CAPITOLO

10

Probabilità condizionata e formula di Bayes

Un'informazione nuova non cambia i fatti: cambia quello che è ragionevole aspettarsi. (testo della redazione)

OBIETTIVI

1. Calcolare probabilità di estrazioni simultanee, carte e lotto con le combinazioni.
2. Definire la probabilità condizionata e calcolarla da una tabella a doppia entrata, da un albero e dalla definizione.
3. Dimostrare e usare il teorema della probabilità composta; riconoscere eventi indipendenti e distinguerli dagli incompatibili.
4. Applicare il teorema della probabilità totale e la formula di Bayes, anche con le frequenze naturali, a test diagnostici, filtri e controlli di qualità.
5. Usare la formula di Bernoulli per le prove ripetute e trovare il numero più probabile di successi.

INDICE DEL CAPITOLO

10.1 Probabilità e calcolo combinatorio 2
10.2 La probabilità condizionata 5
10.3 Probabilità composta e indipendenza 9
10.4 Probabilità totale e formula di Bayes 10
10.5 Bayes al lavoro: test diagnostici e filtri 13
10.6 Le prove ripetute 16

Matematica 4 · Parte IV · Contare, prevedere, misurare la dipendenza · versione 1.0
Capitolo 10: Probabilità condizionata e formula di Bayes · pagine 267–290 del volume completo · estratto per lo studio

Un test per una malattia rara risulta positivo. Quanto è probabile che la persona sia davvero malata? Quasi tutti rispondono «molto», pensando che un test sia affidabile; spesso, invece, la risposta corretta è «meno di una volta su tre». Una casella di posta riceve un messaggio che contiene la parola «gratis»: è più probabile che sia pubblicità indesiderata o una lettera vera? Un pezzo esce difettoso da una fabbrica con due macchine: da quale delle due viene, più probabilmente? In tutte queste domande c'è un'**informazione** nuova, e la domanda è: come cambia una probabilità quando sappiamo qualcosa in più?

In *Matematica 2* (cap. 8) abbiamo imparato le regole fondamentali del calcolo delle probabilità e la regola del prodotto per gli esperimenti in più fasi; la probabilità condizionata vi compariva solo in un riquadro di approfondimento. In questo capitolo la definiamo con precisione, dimostriamo il teorema della probabilità composta e il teorema della probabilità totale, e arriviamo alla **formula di Bayes**, che rovescia il punto di vista: dagli effetti osservati risale alle cause probabili. Con il calcolo combinatorio del capitolo 9 conteremo i casi senza elencarli, e chiuderemo con le **prove ripetute**, dove i coefficienti binomiali e la regola del prodotto si incontrano in un'unica formula.

10.1 Probabilità e calcolo combinatorio

10.1.1 Che cosa sappiamo già

Richiamiamo in una tabella le nozioni di *Matematica 2* che useremo (Tabella 10.1). Là la probabilità di un evento era scritta $p(A)$; da qui in avanti useremo la lettera maiuscola, $P(A)$, come è d'uso. Dal capitolo 9 prendiamo invece i simboli del calcolo combinatorio: $n!$, le disposizioni $D_{n,k} = n(n-1)\dots(n-k+1)$ e le combinazioni $C_{n,k} = \dfrac{n!}{k!\,(n-k)!}$, cioè il coefficiente binomiale, che nel capitolo 9 hai visto scritto anche come $\binom{n}{k}$. In questo capitolo scriveremo sempre $C_{n,k}$.

Tabella 10.1 Le nozioni di base del calcolo delle probabilità (*Matematica 2*, cap. 8).

Nozione	Che cosa dice
Spazio campionario Ω	l'insieme di tutti i risultati possibili di un esperimento aleatorio; un *evento* è un sottoinsieme di Ω
Definizione classica	$P(A) = \dfrac{\text{casi favorevoli}}{\text{casi possibili}}$, se i casi possibili sono in numero finito ed equiprobabili
Definizione frequentista	la frequenza relativa di A in un gran numero di prove ripetute nelle stesse condizioni
Definizione soggettiva	il prezzo che una persona coerente è disposta a pagare per ricevere 1 se A si verifica e 0 altrimenti
Valori	$0 \le P(A) \le 1$; $P(\Omega) = 1$; $P(\varnothing) = 0$
Evento contrario	$P(\overline{A}) = 1 - P(A)$
Unione	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$; se $A \cap B = \varnothing$ (eventi incompatibili), $P(A \cup B) = P(A) + P(B)$
Esperimenti in più fasi	si moltiplicano le probabilità lungo i rami di un albero (regola del prodotto)
Indipendenza (idea)	il risultato di una fase non influenza le probabilità delle altre, come nei lanci successivi di un dado

Nella colonna di destra, il rapporto della definizione classica è scritto con le parole: per usarlo bisogna *contare*. Quando i casi sono pochi bastano un elenco o un albero; quando sono milioni servono gli strumenti del capitolo 9.

10.1.2 Estrazioni simultanee

Un'urna contiene 10 palline: 4 rosse e 6 bianche. Se ne estraggono 3 **contemporaneamente** (oppure una dopo l'altra senza rimetterle nell'urna, che è lo stesso). Qual è la probabilità che esattamente 2 siano rosse?

Il risultato di un'estrazione simultanea è un gruppo di 3 palline, senza ordine: i casi possibili sono le combinazioni $C_{10,3} = \dfrac{10 \cdot 9 \cdot 8}{3!} = 120$, tutte equiprobabili se le palline sono ben mescolate. Un caso favorevole si costruisce in due fasi: si scelgono 2 rosse tra le 4, in $C_{4,2} = 6$ modi, e 1 bianca tra le 6, in $C_{6,1} = 6$ modi; per il principio di moltiplicazione i casi favorevoli sono $6 \cdot 6 = 36$. Quindi

$$P(\text{2 rosse}) = \frac{C_{4,2} \cdot C_{6,1}}{C_{10,3}} = \frac{36}{120} = 0{,}3$$

Lo stesso ragionamento vale per ogni estrazione simultanea. Se tra N oggetti M sono «marcati» (le palline rosse, gli assi di un mazzo, i numeri giocati al lotto) e se ne estraggono r insieme, la probabilità che esattamente k siano marcati è

$$P(\text{k marcati}) = \frac{C_{M,k} \cdot C_{N-M,r-k}}{C_{N,r}}$$

Si poteva anche contare *con l'ordine*: i casi possibili diventano le disposizioni $D_{10,3} = 720$, e i favorevoli si ottengono scegliendo il posto della pallina bianca (3 modi) e poi riempiendo i posti in ordine, $4 \cdot 3 \cdot 6 = 72$ modi, in tutto 216. Il rapporto $\frac{216}{720}$ è ancora 0,3. Contare con o senza ordine dà lo stesso risultato, purché si conti nello stesso modo al numeratore e al denominatore: è l'errore più comune mescolare combinazioni sopra e disposizioni sotto.

10.1.3 Carte, lotto e compleanni

In Italia si gioca con due tipi di mazzi, e bisogna sempre dire quale: il **mazzo di carte italiane** ha 40 carte, in 4 semi (coppe, denari, bastoni, spade) di 10 carte ciascuno, dall'asso al 7 più tre figure (fante, cavallo, re); il **mazzo di carte francesi** ha 52 carte, in 4 semi di 13 carte (dall'asso al 10, poi fante, donna, re).

ESEMPIO SVOLTO

Carte e lotto. a) Da un mazzo di 40 carte italiane se ne distribuiscono 3. Qual è la probabilità di avere almeno un asso? b) Da un mazzo di 52 carte francesi se ne distribuiscono 5. Qual è la probabilità che siano tutte dello stesso seme? c) Al lotto, su ogni ruota si estraggono 5 numeri tra 1 e 90. Chi gioca un *ambo* sceglie 2 numeri e vince se escono entrambi. Qual è la probabilità di vincere?

a) Conviene l'evento contrario, «nessun asso»: le 3 carte sono scelte tra le 36 che non sono assi. $P(\text{nessun asso}) = \frac{C_{36,3}}{C_{40,3}} = \frac{7140}{9880}$, quindi $P(\text{almeno un asso}) = 1 - \frac{7140}{9880} = \frac{137}{494} \approx 0{,}277$.

b) Si sceglie il seme (4 modi) e poi 5 carte tra le 13 di quel seme: $P = \frac{4 \cdot C_{13,5}}{C_{52,5}} = \frac{4 \cdot 1287}{2\,598\,960} \approx 0{,}00198$, circa una volta ogni 500 mani.

c) Le cinquine possibili sono $C_{90,5} = 43\,949\,268$. Quelle favorevoli contengono i 2 numeri giocati e altri 3 scelti tra gli 88 rimanenti: $C_{88,3} = 109\,736$. Quindi $P(\text{ambo}) = \frac{109\,736}{43\,949\,268} = \frac{2}{801} \approx 0{,}0025$: circa una volta ogni 400 estrazioni. Con lo stesso metodo, un terno giocato con 3 numeri esce con probabilità $\frac{C_{87,2}}{C_{90,5}} = \frac{1}{11\,748}$.

Anche il *problema dei compleanni*, risolto in *Matematica 2* con un prodotto di frazioni, si scrive ora in una riga: la probabilità che n persone compiano gli anni in n giorni tutti diversi (con 365

giorni equiprobabili) è il rapporto tra le disposizioni semplici e quelle con ripetizione, $\frac{D_{365,n}}{D'_{365,n}} = \frac{D_{365,n}}{365^n}$. Per $n = 23$ vale circa 0,493, e la probabilità di almeno una coincidenza è circa 0,507.

10.2 La probabilità condizionata

10.2.1 Un'informazione cambia la probabilità

In un liceo di 500 studenti (dati inventati) si è chiesto a ciascuno se pratica uno sport con regolarità. Le risposte, divise tra biennio e triennio, sono raccolte nella **tabella a doppia entrata** della Tabella 10.2: ogni casella conta gli studenti che hanno entrambe le caratteristiche della sua riga e della sua colonna, e i totali stanno ai margini.

