = p(X \le 1{,}5) = p(X = 0) + p(X = 1) = 0{,}7$.

6. a) Falso: per un dado $\mathrm{E}(X) = 3{,}5$, che non è un punteggio possibile. **b)** Falso: vale se X e Y sono indipendenti (per $Y = X$ si ha $\mathrm{Var}(2X) = 4\,\mathrm{Var}(X)$). **c)** Vero: $\mathrm{E}(X) = \mathrm{Var}(X) = \lambda$. **d)** Falso: la moda è l'intero compreso fra $(n+1)p - 1$ e $(n+1)p$; per B(10; 0,25), $np = 2{,}5$ non è nemmeno un valore possibile e la moda è 2.

7. a) $\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2$. **b)** np; $np(1-p)$. **c)** $n \to +\infty$ con $np = \lambda$. **d)** $\frac{1}{9}$.

8. a) Falso: $0 \le F(x) \le 1$ perché $F(x)$ è una probabilità. **b)** Vero. **c)** Vero: $\mathrm{Var}(Z) = 1$. **d)** Vero: $p(X = 1) \approx 0{,}2052$, $p(X = 2) \approx 0{,}2565$, $p(X = 3) \approx 0{,}2138$; la moda è la parte intera di 2,5.

9. $p(X = 0) = \frac{1}{8}$, $p(X = 1) = \frac{3}{8}$, $p(X = 2) = \frac{3}{8}$, $p(X = 3) = \frac{1}{8}$. $F(x) = 0$ per $x < 0$, $\frac{1}{8}$ per $0 \le x < 1$, $\frac{1}{2}$ per $1 \le x < 2$, $\frac{7}{8}$ per $2 \le x < 3$, 1 per $x \ge 3$. $\mathrm{E}(X) = \frac{3}{2}$, $\mathrm{E}(X^2) = 3$, $\mathrm{Var}(X) = 3 - \frac{9}{4} = \frac{3}{4}$ (è B(3; 1/2): $np = 1{,}5$, $np(1-p) = 0{,}75$).

10. a) $M \le k$ se e solo se entrambi i dadi danno al più k: $p(M \le k) = k^2/36$. Quindi $p(M = k) = \frac{k^2 - (k-1)^2}{36} = \frac{2k-1}{36}$. **b)** $\mathrm{E}(M) = \frac{1 \cdot 1 + 2 \cdot 3 + 3 \cdot 5 + 4 \cdot 7 + 5 \cdot 9 + 6 \cdot 11}{36} = \frac{161}{36} \approx 4{,}47$. **c)** $p(M \le 3) = \frac{9}{36} = \frac{1}{4}$.

11. $\mathrm{E}(X) = -0{,}2 + 0 + 0{,}8 + 0{,}5 = 1{,}1$; $\mathrm{E}(X^2) = 0{,}2 + 1{,}6 + 2{,}5 = 4{,}3$; $\mathrm{Var}(X) = 4{,}3 - 1{,}21 = 3{,}09$; $\sigma(X) \approx 1{,}76$; $F(1) = 0{,}2 + 0{,}3 = 0{,}5$; $p(X > 0) = 0{,}4 + 0{,}1 = 0{,}5$.

12. I salti di F danno la distribuzione: $p(X = 1) = 0{,}25$, $p(X = 2) = 0{,}35$, $p(X = 4) = 0{,}4$. $\mathrm{E}(X) = 0{,}25 + 0{,}7 + 1{,}6 = 2{,}55$. $p(1 < X \le 4) = F(4) - F(1) = 1 - 0{,}25 = 0{,}75$.

13. $k(1 + 4 + 9) = 1$, quindi $k = \frac{1}{14}$.
$\mathrm{E}(X) = \frac{1 + 8 + 27}{14} = \frac{18}{7} \approx 2{,}57$;
$\mathrm{E}(X^2) = \frac{1 + 16 + 81}{14} = 7$; $\mathrm{Var}(X) = 7 - \frac{324}{49} = \frac{19}{49} \approx 0{,}39$; $p(X \ge 2) = \frac{13}{14}$.

14. $\mathrm{E}(3X - 2) = 12 - 2 = 10$; $\mathrm{Var}(3X - 2) = 9 \cdot 2{,}5 = 22{,}5$; $\sigma(7 - X) = |-1| \cdot \sqrt{2{,}5} \approx 1{,}58$; $\mathrm{E}(X^2) = \mathrm{Var}(X) + [\mathrm{E}(X)]^2 = 2{,}5 + 16 = 18{,}5$.

15. $X \sim \mathrm{B}(8; 1/4)$. $p(X = 2) = 28 \cdot \frac{1}{16} \cdot (\frac{3}{4})^6 = \frac{5103}{16384} \approx 0{,}3115$; $p(X \ge 1) = 1 - (\frac{3}{4})^8 = \frac{58975}{65536} \approx 0{,}8999$; $\mathrm{E}(X) = 2$; $\sigma(X) = \sqrt{1{,}5} \approx 1{,}22$; $(n+1)p = 2{,}25$, moda 2.

16. $np = 6$ e $np(1-p) = 4{,}2$; dividendo, $1 - p = 0{,}7$, quindi $p = 0{,}3$ e $n = 20$. $p(X = 6) = \binom{20}{6}\, 0{,}3^6 \cdot 0{,}7^{14} \approx 0{,}1916$.

17. $p(X = 0) = e^{-3} \approx 0{,}0498$; $p(X = 2) = 4{,}5\, e^{-3} \approx 0{,}2240$; $p(X \le 2) = (1 + 3 + 4{,}5)\, e^{-3} = 8{,}5\, e^{-3} \approx 0{,}4232$; $p(X \ge 2) = 1 - 4e^{-3} \approx 0{,}8009$.

18. Binomiale: $0{,}99^{200} \approx 0{,}1340$, $200 \cdot 0{,}01 \cdot 0{,}99^{199} \approx 0{,}2707$, $\binom{200}{2}\, 0{,}01^2 \cdot 0{,}99^{198} \approx 0{,}2720$. Poisson con $\lambda = 2$: $e^{-2} \approx 0{,}1353$, $2e^{-2} \approx 0{,}2707$, $2e^{-2} \approx 0{,}2707$. Le differenze non superano 0,0014: con $n = 200$ e $p = 0{,}01$

l'approssimazione è ancora migliore che per B(100; 0,02).

19. a) $p(|X - 50| \geq 10) \leq 16/100 = 0{,}16$, quindi $p(40 < X < 60) \geq 0{,}84$. **b)** Occorre $16/\varepsilon^2 \leq 0{,}04$, cioè $\varepsilon^2 \geq 400$: $\varepsilon = 20$ (cinque deviazioni standard).

20. $\mathrm{Var}(X) = \Sigma\,(x_i - \mu)^2\, p_i$ è una somma di termini non negativi, quindi $\mathrm{Var}(X) \geq 0$; per il teorema $\mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = \mathrm{Var}(X) \geq 0$. L'uguaglianza vale se e solo se ogni termine è nullo, cioè se ogni valore con probabilità positiva è uguale a μ: X è costante (con probabilità 1).

21. I termini con $k = 0$ e $k = 1$ sono nulli; per $k \geq 2$, $k(k-1)/k! = 1/(k-2)!$. Quindi $E[X(X-1)] = \lambda^2\, e^{-\lambda}\, \Sigma_{k\geq 2}\, \lambda^{k-2}/(k-2)! = \lambda^2\, e^{-\lambda}\, e^{\lambda} = \lambda^2$. Allora $\mathrm{E}(X^2) = E[X(X-1)] + \mathrm{E}(X) = \lambda^2 + \lambda$ e $\mathrm{Var}(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$. c.v.d.