Tabella 10.2 Studenti di un liceo secondo la classe e la pratica sportiva (dati inventati).

	Fa sport (S)	Non fa sport	Totale
Biennio (B)	120	80	200
Triennio (T)	150	150	300
Totale	270	230	500

Scegliamo a caso uno studente. La probabilità che faccia sport è $P(S) = \frac{270}{500} = 0{,}54$. Ma se sappiamo che lo studente è del biennio, i casi possibili non sono più 500: sono i 200 studenti della prima riga, e tra questi 120 fanno sport. La probabilità diventa $\frac{120}{200} = 0{,}6$. L'informazione «è del biennio» ha *ristretto lo spazio campionario* all'evento B, e dentro B i casi favorevoli sono quelli di $S \cap B$.

Dividendo numeratore e denominatore per 500, lo stesso numero si scrive con le probabilità di partenza: $\frac{120}{200} = \frac{120/500}{200/500} = \frac{P(S \cap B)}{P(B)}$. Questa scrittura non dipende più dal conteggio dei casi, e diventa la definizione generale (Figura 10.1).

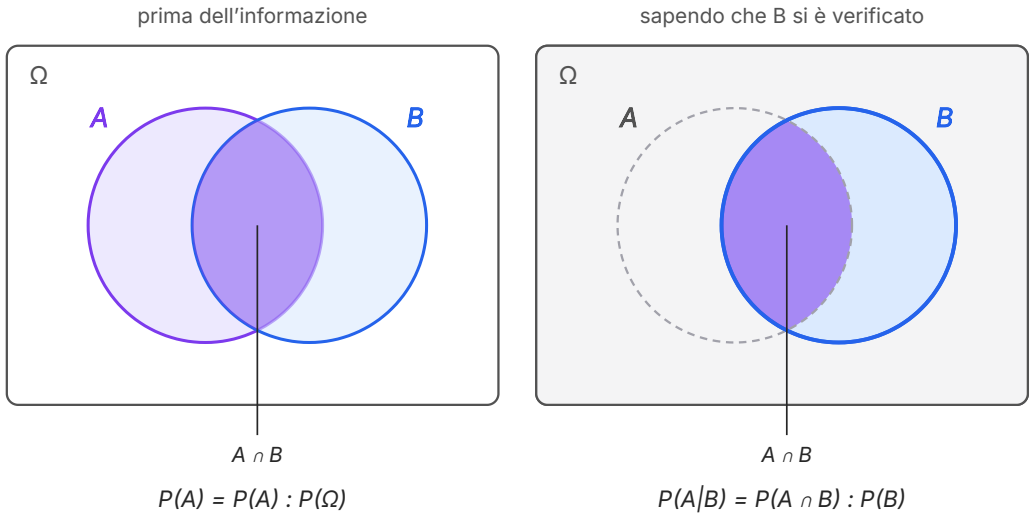

Figura 10.1 A sinistra, due eventi A e B nello spazio campionario Ω. A destra, sapendo che B si è verificato: B diventa il nuovo spazio campionario e di A resta solo la parte $A \cap B$. La probabilità condizionata è la frazione di B occupata da $A \cap B$.

10.2.2 La definizione

IN SINTESI

Probabilità condizionata. Siano A e B due eventi, con $P(B) > 0$. La **probabilità condizionata** di A dato B è

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Si legge «probabilità di A dato B» (o «sapendo che B si è verificato»). Se $P(B) = 0$, $P(A|B)$ non è definita.

Nell'esempio, $P(S|B) = 0{,}6$, mentre per il triennio $P(S|T) = \frac{150}{300} = 0{,}5$. Le due probabilità condizionate sono diverse tra loro e diverse da $P(S) = 0{,}54$: sapere in quale parte della scuola si trova lo studente *cambia* la nostra valutazione. La probabilità $P(S)$, calcolata senza informazioni aggiuntive, si dice anche **probabilità assoluta** (o non condizionata).

ESEMPIO SVOLTO

Due dadi. Si lanciano due dadi equilibrati. a) Sapendo che la somma è 8, qual è la probabilità che almeno uno dei due dadi mostri 6? b) Sapendo che almeno un dado mostra 6, qual è la probabilità che la somma sia 8?

Lo spazio campionario ha 36 coppie equiprobabili. Siano A = «almeno un 6» e E = «la somma è 8».

a) $E = \{(2;6), (3;5), (4;4), (5;3), (6;2)\}$: 5 coppie. Di queste, 2 contengono un 6. Quindi $P(A|E) = \frac{2/36}{5/36} = \frac{2}{5} = 0{,}4$. Senza l'informazione, invece, $P(A) = \frac{11}{36} \approx 0{,}31$.

b) A contiene 11 coppie (6 con il primo dado uguale a 6, 6 con il secondo, meno la coppia (6; 6) contata due volte), e $A \cap E$ è ancora formato dalle 2 coppie (2; 6) e (6; 2). Quindi $P(E|A) = \frac{2}{11} \approx 0{,}18$.

Osservazione. I due risultati sono diversi: scambiare i ruoli dei due eventi cambia il denominatore.

10.2.3 Una probabilità a tutti gli effetti

Fissato l'evento B, la funzione che a ogni evento A associa $P(A|B)$ rispetta tutte le regole di una probabilità: è compresa tra 0 e 1, perché $A \cap B$ è contenuto in B; vale $P(B|B) = 1$, perché ora B è lo spazio campionario; e se A_1, A_2 sono incompatibili, anche $A_1 \cap B$ e $A_2 \cap B$ lo sono, e quindi $P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B)$. In particolare vale la regola dell'evento contrario:

$$P(\overline{A}|B) = 1 - P(A|B)$$

Nella scuola dell'esempio, la probabilità che uno studente del biennio non faccia sport è $1 - 0{,}6 = 0{,}4$, come si legge nella tabella ($\frac{80}{200}$). Tutte le regole che conosciamo si possono quindi usare anche «dentro» un'informazione: basta condizionare ogni probabilità allo stesso evento.

Le probabilità condizionate compaiono anche sugli **alberi**. Nella Figura 10.2 la scuola è descritta come un esperimento in due fasi: prima si sceglie la classe ($P(B) = 0{,}4$, $P(T) = 0{,}6$), poi si guarda se lo studente fa sport. I rami della seconda fase portano probabilità condizionate: su quelli che partono da B c'è $P(S|B) = 0{,}6$, non $P(S)$. Le probabilità alle foglie, ottenute moltiplicando lungo i rami, sono le frequenze relative delle quattro caselle interne della tabella: $0{,}24 = \frac{120}{500}$, $0{,}16 = \frac{80}{500}$, $0{,}30 = \frac{150}{500}$ e ancora 0,30.

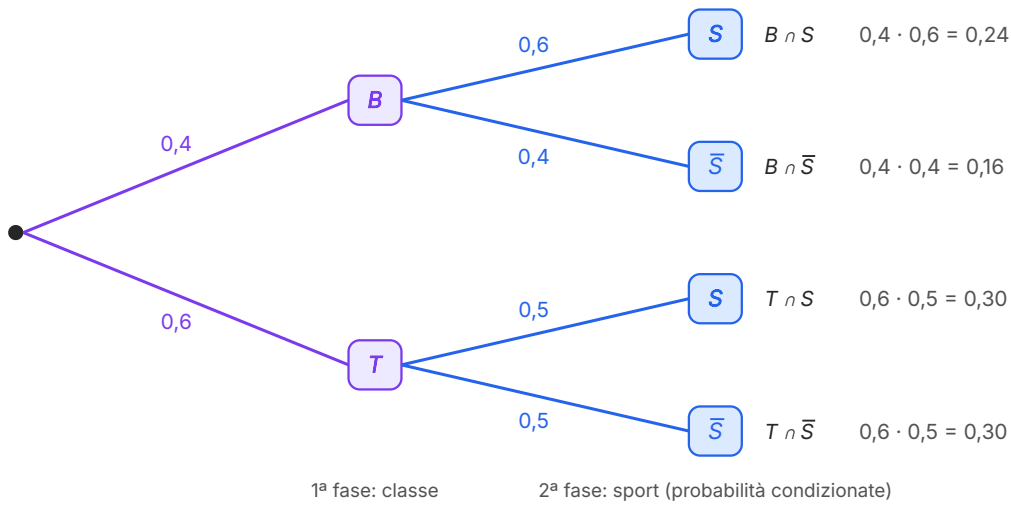

Figura 10.2 La Tabella 10.2 come albero. Sui rami della prima fase ci sono probabilità assolute, su quelli della seconda probabilità condizionate; ogni foglia porta il prodotto delle probabilità lungo il suo percorso, e le quattro foglie sommano a 1.

ERRORE FREQUENTE

Scambiare $P(A|B)$ con $P(B|A)$. Nella scuola, la probabilità che uno studente del biennio faccia sport è $P(S|B) = 0{,}6$; la probabilità che uno sportivo sia del biennio è $P(B|S) = \frac{120}{270} \approx 0{,}44$. Sono domande diverse, con risposte diverse.

Lo scambio può avere conseguenze serie. In un processo (esempio inventato) si trova sulla scena del crimine una traccia che si presenta in una persona su 10 000, e l'imputato ce l'ha. L'accusa dice: «la probabilità che un innocente abbia questa traccia è 1 su 10 000, quindi l'imputato è colpevole con probabilità 0,9999». È la **fallacia del procuratore**: 1 su 10 000 è $P(\text{traccia}|\text{innocente})$, mentre il giudice deve valutare $P(\text{innocente}|\text{traccia})$. Se il colpevole potrebbe essere uno qualunque dei 500 000 adulti di una città, circa 50 di loro hanno la traccia: senza altre prove, l'imputato è solo uno di questi 50, e la probabilità che sia il colpevole è circa $\frac{1}{50}$.

10.3 Probabilità composta e indipendenza

10.3.1 Il teorema della probabilità composta

Moltiplicando per $P(B)$ i due membri della definizione si ottiene un modo per calcolare la probabilità che due eventi si verifichino *entrambi*. È la regola del prodotto che in *Matematica 2* abbiamo usato in modo intuitivo; ora possiamo dimostrarla.

IN SINTESI

Teorema della probabilità composta.

Ipotesi: A e B sono eventi dello stesso spazio campionario, con $P(B) > 0$.

Tesi: $P(A \cap B) = P(B) \cdot P(A|B)$. Se anche $P(A) > 0$, vale pure $P(A \cap B) = P(A) \cdot P(B|A)$.

Dimostrazione. Per definizione $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$; moltiplicando per $P(B)$, che è positivo, si ha $P(B) \cdot P(A|B) = P(A \cap B)$. Scambiando i nomi dei due eventi, e ricordando che $B \cap A = A \cap B$, si ottiene la seconda uguaglianza. c.v.d.

Applicando il teorema due volte si passa a tre eventi: $P(A \cap B \cap C) = P(A) \cdot P(B|A) \cdot P(C|A \cap B)$, e così via. Per esempio, la probabilità che 3 carte distribuite da un mazzo italiano siano tutte di denari è

$$\frac{10}{40} \cdot \frac{9}{39} \cdot \frac{8}{38} = \frac{720}{59\,280} = \frac{3}{247} \approx 0{,}012$$

dove ogni fattore è una probabilità condizionata: la seconda carta è di denari con probabilità $\dfrac{9}{39}$ *sapendo che* la prima lo era. Contando con le combinazioni si trova lo stesso numero, $\dfrac{C_{10,3}}{C_{40,3}} = \dfrac{120}{9880} = \dfrac{3}{247}$: i due metodi si controllano a vicenda.

10.3.2 Eventi indipendenti

Può accadere che l'informazione B non cambi la probabilità di A: $P(A|B) = P(A)$. Diremo allora che A è **indipendente** da B. Per il teorema della probabilità composta, se $P(B) > 0$:

$$P(A|B) = P(A) \Leftrightarrow P(A \cap B) = P(A) \cdot P(B)$$

Infatti $P(A \cap B) = P(B) \cdot P(A|B)$, e il secondo membro è uguale a $P(B) \cdot P(A)$ esattamente quando $P(A|B) = P(A)$. La forma a destra è simmetrica in A e B, vale anche quando una delle due probabilità è nulla, e si usa come **definizione**: due eventi sono **indipendenti** se $P(A \cap B) = P(A) \cdot P(B)$. In particolare, se A è indipendente da B, anche B è indipendente da A.

L'indipendenza è un fatto numerico, da verificare con il calcolo, non un'impressione. Lanciamo due dadi e sia A = «il primo dado è pari». Con E_7 = «la somma è 7» si ha $P(A \cap E_7) = \dfrac{3}{36}$ (le coppie (2; 5), (4; 3), (6; 1)) e $P(A) \cdot P(E_7) = \dfrac{1}{2} \cdot \dfrac{6}{36} = \dfrac{3}{36}$: gli eventi sono indipendenti, anche se la somma «coinvolge» il primo dado. Con E_8 = «la somma è 8», invece, $P(A \cap E_8) = \dfrac{3}{36}$ (le coppie (2; 6), (4; 4), (6; 2)), mentre $P(A) \cdot P(E_8) = \dfrac{1}{2} \cdot \dfrac{5}{36} = \dfrac{5}{72}$: non sono indipendenti. Allo stesso modo, in un mazzo italiano gli eventi «la carta è un asso» e «la carta è di coppe» sono indipendenti: $\dfrac{1}{40} = \dfrac{4}{40} \cdot \dfrac{10}{40}$.

Nella scuola della Tabella 10.2, $P(S|B) = 0{,}6 \neq P(S) = 0{,}54$: la pratica sportiva e la classe frequentata *non* sono indipendenti. Nel capitolo 11 ritroveremo questa idea per i dati statistici: due caratteri sono indipendenti quando le distribuzioni condizionate sono tutte uguali.

ERRORE FREQUENTE

Incompatibili non vuol dire indipendenti. Due eventi *incompatibili* non possono verificarsi insieme: $A \cap B = \varnothing$. Se hanno entrambi probabilità positiva, allora $P(A \cap B) = 0$ mentre $P(A) \cdot P(B) > 0$: **non** sono indipendenti. Anzi, la dipendenza è la più forte possibile: sapere che B si è verificato rende A impossibile, $P(A|B) = 0$. Per esempio, con un dado, «esce 1» ed «esce 6» sono incompatibili e fortemente dipendenti. L'incompatibilità si vede sul diagramma di Venn (i due insiemi non si toccano); l'indipendenza si verifica solo con i numeri.

10.4 Probabilità totale e formula di Bayes

10.4.1 Partizioni e probabilità totale

Spesso un evento può verificarsi per strade diverse. Un pezzo difettoso può venire dalla prima o dalla seconda macchina di una fabbrica; un test può risultare positivo su una persona malata o su una sana. Per calcolarne la probabilità conviene dividere lo spazio campionario secondo le possibili «cause».

Si chiama **partizione** dello spazio campionario Ω una famiglia di eventi $H_1, H_2, \ldots, H_n$ a due a due incompatibili, ciascuno con probabilità positiva, la cui unione è Ω: in ogni prova si verifica uno e uno solo degli H_i, quindi $P(H_1) + P(H_2) + \ldots + P(H_n) = 1$. Un qualunque evento E viene allora diviso nei pezzi $E \cap H_1$, $E \cap H_2$, ..., a due a due incompatibili (Figura 10.3).

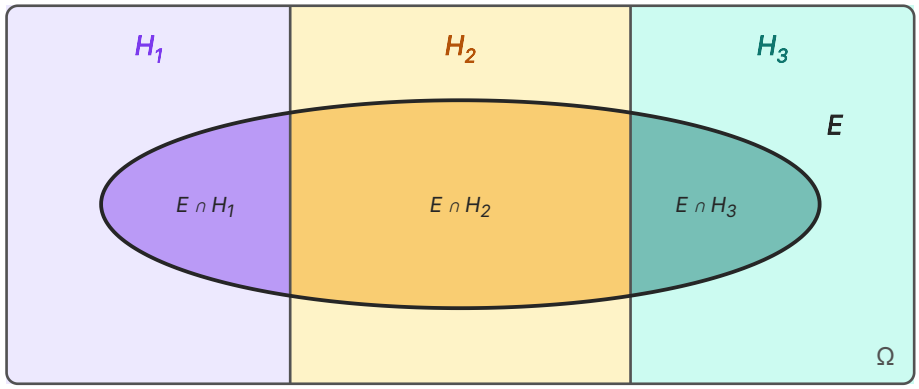

Figura 10.3 Una partizione di Ω in tre eventi H_1, H_2, H_3. L'evento E è diviso in tre parti incompatibili, una per ciascun H_i; la sua probabilità è la somma delle probabilità delle parti.

Sommando le probabilità dei pezzi e calcolando ciascuno con il teorema della probabilità composta, $P(E \cap H_i) = P(H_i) \cdot P(E|H_i)$, si ottiene il **teorema della probabilità totale**:

$$P(E) = P(H_1) \cdot P(E|H_1) + P(H_2) \cdot P(E|H_2) + \ldots + P(H_n) \cdot P(E|H_n)$$

Sull'albero, è la somma dei prodotti lungo tutti i percorsi che terminano con E. Con la partizione più semplice, B e $\overline{B}$, la formula diventa $P(E) = P(B) \cdot P(E|B) + P(\overline{B}) \cdot P(E|\overline{B})$: nella scuola, $P(S) = 0{,}4 \cdot 0{,}6 + 0{,}6 \cdot 0{,}5 = 0{,}54$, come nella tabella.

10.4.2 La formula di Bayes

Il teorema della probabilità totale va dalle cause agli effetti: conoscendo le probabilità di ciascuna causa H_i e la probabilità dell'effetto E per ciascuna causa, calcola la probabilità dell'effetto. Spesso però la domanda è rovesciata: abbiamo *osservato* l'effetto E, e ci chiediamo quale causa sia la più probabile, cioè vogliamo $P(H_i|E)$. Per la definizione e per il teorema della probabilità composta, $P(H_i|E) = \dfrac{P(H_i \cap E)}{P(E)} = \dfrac{P(H_i) \cdot P(E|H_i)}{P(E)}$; al denominatore mettiamo la probabilità totale.

IN SINTESI

Probabilità totale e formula di Bayes. Siano $H_1, \dots, H_n$ una partizione di Ω ed E un evento con $P(E) > 0$. Allora

$$P(E) = \sum_j P(H_j) \cdot P(E|H_j)$$

e, per ogni i, vale la **formula di Bayes**

$$P(H_i|E) = \frac{P(H_i) \cdot P(E|H_i)}{P(H_1) \cdot P(E|H_1) + \dots + P(H_n) \cdot P(E|H_n)}$$

$P(H_i)$ è la **probabilità a priori** della causa H_i, $P(E|H_i)$ la **verosimiglianza** dell'effetto osservato, $P(H_i|E)$ la **probabilità a posteriori**. Le probabilità a posteriori, come quelle a priori, hanno somma 1.

La formula ha un significato semplice: la probabilità a posteriori di ciascuna causa è proporzionale al prodotto «probabilità a priori × verosimiglianza», e il denominatore serve solo a fare in modo che le probabilità sommino a 1. Una causa diventa più probabile se rende l'effetto osservato più probabile delle altre.

ESEMPIO SVOLTO

Due macchine. In una fabbrica (dati inventati) la macchina M_1 produce il 60% dei pezzi, di cui il 3% difettosi; la macchina M_2 produce il restante 40%, di cui il 5% difettosi. Si sceglie a caso un pezzo. a) Qual è la probabilità che sia difettoso? b) Il pezzo è difettoso: qual è la probabilità che venga da M_1? E da M_2?