22. $\binom{n}{k+1} : \binom{n}{k} = (n-k)/(k+1)$, e il rapporto delle potenze è $p/(1-p)$: si ottiene la formula. Il rapporto è maggiore di 1 se e solo se $(n-k)p > (k+1)(1-p)$, cioè $np - kp > k + 1 - kp - p$, ossia $k < (n+1)p - 1$. Quindi $p(X = k+1) > p(X = k)$ esattamente per quei k, e da lì in poi le probabilità non crescono. Per B(15; 0,3): $(n+1)p = 4{,}8$, la moda è 4 ($p(X = 4) \approx 0{,}2186$, $p(X = 5) \approx 0{,}2061$).

23. La varianza non è lineare: $\mathrm{Var}(X - Y) = \mathrm{Var}(X) + (-1)^2\,\mathrm{Var}(Y) = 3 + 2 = 5$. Togliere una variabile indipendente aggiunge incertezza, non la riduce.

24. $\mathrm{E}(XY) = \mathrm{E}(X^2) = \dfrac{91}{6} \approx 15{,}17$, mentre $\mathrm{E}(X)\mathrm{E}(Y) = \dfrac{49}{4} = 12{,}25$. $\mathrm{Var}(X+Y) = \mathrm{Var}(2X) = 4 \cdot \dfrac{35}{12} = \dfrac{35}{3}$, mentre $\mathrm{Var}(X) + \mathrm{Var}(Y) = \dfrac{35}{6}$. Nessuna contraddizione: l'additività della varianza richiede l'indipendenza, e X e $Y = X$ non sono indipendenti ($p(X = 1, Y = 1) = 1/6 \neq 1/36$). L'additività del valore atteso, $\mathrm{E}(X+Y) = 7 = \mathrm{E}(X) + \mathrm{E}(Y)$, vale invece anche qui.

25. $\mathrm{E}(V) = (1000 + 5 \cdot 200 + 20 \cdot 50) : 1000 = 3$ €. $\mathrm{E}(V^2) = (1\,000\,000 + 5 \cdot 40\,000 + 20 \cdot 2500) : 1000 = 1250$, $\mathrm{Var}(V) = 1250 - 9 = 1241$, $\sigma(V) \approx 35{,}2$ €. Guadagno medio di chi compra: $3 - 5 = -2$ €. Il prezzo equo è 3 €.

26. a) $e^{-0{,}5} \approx 0{,}6065$. **b)** In 4 pagine $\lambda = 2$: $p(X \leq 1) = 3e^{-2} \approx 0{,}4060$. **c)** $200 \cdot e^{-0{,}5} \approx 121$ pagine.

27. $X \sim \mathrm{B}(52; 0{,}9)$: $\mathrm{E}(X) = 46{,}8$, $\sigma(X) = \sqrt{4{,}68} \approx 2{,}16$. Qualcuno resta a terra se $X \geq 51$: $p(X = 51) = 52 \cdot 0{,}9^{51} \cdot 0{,}1 \approx 0{,}0241$, $p(X = 52) = 0{,}9^{52} \approx 0{,}0042$, in totale $\approx 0{,}0283$, poco meno del 3%.

28. Ipergeometrica: $\binom{46}{5} : \binom{50}{5} =$ $1\,370\,754 : 2\,118\,760 \approx 0{,}6470$. Binomiale: $0{,}92^5 \approx 0{,}6591$. L'errore (circa 0,012, il 2% del valore) è accettabile per una stima rapida, ma non trascurabile, perché 5 estrazioni su 50 cambiano la composizione del lotto. Con 5000 pezzi e 400 difettosi l'ipergeometrica dà $\approx 0{,}6590$, praticamente uguale alla binomiale.

29. Traccia. Per $k = 0, 1, 2, 3$ si trova, per la binomiale e per la Poisson con $\lambda = 3$: 0,0388 e 0,0498; 0,1368 e 0,1494; 0,2293 e 0,2240; 0,2428 e 0,2240. La differenza massima è in $k = 3$, circa 0,019, quasi sette volte quella della Tabella 15.2: $p = 0{,}15$ non è piccolo e $n = 20$ non è grande.