Sia D = «il pezzo è difettoso». La partizione è M_1, M_2, con $P(M_1) = 0{,}6$, $P(M_2) = 0{,}4$, $P(D|M_1) = 0{,}03$, $P(D|M_2) = 0{,}05$.

a) Probabilità totale: $P(D) = 0{,}6 \cdot 0{,}03 + 0{,}4 \cdot 0{,}05 = 0{,}018 + 0{,}020 = 0{,}038$, cioè il 3,8% dei pezzi.

b) Formula di Bayes: $P(M_1|D) = \dfrac{0{,}018}{0{,}038} = \dfrac{9}{19} \approx 0{,}474$ e $P(M_2|D) = \dfrac{0{,}020}{0{,}038} = \dfrac{10}{19} \approx 0{,}526$. Controllo: $\dfrac{9}{19} + \dfrac{10}{19} = 1$.

Commento. A priori il pezzo viene più probabilmente da M_1 (0,6 contro 0,4); dopo aver visto che è difettoso, la causa più probabile diventa M_2, perché produce difetti con frequenza maggiore.

10.4.3 Monty Hall rivisitato

In *Matematica 2* abbiamo incontrato il problema di Monty Hall: dietro una di tre porte c'è un'automobile, dietro le altre due una capra; il concorrente sceglie una porta, il conduttore, che sa

dov'è l'auto, ne apre un'altra con una capra e offre di cambiare. Conviene cambiare, e la formula di Bayes lo mostra in poche righe.

Supponiamo che il concorrente scelga la porta 1 e che il conduttore apra la porta 3. Le cause sono H_1, H_2, H_3 = «l'auto è dietro la porta 1, 2, 3», ciascuna con probabilità a priori $\frac{1}{3}$; l'effetto è E = «il conduttore apre la porta 3». Se l'auto è dietro la porta 1, il conduttore può aprire la 2 o la 3, e sceglie a caso: $P(E|H_1) = \frac{1}{2}$. Se è dietro la 2, è costretto ad aprire la 3: $P(E|H_2) = 1$. Se è dietro la 3, non la apre mai: $P(E|H_3) = 0$. La probabilità totale è $P(E) = \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot 1 + 0 = \frac{1}{2}$, e quindi

$$P(H_1|E) = \frac{1/6}{1/2} = \frac{1}{3}, \quad P(H_2|E) = \frac{1/3}{1/2} = \frac{2}{3}, \quad P(H_3|E) = 0$$

Cambiando porta il concorrente vince con probabilità $\frac{2}{3}$. Il calcolo mostra anche *perché*: l'apertura della porta 3 è un'informazione che il conduttore dà sapendo dov'è l'auto, e la sua verosimiglianza è doppia se l'auto è dietro la porta 2.

APPROFONDIMENTO · BAYES, LAPLACE E LA PROBABILITÀ DELLE CAUSE

Thomas Bayes era un pastore presbiteriano inglese, appassionato di matematica, morto nel 1761. Tra le sue carte l'amico Richard Price trovò un saggio sul «problema inverso» della probabilità: dato che un evento si è verificato un certo numero di volte in un certo numero di prove, che cosa si può dire della sua probabilità sconosciuta? Price lo fece pubblicare nel 1763 dalla Royal Society di Londra. La formula che oggi porta il nome di Bayes vi compare solo in una forma particolare.

Pochi anni dopo, in modo indipendente, il francese Pierre-Simon Laplace formulò il principio in generale e lo chiamò calcolo della *probabilità delle cause*: dagli eventi osservati si risale alle cause che li hanno prodotti, ciascuna pesata per la probabilità che dà a ciò che si è visto. Laplace lo applicò a problemi di astronomia e di demografia, e fece della probabilità uno strumento per ragionare in condizioni di incertezza in tutte le scienze.

10.5 Bayes al lavoro: test diagnostici e filtri

10.5.1 Sensibilità, specificità, prevalenza

Un test diagnostico divide le persone in *positive* (+) e *negative* (−), ma nessun test è perfetto. Chiamiamo M l'evento «la persona ha la malattia». Tre numeri descrivono la situazione:

- la **prevalenza** $P(M)$, cioè la frazione della popolazione che ha la malattia;
- la **sensibilità** $P(+|M)$, cioè la probabilità che il test sia positivo su una persona malata;

- la **specificità** $P(-|\overline{M})$, cioè la probabilità che il test sia negativo su una persona sana.

Un positivo su una persona sana è un **falso positivo**, e capita con probabilità 1 − specificità; un negativo su una persona malata è un **falso negativo**. Ma a chi riceve un risultato positivo non interessa nessuno di questi tre numeri: interessa $P(M|+)$, la probabilità di essere malato sapendo di essere positivo, che si chiama **valore predittivo positivo**. È una probabilità «rovesciata», e si calcola con Bayes.

Prendiamo un test con sensibilità 90% e specificità 95%, per una malattia con prevalenza 2% (numeri inventati, ma di un ordine di grandezza realistico). Con la formula:

$$P(M|+) = \frac{0{,}02 \cdot 0{,}90}{0{,}02 \cdot 0{,}90 + 0{,}98 \cdot 0{,}05} = \frac{0{,}018}{0{,}018 + 0{,}049} = \frac{0{,}018}{0{,}067} \approx 0{,}27$$

Solo il 27% dei positivi è malato. Il risultato sorprende quasi tutti, medici compresi, ed è più facile da capire con le **frequenze naturali**: invece delle probabilità, si immagina un gruppo di persone e si conta. Su 10 000 persone sottoposte al test, 200 sono malate (il 2%) e 9800 sane; delle malate, 180 risultano positive; delle sane, il 5%, cioè 490, risultano positive per errore (Figura 10.4). I positivi sono in tutto 180 + 490 = 670, e tra loro i malati sono 180: $\frac{180}{670} \approx 0{,}27$. È lo stesso calcolo della formula di Bayes, con numeri interi al posto delle probabilità.

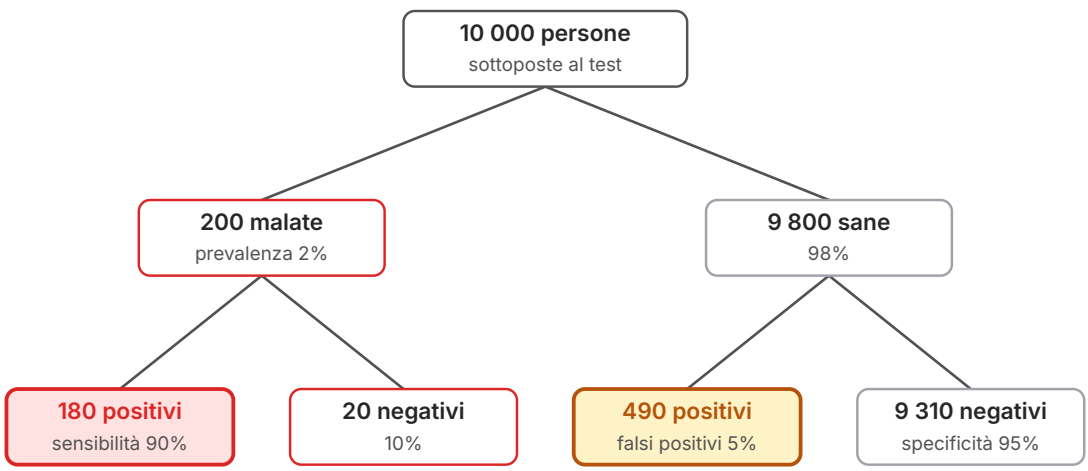

Figura 10.4 Il test diagnostico con le frequenze naturali (dati inventati: prevalenza 2%, sensibilità 90%, specificità 95%). I falsi positivi, il 5% di una folla di sani, sono più numerosi dei veri positivi, il 90% di pochi malati.

La ragione del paradosso è che i malati sono pochi: anche una piccola frazione di errori sui moltissimi sani produce più positivi di tutti i malati messi insieme. La Figura 10.5 lo mostra a

colpo d'occhio su 1000 persone. Lo stesso test è invece molto affidabile quando è *negativo*: $P(\overline{M}|-) = \dfrac{9310}{9310 + 20} \approx 0{,}998$.

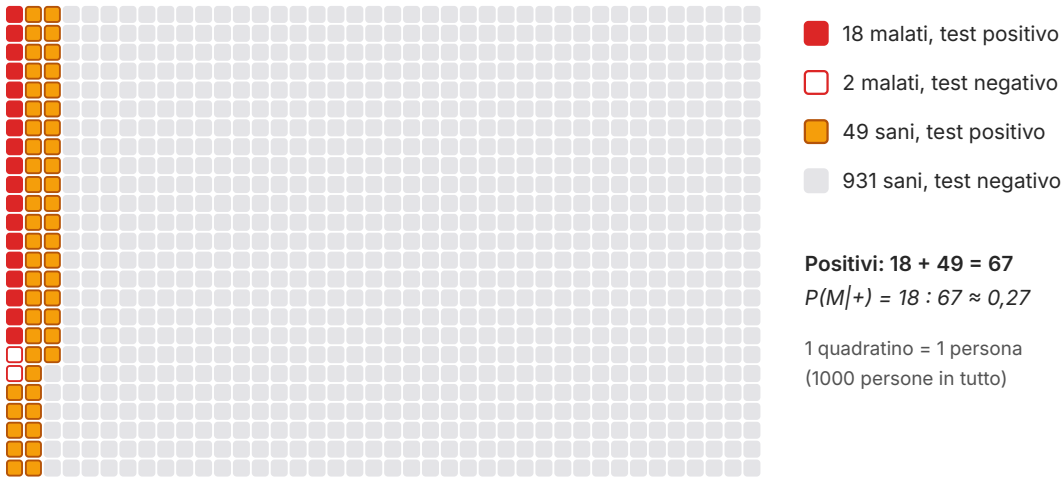

Figura 10.5 Le stesse proporzioni della Figura 10.4 su 1000 persone, una per quadratino: 20 malati (in rosso), 980 sani. I positivi sono i 18 quadratini rossi pieni e i 49 arancioni.

10.5.2 Un secondo test

Che cosa si fa, allora, dopo un risultato positivo? Si ripete l'esame, o se ne fa un altro. La formula di Bayes si può applicare di nuovo, usando come probabilità a priori la probabilità a posteriori del primo passo: *ciò che si è imparato diventa il punto di partenza*. Se il secondo test ha le stesse caratteristiche e, per ogni persona, il suo esito è indipendente da quello del primo una volta noto lo stato di salute (è un'ipotesi del modello, da verificare in medicina), per chi risulta positivo due volte

$$P(M|{+}{+}) = \frac{0{,}27 \cdot 0{,}90}{0{,}27 \cdot 0{,}90 + 0{,}73 \cdot 0{,}05} \approx 0{,}87$$

(con il valore esatto $\dfrac{18}{67}$ al posto di 0,27 si ottiene ancora 0,87). Due risultati positivi rendono la malattia molto probabile; un solo positivo, in una malattia rara, è soprattutto un motivo per fare altri esami.

10.5.3 Il filtro antispam

Gli stessi conti sono alla base dei filtri che separano la posta indesiderata (*spam*) da quella vera. Supponiamo (dati inventati) che il 40% dei messaggi che arrivano a una casella sia spam, che la

parola «gratis» compaia nel 25% dei messaggi di spam e solo nel 2% degli altri. Un messaggio contiene «gratis»: la probabilità che sia spam è

$$P(\text{spam}|\text{gratis}) = \frac{0{,}40 \cdot 0{,}25}{0{,}40 \cdot 0{,}25 + 0{,}60 \cdot 0{,}02} = \frac{0{,}100}{0{,}112} \approx 0{,}89$$

Un filtro reale non guarda una sola parola, ma ne combina moltissime, aggiornando la probabilità a ogni parola letta, come abbiamo fatto con il secondo test; e impara le verosimiglianze dai messaggi che l'utente ha già classificato.

COLLEGAMENTI

Filosofia: la probabilità come grado di fiducia. In *Matematica 2* abbiamo incontrato la definizione soggettiva: la probabilità di un evento è il grado di fiducia che una persona coerente ha nel suo verificarsi, misurato dal prezzo di una scommessa. Il matematico italiano Bruno de Finetti (1906–1985) ne fece il fondamento di tutta la teoria. In questa lettura la formula di Bayes acquista un significato filosofico: è la regola con cui un soggetto razionale *rivede le proprie convinzioni alla luce dell'esperienza*. La probabilità a priori è ciò che si crede prima; l'osservazione pesa le ipotesi con le loro verosimiglianze; la probabilità a posteriori è ciò che si deve credere dopo.

Che il ragionamento sotto incertezza abbia una logica è un'idea antica. Nei *Pensieri*, pubblicati dopo la sua morte, Blaise Pascal sostiene la sua celebre «scommessa» confrontando ciò che si guadagna e ciò che si perde secondo come stanno le cose: un uso del valore atteso per decidere. Chi quest'anno incontra in filosofia Pascal, Hume e Kant può chiedersi: le probabilità descrivono il mondo (le frequenze) o la nostra conoscenza del mondo (i gradi di fiducia)? La matematica, con gli stessi assiomi, serve entrambe le risposte.

10.6 Le prove ripetute

10.6.1 Lo schema di Bernoulli

Molti esperimenti consistono nel ripetere più volte la stessa prova, che ha solo due esiti: il lancio di un dado in cui interessa solo se esce 6, una serie di tiri liberi a canestro, il controllo di pezzi prodotti da una macchina. Chiamiamo **successo** l'esito che ci interessa e **insuccesso** l'altro. Si parla di **prove ripetute** (o *schema di Bernoulli*) quando le prove sono n, sono indipendenti e in ciascuna il successo ha la stessa probabilità p.

Lanciamo un dado 5 volte: qual è la probabilità di ottenere esattamente due 6? Una sequenza particolare, per esempio «6, 6, non 6, non 6, non 6», ha probabilità $\frac{1}{6} \cdot \frac{1}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} = \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^3$, per l'indipendenza dei lanci. Ogni altra sequenza con due 6 e tre non-6 ha la stessa probabilità, perché contiene gli stessi fattori in

un altro ordine. Quante sono? Una sequenza è individuata dalle posizioni dei due successi, scelte tra le 5: sono $C_{5,2} = 10$. Le sequenze sono incompatibili, quindi le probabilità si sommano:

$$P(\text{due 6 in 5 lanci}) = C_{5,2} \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^3 = \frac{10 \cdot 125}{7776} = \frac{625}{3888} \approx 0{,}161$$

Lo stesso ragionamento, con n prove e k successi, dà la formula generale, che porta il nome di Jakob Bernoulli (di cui in *Matematica 2* abbiamo incontrato la legge dei grandi numeri).

IN SINTESI

Formula di Bernoulli. In n prove ripetute indipendenti, ciascuna con probabilità di successo p, la probabilità di ottenere esattamente k successi ($k = 0, 1, \ldots, n$) è

$$P(\text{k successi}) = C_{n,k} \cdot p^k \cdot (1-p)^{n-k}$$

Il coefficiente binomiale conta le sequenze con k successi; il prodotto delle potenze è la probabilità di ciascuna sequenza. Per $k = 0$ la formula dà $(1-p)^n$, e quindi $P(\text{almeno un successo}) = 1 - (1-p)^n$.

10.6.2 Il numero più probabile di successi

Quanti successi aspettarsi? Calcoliamo tutte le probabilità per $n = 10$ prove con $p = 0{,}3$, per esempio 10 tiri da tre punti di un giocatore che segna il 30% delle volte (Figura 10.6).

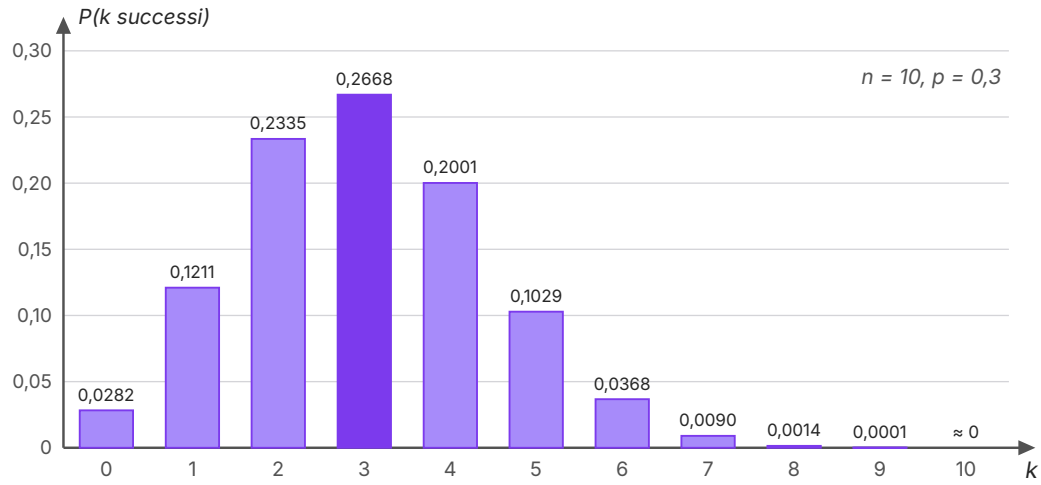

Figura 10.6 Le probabilità di k successi in 10 prove ripetute con $p = 0{,}3$, calcolate con la formula di Bernoulli e arrotondate a quattro cifre. Il numero più probabile di successi è 3, cioè np; le probabilità esatte sommano a 1.

Il valore più probabile è $k = 3$, con probabilità $C_{10,3} \cdot 0{,}3^3 \cdot 0{,}7^7 \approx 0{,}267$: coincide con $np = 10 \cdot 0{,}3$. Ma attenzione: anche il valore più probabile ha probabilità molto minore di 1, e 2 o 4 successi hanno probabilità non molto minori (circa 0,23 e 0,20). La probabilità di almeno un successo è $1 - 0{,}7^{10} \approx 0{,}972$. In generale il numero più probabile di successi è vicino a np; lo si trova calcolando la tabella delle probabilità e scegliendo il valore massimo.

Sommando le probabilità per k da 0 a n si ottiene sempre 1: per la formula del binomio di Newton del capitolo 9, la somma è $(p + (1 - p))^n = 1^n = 1$. I coefficienti binomiali della formula di Bernoulli sono i numeri di una riga del triangolo di Tartaglia: per $n = 10$, 1, 10, 45, 120, 210, 252, 210, 120, 45, 10, 1.

ESEMPIO SVOLTO

Un esame a crocette. Un test ha 8 domande, ciascuna con 4 risposte di cui una sola esatta. Uno studente risponde a caso a tutte. a) Qual è la probabilità che indovini esattamente 2 risposte? b) Qual è la probabilità che ne indovini almeno 6 (soglia di sufficienza)? c) Qual è il numero più probabile di risposte esatte?

Le prove sono $n = 8$, indipendenti, con $p = \frac{1}{4}$ e $1 - p = \frac{3}{4}$.

a) $P(2) = C_{8,2} \cdot \left(\frac{1}{4}\right)^2 \left(\frac{3}{4}\right)^6 = 28 \cdot \frac{729}{65\,536} = \frac{20\,412}{65\,536} \approx 0{,}311$.

b) $P(6) = 28 \cdot \frac{9}{65\,536}$, $P(7) = 8 \cdot \frac{3}{65\,536}$, $P(8) = \frac{1}{65\,536}$; in tutto $\frac{252 + 24 + 1}{65\,536} = \frac{277}{65\,536} \approx 0{,}0042$: rispondere a caso non conviene.

c) $P(0) = \frac{6561}{65\,536} \approx 0{,}100$, $P(1) = \frac{17\,496}{65\,536} \approx 0{,}267$, $P(2) \approx 0{,}311$, $P(3) = \frac{56 \cdot 243}{65\,536} \approx 0{,}208$, e da lì le probabilità diminuiscono: il numero più probabile è 2, vicino a $np = 2$.