30. Traccia. La media dei 1000 conteggi è vicina a $np = 3$ e la varianza a $np(1-p) = 2{,}1$. La media di 1000 conteggi indipendenti ha deviazione standard $\sqrt{2{,}1/1000} \approx 0{,}046$: da una simulazione all'altra la media cambia tipicamente di qualche centesimo, e quasi sempre resta fra 2,9 e 3,1.

31. $a + b + \dfrac{1}{4} = 1$ e $a + 2b + \dfrac{3}{4} = 2$, da cui $b = \dfrac{1}{2}$, $a = \dfrac{1}{4}$. $\mathrm{E}(X^2) = \dfrac{1}{4} + 2 + \dfrac{9}{4} = \dfrac{9}{2}$, $\mathrm{Var}(X) = \dfrac{9}{2} - 4 = \dfrac{1}{2}$.

32. $\lambda\, e^{-\lambda} = \dfrac{\lambda^2}{2}\, e^{-\lambda}$ con $\lambda > 0$ dà $\lambda = 2$. $p(X \geq 3) =$ $1 - e^{-2}(1 + 2 + 2) = 1 - 5e^{-2} \approx 0{,}3233$.

33. $1 - (5/6)^n > 0{,}99$ equivale a $(5/6)^n < 0{,}01$, cioè $n > \ln 0{,}01 \,/\, \ln(5/6) \approx 25{,}3$: servono 26 lanci ($(5/6)^{25} \approx 0{,}0105$, $(5/6)^{26} \approx 0{,}0087$). Il numero di sei è B(26; 1/6), con valore atteso $26/6 = 13/3 \approx 4{,}33$.

34. a) $g(p) = np(1-p)^{n-1}$, $g'(p) = n(1-p)^{n-2}[(1-p) - (n-1)p] =$ $n(1-p)^{n-2}(1 - np)$. In $]0; 1[$ la derivata è positiva per $p < 1/n$ e negativa per $p > 1/n$: il massimo è in $p = 1/n$

(agli estremi $g(0) = g(1) = 0$). **b)** $g(1/n) = (1 - 1/n)^{n-1} = (1 - 1/n)^n : (1 - 1/n)$, che tende a $e^{-1} : 1 = 1/e \approx 0{,}368$.

35. a) $X \sim \mathrm{B}(50; 0{,}02)$: $p(X = 0) = 0{,}98^{50} \approx 0{,}3642$; $p(X = 1) = 50 \cdot 0{,}02 \cdot 0{,}98^{49} \approx 0{,}3716$; $p(X \le 1) \approx 0{,}7358$; $\mathrm{E}(X) = 1$, $\sigma(X) = \sqrt{0{,}98} \approx 0{,}99$. **b)** $\lambda = np = 1$: $p(X = 0) \approx e^{-1} \approx 0{,}3679$ (errore $\approx 0{,}0037$, circa l'1%); $p(X \le 1) \approx 2e^{-1} \approx 0{,}7358$ (errore dell'ordine di un centomillesimo: gli errori su $k = 0$ e $k = 1$ quasi si compensano). **c)** $p(X = 2) = \binom{50}{2} 0{,}02^2 \cdot 0{,}98^{48} \approx 0{,}1858$, quindi $p(X \ge 3) = 1 - 0{,}3642 - 0{,}3716 - 0{,}1858 \approx 0{,}0784$; il costo medio è $40 \cdot 0{,}0784 \approx 3{,}14$ € per confezione. **d)** $g'(p) = 50(1 - p)^{48}[(1 - p) - 49p] = 50(1 - p)^{48}(1 - 50p)$: g cresce in $[0; 0{,}02]$ e decresce in $[0{,}02; 1]$; il massimo è $g(0{,}02) = 0{,}98^{49} \approx 0{,}3716$. La probabilità di trovare esattamente un chip difettoso è massima quando il numero atteso di difettosi, $50p$, vale 1. Infine $p(X = 0) = (1 - p)^{50} \ge 0{,}5$ se e solo se $p \le 1 - 0{,}5^{1/50} \approx 0{,}0138$, cioè circa l'1,38%.