APPROFONDIMENTO · VERSO LA DISTRIBUZIONE BINOMIALE

La tabella dei valori di P(k successi) al variare di k è un primo esempio di *distribuzione di probabilità*: a ogni possibile numero di successi associa la sua probabilità, come una distribuzione di frequenze associa a ogni valore la sua frequenza. Si chiama **distribuzione binomiale**. Nel quinto anno studierai le distribuzioni di probabilità in generale, con il loro valore medio (per la binomiale è proprio np) e la loro dispersione, e vedrai che quando n è grande il diagramma a barre della binomiale assume una forma a campana.

Lo schema delle prove ripetute è un **modello**: chiede che le prove siano indipendenti e che la probabilità di successo non cambi. Per i lanci di un dado le ipotesi sono ragionevoli; per i tiri di un giocatore stanco o per estrazioni senza reimmissione da un'urna piccola non lo sono, e allo-

ra si usano altri modelli, come quello delle estrazioni simultanee del primo paragrafo. Scegliere il modello giusto, e saper dire quali ipotesi assume, è la parte più importante del lavoro.

SINTESI

- Nelle estrazioni simultanee i casi sono combinazioni: la probabilità di k oggetti marcati tra r estratti, su N oggetti di cui M marcati, è $\frac{C_{M,k} \cdot C_{N-M,r-k}}{C_{N,r}}$. Si può contare anche con l'ordine, purché sopra e sotto allo stesso modo.
- Il mazzo di carte italiane ha 40 carte, quello francese 52: bisogna sempre dire quale si usa.
- Se $P(B) > 0$, la probabilità di A dato B è $P(A|B) = \frac{P(A \cap B)}{P(B)}$: l'informazione restringe lo spazio campionario a B.
- Fissato B, $P(\cdot|B)$ è una probabilità: in particolare $P(\overline{A}|B) = 1 - P(A|B)$. Sugli alberi, i rami dalla seconda fase in poi portano probabilità condizionate.
- In generale $P(A|B) \neq P(B|A)$: confonderle è la fallacia del procuratore.
- Teorema della probabilità composta: $P(A \cap B) = P(B) \cdot P(A|B) = P(A) \cdot P(B|A)$; per tre eventi $P(A \cap B \cap C) = P(A) \cdot P(B|A) \cdot P(C|A \cap B)$.
- A e B sono indipendenti se $P(A \cap B) = P(A) \cdot P(B)$, cioè (con $P(B) > 0$) se $P(A|B) = P(A)$. Due eventi incompatibili con probabilità positiva non sono mai indipendenti.
- Una partizione è una famiglia di eventi a due a due incompatibili, con probabilità positive, che riempiono Ω. Probabilità totale: $P(E) = \sum P(H_i) \cdot P(E|H_i)$.
- Formula di Bayes: $P(H_i|E) = \frac{P(H_i) \cdot P(E|H_i)}{P(E)}$; dalle probabilità a priori e dalle verosimiglianze si passa alle probabilità a posteriori, che sommano a 1.
- In un test diagnostico, sensibilità $P(+|M)$, specificità $P(-|\overline{M})$ e prevalenza $P(M)$ danno con Bayes il valore predittivo $P(M|+)$, che per una malattia rara può essere basso; le frequenze naturali rendono il calcolo intuitivo.
- La probabilità a posteriori diventa la nuova probabilità a priori quando arriva un'altra informazione: è il modo in cui Bayes descrive l'apprendimento dall'esperienza.
- Formula di Bernoulli: in n prove indipendenti con probabilità di successo p, $P(\text{k successi}) = C_{n,k}\, p^k\, (1-p)^{n-k}$; il numero più probabile di successi è vicino a np.

ESERCIZI

Comprendere

1. ● Se $P(B) > 0$, la probabilità condizionata $P(A|B)$ è uguale a:
   - **A.** $\frac{P(A \cap B)}{P(A)}$
   - **B.** $\frac{P(A \cap B)}{P(B)}$
   - **C.** $P(A) \cdot P(B)$
   - **D.** $\frac{P(A)}{P(B)}$

2. ● Da un mazzo di 40 carte italiane si estrae una carta. Sapendo che è una figura (fante, cavallo o re), la probabilità che sia un re è:
   - **A.** $\frac{1}{10}$
   - **B.** $\frac{1}{12}$
   - **C.** $\frac{1}{3}$
   - **D.** $\frac{3}{10}$

3. ●● A e B sono eventi incompatibili con $P(A) = 0{,}3$ e $P(B) = 0{,}4$. Allora:
   - **A.** $P(A|B) = 0$
   - **B.** A e B sono indipendenti
   - **C.** $P(A \cap B) = 0{,}12$
   - **D.** $P(A|B) = 0{,}3$

4. ●● A e B sono indipendenti, con $P(A) = 0{,}5$ e $P(B) = 0{,}4$. Allora $P(A \cap B)$ vale:
   - **A.** 0,9
   - **B.** 0,1
   - **C.** 0,45
   - **D.** 0,2

5. ●● Si lancia 4 volte una moneta equilibrata. La probabilità di ottenere esattamente 2 teste è:
   - **A.** $\frac{1}{2}$
   - **B.** $\frac{3}{8}$
   - **C.** $\frac{1}{4}$
   - **D.** $\frac{1}{16}$

6. ●● H_1, H_2 formano una partizione, con $P(H_1) = 0{,}3$; inoltre $P(E|H_1) = 0{,}5$ e $P(E|H_2) = 0{,}1$. Allora $P(E)$ vale:

   A. 0,22
   B. 0,6
   C. 0,3
   D. 0,035

7. ● Vero o falso? Correggi le affermazioni false.
   a) Per ogni coppia di eventi con probabilità positiva, $P(A|B) = P(B|A)$.
   b) $P(\overline{A}|B) = 1 - P(A|B)$, se $P(B) > 0$.
   c) Se A e B sono indipendenti, anche $\overline{A}$ e B lo sono.
   d) Due eventi incompatibili con probabilità positiva sono indipendenti.
   e) Le probabilità degli eventi di una partizione hanno somma 1.
   f) Le probabilità a posteriori date dalla formula di Bayes possono avere somma maggiore di 1.

8. ● Completa.
   a) Teorema della probabilità composta: $P(A \cap B) = P(B) \cdot$ _____ .
   b) A e B sono indipendenti se $P(A \cap B) =$ _____ .
   c) Se H_1, H_2, H_3 sono una partizione, $P(E) =$ _____ .
   d) In un test diagnostico, la probabilità che il test sia positivo su una persona malata si chiama _____ .
   e) In n prove ripetute con probabilità di successo p, $P(\text{k successi}) =$ _____ .

Calcolare

9. ● Un'urna contiene 5 palline rosse e 3 nere. Se ne estraggono 2 contemporaneamente. Calcola la probabilità che a) siano entrambe rosse; b) siano di colori diversi; c) siano entrambe nere.

10. ●● Da un mazzo di 40 carte italiane se ne estraggono 4 contemporaneamente. Calcola la probabilità di avere a) 4 carte di denari; b) esattamente 2 assi; c) almeno un asso.

11. ●● Al lotto si estraggono 5 numeri tra 1 e 90 su ogni ruota. Calcola la probabilità di vincere su una ruota giocando a) un solo numero (estratto semplice); b) un ambo; c) una quaterna (4 numeri, che devono uscire tutti).

12. ●● Da un mazzo di 52 carte francesi si distribuiscono 5 carte. Calcola la probabilità a) di avere i 4 assi; b) di non avere nessuna carta di cuori.

13. ●● In un'indagine su 250 persone (dati inventati), 150 si sono vaccinate contro l'influenza e 15 di loro l'hanno presa; delle 100 non vaccinate, 25 l'hanno presa. Costruisci la tabella a doppia entrata e, scelta a caso una persona, calcola: a) $P(I)$, dove I = «ha preso l'influenza»; b) $P(I|V)$ e $P(I|\overline{V})$, dove V = «si è vaccinata»; c) $P(V|I)$. d) Gli eventi I e V sono indipendenti?

14. ●● Si lanciano due dadi. Siano A = «esce almeno un 6» e B = «la somma è almeno 10». Calcola $P(A|B)$ e $P(B|A)$.

15. ●● Da un mazzo di 40 carte italiane si estrae una carta. Stabilisci con il calcolo se sono indipendenti: a) «è un asso» ed «è di coppe»; b) «è una figura» ed «è un re»; c) «è di denari» e «non è di denari».

16. ●● Un'urna contiene 3 palline bianche e 2 nere; se ne estraggono 3, una dopo l'altra, senza reimmissione. Calcola con il teorema della probabilità composta a) la probabilità della sequenza bianca, bianca, nera; b) la probabilità di estrarre esattamente una pallina nera, e controlla il risultato con le combinazioni.

17. ●● Tre macchine producono rispettivamente il 50%, il 30% e il 20% dei pezzi di una fabbrica (dati inventati), con percentuali di pezzi difettosi del 2%, del 3% e del 5%. a) Calcola la probabilità che un pezzo scelto a caso sia difettoso. b) Un pezzo è difettoso: calcola la probabilità che venga da ciascuna macchina e controlla che le tre probabilità sommino a 1.

18. ●● Si lancia un dado 6 volte. Calcola la probabilità di ottenere a) esattamente un 6; b) almeno un 6; c) al più un 6.

Dimostrare e argomentare

19. ●● Dimostra che se A e B sono indipendenti, allora anche $\overline{A}$ e B sono indipendenti. (Suggerimento: B è l'unione degli eventi incompatibili $A \cap B$ e $\overline{A} \cap B$.)

20. ●● Dimostra che, se $P(A) > 0$ e $P(B) > 0$, vale $P(A|B) = \dfrac{P(B|A) \cdot P(A)}{P(B)}$. Deduci che, se $P(A \cap B) > 0$, si ha $P(A|B) = P(B|A)$ se e solo se $P(A) = P(B)$.

21. ●● *Trova l'errore.* «Una famiglia ha due figli e sappiamo che almeno uno è maschio. La probabilità che siano due maschi è $\dfrac{1}{2}$, perché l'altro figlio può essere maschio o femmina.» Supponendo che le quattro sequenze MM, MF, FM, FF siano equiprobabili, trova l'errore e calcola la probabilità corretta.

22. ●● Spiega perché, nella formula di Bernoulli, la somma delle probabilità per $k = 0, 1, \ldots, n$ è uguale a 1, usando il binomio di Newton. Verifica la proprietà scrivendo esplicitamente i quattro termini per $n = 3$ e $p = \dfrac{1}{2}$.

Problemi

23. ●● Per una malattia con prevalenza 0,5% si usa un test con sensibilità 98% e specificità 97% (dati inventati). a) Con le frequenze naturali su 100 000 persone, completa un albero come quello della Figura 10.4. b) Calcola la probabilità che una persona positiva sia malata. c) Calcola la probabilità che una persona negativa sia malata. d) Commenta i due risultati.

24. ●● In una casella di posta (dati inventati) il 30% dei messaggi è spam. La parola «offerta» compare nel 40% dei messaggi di spam e nel 5% degli altri. Un messaggio contiene «offerta»: qual è la probabilità che sia spam?

25. ●● Una giocatrice di pallacanestro segna un tiro libero con probabilità 0,8, indipendentemente dagli altri tiri. Tira 5 tiri liberi. a) Calcola la probabilità che li segni tutti. b) Calcola la probabilità che ne segni esattamente 4. c) Costruisci la tabella delle probabilità per $k = 0, \dots, 5$ e trova il numero più probabile di canestri.

26. ●● L'urna U_1 contiene 3 palline rosse e 2 bianche, l'urna U_2 1 rossa e 4 bianche. Si lancia una moneta: con testa si estrae una pallina da U_1, con croce da U_2. La pallina estratta è rossa. Qual è la probabilità che venga da U_1? Disegna l'albero.

Con il computer

27. ●● Simula con un foglio di calcolo il test diagnostico del testo (prevalenza 2%, sensibilità 90%, specificità 95%). Nella colonna A genera 1000 numeri casuali tra 0 e 1 e scrivi «malato» se il numero è minore di 0,02. Nella colonna B genera un secondo numero casuale e decidi l'esito del test: per un malato «positivo» se il numero è minore di 0,90, per un sano «positivo» se è minore di 0,05. a) Conta i positivi e, tra questi, i malati: quale frazione ottieni? b) Ripeti la simulazione più volte (ricalcolando il foglio) e poi con 10 000 righe: come si comportano i risultati rispetto al valore $\frac{18}{67} \approx 0{,}27$ dato da Bayes?

Verso la verifica

28. ●● In una classe di 25 studenti, 15 ragazze e 10 ragazzi, si sorteggia una commissione di 3 persone. Calcola la probabilità che a) sia formata da sole ragazze; b) contenga almeno un ragazzo; c) contenga esattamente 2 ragazze.

29. ●● Un'azienda (dati inventati) compra il 70% dei componenti dal fornitore X, che ne consegna il 4% difettosi, e il 30% dal fornitore Y, che ne consegna il 10% difettosi. a) Calcola la probabilità che un componente scelto a caso sia difettoso. b) Un componente è difettoso: qual è la probabilità che venga da Y? c) In una scatola di 10 componenti del fornitore X, considerati come prove indipendenti, qual è la probabilità che ce ne sia al più uno difettoso?

30. ●●● *Sfida.* L'urna A contiene 4 palline rosse e 6 bianche, l'urna B 7 rosse e 3 bianche. Si lancia un dado: se esce 1 o 2 si usa l'urna A, altrimenti l'urna B. Dall'urna scelta si estraggono 2 palline contemporaneamente, ed entrambe sono rosse. a) Calcola la probabilità di estrarre due rosse da ciascuna urna. b) Calcola la probabilità totale di estrarre due rosse. c) Qual è la probabilità che sia stata usata l'urna A?

PER APPROFONDIRE

- [HAC] Hacking, *The Emergence of Probability*: come nel Seicento nacque l'idea di probabilità, con le sue due facce, frequenza e grado di fiducia; i capitoli su Pascal e sulla «scommessa».
- [GIG] Gigerenzer, *Calculated Risks*: i test diagnostici, la fallacia del procuratore e il metodo delle frequenze naturali per capire la formula di Bayes senza formule.
- [BOY] Boyer, *Storia della matematica*: le pagine sul calcolo delle probabilità nel Settecento, da Jakob Bernoulli a Laplace.

Glossario del capitolo

Estrazione simultanea
Estrazione di più oggetti insieme (o uno dopo l'altro senza reimmissione), in cui i casi possibili sono combinazioni.

Eventi incompatibili
Eventi che non possono verificarsi insieme: $A \cap B = \varnothing$. Se hanno probabilità positiva, non sono indipendenti.

Eventi indipendenti
Eventi per cui $P(A \cap B) = P(A) \cdot P(B)$; se $P(B) > 0$, equivale a $P(A|B) = P(A)$.

Fallacia del procuratore
L'errore di scambiare $P(A|B)$ con $P(B|A)$, per esempio la probabilità di un indizio per un innocente con la probabilità di innocenza dato l'indizio.

Formula di Bayes
$P(H_i|E) = \dfrac{P(H_i)\,P(E|H_i)}{P(E)}$: dà la probabilità di una causa dopo averne osservato l'effetto.

Frequenze naturali
Modo di rappresentare un problema di probabilità contando i casi in un gruppo immaginario (per esempio 10 000 persone) invece di usare frazioni.

Partizione
Famiglia di eventi a due a due incompatibili, con probabilità positive, la cui unione è lo spazio campionario.

Prevalenza
La frazione di una popolazione che ha una certa malattia, $P(M)$.

Probabilità a posteriori
La probabilità di un'ipotesi dopo l'osservazione E, $P(H_i|E)$, calcolata con la formula di Bayes.

Probabilità a priori
La probabilità di un'ipotesi prima di un'osservazione, $P(H_i)$.

Probabilità assoluta
La probabilità di un evento valutata senza informazioni aggiuntive, per distinguerla da una probabilità condizionata.

Probabilità condizionata
Per $P(B) > 0$, la probabilità di A sapendo che B si è verificato: $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$.

Prove ripetute
Prove indipendenti, ciascuna con due esiti e con la stessa probabilità di successo p; vi si applica la formula di Bernoulli $C_{n,k}\,p^k\,(1-p)^{n-k}$.

Sensibilità
In un test diagnostico, la probabilità che il test sia positivo su una persona malata, $P(+|M)$.

Specificità
In un test diagnostico, la probabilità che il test sia negativo su una persona sana, $P(-|\overline{M})$.

Teorema della probabilità composta
$P(A \cap B) = P(B) \cdot P(A|B) = P(A) \cdot P(B|A)$: la probabilità che due eventi si verifichino entrambi.

Teorema della probabilità totale
Se $H_1, \ldots, H_n$ sono una partizione, $P(E) = P(H_1)\,P(E|H_1) + \ldots + P(H_n)\,P(E|H_n)$.

Valore predittivo positivo
La probabilità di essere malati sapendo che il test è positivo, $P(M|+)$; si calcola con la formula di Bayes.

Verosimiglianza
La probabilità dell'effetto osservato nell'ipotesi H_i, $P(E|H_i)$.

Soluzioni

1. B. È la definizione.

2. C. Le figure sono 12, di cui 4 re: $\frac{4}{12} = \frac{1}{3}$. (La risposta A è la probabilità assoluta $P(\text{re}) = \frac{4}{40}$.)

3. A. $A \cap B = \varnothing$, quindi $P(A|B) = \frac{0}{0{,}4} = 0$. Non sono indipendenti, perché $0 \neq 0{,}3 \cdot 0{,}4 = 0{,}12$.

4. D. $P(A \cap B) = 0{,}5 \cdot 0{,}4 = 0{,}2$.

5. B. Formula di Bernoulli: $C_{4,2} \cdot (\frac{1}{2})^4 = \frac{6}{16} = \frac{3}{8}$.

6. A. $P(H_2) = 0{,}7$;
$P(E) = 0{,}3 \cdot 0{,}5 + 0{,}7 \cdot 0{,}1 = 0{,}15 + 0{,}07 = 0{,}22$.

7. a) F: in generale sono diverse (cambia il denominatore); sono uguali solo se $P(A) = P(B)$ oppure se $P(A \cap B) = 0$. b) V. c) V (si dimostra nella sezione «Dimostrare e argomentare»). d) F: non sono indipendenti, perché $P(A \cap B) = 0 \neq P(A) \cdot P(B)$. e) V. f) F: le probabilità a posteriori hanno sempre somma 1.

8. a) $P(A|B)$. b) $P(A) \cdot P(B)$. c)
$P(H_1) \cdot P(E|H_1) + P(H_2) \cdot P(E|H_2) + P(H_3) \cdot P(E|H_3)$. d) sensibilità. e) $C_{n,k} \cdot p^k \cdot (1-p)^{n-k}$.

9. I casi possibili sono $C_{8,2} = 28$. a)
$\frac{C_{5,2}}{28} = \frac{10}{28} = \frac{5}{14}$. b) $\frac{5 \cdot 3}{28} = \frac{15}{28}$. c)
$\frac{C_{3,2}}{28} = \frac{3}{28}$. Controllo: $\frac{10 + 15 + 3}{28} = 1$.

10. $C_{40,4} = 91\,390$. a)
$\frac{C_{10,4}}{91\,390} = \frac{210}{91\,390} = \frac{21}{9139} \approx 0{,}0023$. b)
$\frac{C_{4,2} \cdot C_{36,2}}{91\,390} = \frac{6 \cdot 630}{91\,390} = \frac{378}{9139} \approx 0{,}0414$. c)
$1 - \frac{C_{36,4}}{91\,390} = 1 - \frac{58\,905}{91\,390} = \frac{6497}{18\,278} \approx 0{,}355$.

11. $C_{90,5} = 43\,949\,268$. a) $\frac{C_{89,4}}{C_{90,5}} = \frac{5}{90} = \frac{1}{18}$. b)
$\frac{C_{88,3}}{C_{90,5}} = \frac{2}{801} \approx 0{,}0025$. c)
$\frac{C_{86,1}}{C_{90,5}} = \frac{86}{43\,949\,268} = \frac{1}{511\,038}$.

12. $C_{52,5} = 2\,598\,960$. a) La quinta carta è una qualunque delle altre 48: $\frac{48}{2\,598\,960} = \frac{1}{54\,145}$. b)
$\frac{C_{39,5}}{C_{52,5}} = \frac{575\,757}{2\,598\,960} = \frac{2109}{9520} \approx 0{,}222$.

13. Tabella: vaccinati 15 con influenza e 135 senza (totale 150); non vaccinati 25 e 75 (totale 100); totali di colonna 40 e 210; totale 250. a) $P(I) = \frac{40}{250} = 0{,}16$.
b) $P(I|V) = \frac{15}{150} = 0{,}1$; $P(I|\overline{V}) = \frac{25}{100} = 0{,}25$. c)
$P(V|I) = \frac{15}{40} = 0{,}375$. d) No: $P(I|V) = 0{,}1 \neq P(I) = 0{,}16$ (equivalentemente $P(I \cap V) = 0{,}06 \neq 0{,}16 \cdot 0{,}6 = 0{,}096$).

14. A ha 11 coppie, $B = \{(4; 6), (5; 5), (6; 4), (5; 6), (6; 5), (6; 6)\}$ ne ha 6, $A \cap B$ ne ha 5 (tutte tranne (5; 5)).
$P(A|B) = \frac{5}{6}$; $P(B|A) = \frac{5}{11}$.

15. a) Sì: $P(\text{asso di coppe}) = \frac{1}{40} = \frac{4}{40} \cdot \frac{10}{40}$. b) No:
$P(\text{figura} \cap \text{re}) = P(\text{re}) = \frac{4}{40} = 0{,}1$, mentre
$\frac{12}{40} \cdot \frac{4}{40} = 0{,}03$. c) No: sono incompatibili con probabilità positive, $P(\cap) = 0 \neq \frac{1}{4} \cdot \frac{3}{4}$.

16. a) $\frac{3}{5} \cdot \frac{2}{4} \cdot \frac{2}{3} = \frac{12}{60} = \frac{1}{5}$. b) Le sequenze con una sola nera sono 3 (BBN, BNB, NBB), ciascuna con probabilità $\frac{1}{5}$ (stessi fattori in ordine diverso): $\frac{3}{5}$. Con le combinazioni: $\frac{C_{3,2} \cdot C_{2,1}}{C_{5,3}} = \frac{3 \cdot 2}{10} = \frac{3}{5}$.

17. a)
$P(D) = 0{,}5 \cdot 0{,}02 + 0{,}3 \cdot 0{,}03 + 0{,}2 \cdot 0{,}05 = 0{,}010 + 0{,}009 + 0{,}010 = 0{,}02$
b) $P(M_1|D) = \frac{0{,}010}{0{,}029} = \frac{10}{29} \approx 0{,}345$;
$P(M_2|D) = \frac{9}{29} \approx 0{,}310$; $P(M_3|D) = \frac{10}{29} \approx 0{,}345$.
Somma: $\frac{29}{29} = 1$.

18. $p = \frac{1}{6}$, $n = 6$. a)
$6 \cdot \frac{1}{6} \cdot (\frac{5}{6})^5 = (\frac{5}{6})^5 = \frac{3125}{7776} \approx 0{,}402$. b)
$1 - (\frac{5}{6})^6 = \frac{31\,031}{46\,656} \approx 0{,}665$. c)
$(\frac{5}{6})^6 + (\frac{5}{6})^5 = \frac{15\,625 + 18\,750}{46\,656} = \frac{34\,375}{46\,656} \approx 0{,}737$.

19. $A \cap B$ e $\overline{A} \cap B$ sono incompatibili e la loro unione è B, quindi $P(\overline{A} \cap B) = P(B) - P(A \cap B)$. Per l'indipendenza $P(A \cap B) = P(A) \cdot P(B)$, perciò
$P(\overline{A} \cap B) = P(B) - P(A) \cdot P(B) = (1 - P(A)) \cdot P(B) = P(\overline{A}) \cdot P(B)$. c.v.d.

20. Per il teorema della probabilità composta $P(A \cap B) = P(B) \cdot P(A|B) = P(A) \cdot P(B|A)$; dividendo per $P(B) > 0$ si ottiene la tesi. Se $P(A \cap B) > 0$, anche $P(B|A) > 0$, e dall'uguaglianza $P(A|B) = P(B|A) \cdot \frac{P(A)}{P(B)}$ si vede che $P(A|B) = P(B|A)$ se e solo se $\frac{P(A)}{P(B)} = 1$, cioè $P(A) = P(B)$. c.v.d.

21. L'informazione «almeno un maschio» esclude solo FF: restano MM, MF, FM, equiprobabili. L'errore è pensare a «l'altro figlio» come a un figlio determinato: l'informazione non dice quale dei due è maschio. $P(\text{MM}|\text{almeno un M}) = \frac{1/4}{3/4} = \frac{1}{3}$.

22. La somma è $\sum C_{n,k}\, p^k\, (1-p)^{n-k}$, che per il binomio di Newton è lo sviluppo di $(p + (1-p))^n = 1^n = 1$.
Per $n = 3$, $p = \frac{1}{2}$: $\frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1$.

23. a) 500 malati: 490 positivi e 10 negativi; 99 500 sani: 2985 positivi (3%) e 96 515 negativi. b)
$P(M|+) = \frac{490}{490 + 2985} = \frac{490}{3475} \approx 0{,}141$. c)
$P(M|-) = \frac{10}{10 + 96\,515} \approx 0{,}0001$. d) Un positivo è malato solo nel 14% circa dei casi, perché la malattia è rara e i falsi positivi sono molti; un negativo invece rassicura quasi completamente.

24.
$P(\text{spam}|\text{offerta}) = \frac{0{,}3 \cdot 0{,}4}{0{,}3 \cdot 0{,}4 + 0{,}7 \cdot 0{,}05} = \frac{0{,}12}{0{,}155} \approx 0{,}774$.

25. a) $0{,}8^5 = 0{,}32768 \approx 0{,}328$. b)
$5 \cdot 0{,}8^4 \cdot 0{,}2 = 0{,}4096 \approx 0{,}410$. c) $P(0) = 0{,}00032$, $P(1) = 0{,}0064$, $P(2) = 0{,}0512$, $P(3) = 0{,}2048$, $P(4) = 0{,}4096$, $P(5) = 0{,}32768$ (somma 1). Il numero più probabile è 4, uguale a $np = 4$.

26. $P(R) = \frac{1}{2} \cdot \frac{3}{5} + \frac{1}{2} \cdot \frac{1}{5} = \frac{2}{5}$;
$P(U_1|R) = \frac{3/10}{2/5} = \frac{3}{4}$. L'albero ha due rami da $\frac{1}{2}$ (le urne) e, da ciascuno, i rami «rossa» e «bianca» con $\frac{3}{5}$, $\frac{2}{5}$ e $\frac{1}{5}$, $\frac{4}{5}$.

27. Con 1000 righe i positivi sono circa 67 e i malati tra i positivi circa 18, ma i numeri cambiano molto da una simulazione all'altra (la frazione può uscire, per esempio, tra 0,15 e 0,40). Con 10 000 righe i positivi sono circa 670 e la frazione si avvicina in genere a 0,27: le frequenze si stabilizzano intorno alla probabilità calcolata con Bayes, tanto più quanto più grande è il campione.

28. $C_{25,3} = 2300$. a) $\frac{C_{15,3}}{2300} = \frac{455}{2300} = \frac{91}{460} \approx 0{,}198$.
b) $1 - \frac{91}{460} = \frac{369}{460} \approx 0{,}802$. c)
$\frac{C_{15,2} \cdot 10}{2300} = \frac{1050}{2300} = \frac{21}{46} \approx 0{,}457$.

29. a)
$P(D) = 0{,}7 \cdot 0{,}04 + 0{,}3 \cdot 0{,}10 = 0{,}028 + 0{,}030 = 0{,}058$. b)
$P(Y|D) = \frac{0{,}030}{0{,}058} = \frac{15}{29} \approx 0{,}517$. c)
$0{,}96^{10} + 10 \cdot 0{,}04 \cdot 0{,}96^9 \approx 0{,}665 + 0{,}277 \approx 0{,}942$.

30. a) $P(RR|A) = \frac{C_{4,2}}{C_{10,2}} = \frac{6}{45} = \frac{2}{15}$;
$P(RR|B) = \frac{C_{7,2}}{C_{10,2}} = \frac{21}{45} = \frac{7}{15}$. b) $P(A) = \frac{1}{3}$, $P(B) = \frac{2}{3}$:
$P(RR) = \frac{1}{3} \cdot \frac{2}{15} + \frac{2}{3} \cdot \frac{7}{15} = \frac{2 + 14}{45} = \frac{16}{45}$. c)
$P(A|RR) = \frac{2/45}{16/45} = \frac{1}{8}$.