6 Numărare şi probabilitate

În acest capitol trecem în revistă combinatorica elementară şi teoria elementară a probabilităţilor. Dacă aveţi o bună fundamentare în aceste domenii, puteţi trece mai uşor peste începutul capitolului şi să vă concentraţi asupra secţiunilor următoare. Multe dintre capitole nu necesită probabilităţi, dar pentru anumite capitole ele sunt esenţiale.

În secţiunea 6.1 trecem în revistă principalele rezultate ale teoriei numărării (combinatorica, analiza combinatorie), inclusiv formule standard pentru numărarea permutărilor şi combinărilor. În secţiunea 6.2 sunt prezentate axiomele probabilităţii şi rezultatele de bază legate de distribuţiile de probabilitate. Variabilele aleatoare sunt introduse în secţiunea 6.3, unde se dau de asemenea proprietăţile mediei şi ale dispersiei. Secţiunea 6.4 investighează distribuţia geometrică şi cea binomială care apar în studiul probelor bernoulliene. Studiul distribuţiei binomiale este continuat în secţiunea 6.5, care este o discuţie avansată despre "cozile" distribuţiei. În fine, în secţiunea 6.6 ilustrăm analiza probabilistică prin intermediul a trei exemple: paradoxul zilei de naştere, aruncarea aleatoare a bilelor în cutii şi câştigarea liniilor (secvenţelor).

6.1. Numărare

Teoria numărării încearcă să răspundă la întrebarea "câţi", fără a număra efectiv. De exemplu, am putea să ne întrebăm "câte numere distincte de n biţi există", sau "în câte moduri pot fi ordonate n elemente distincte". În această secţiune, vom trece în revistă elementele teoriei numărării. Deoarece anumite părţi ale materialului necesită înţelegerea noţiunilor şi rezultatelor de bază referitoare la mulţimi, cititorul este sfătuit să înceapă prin a revedea materialul din secţiunea 5.1.

Regulile sumei şi produsului

O mulţime ale cărei elemente dorim să le numărăm poate fi exprimată uneori ca o reuniune de mulţimi disjuncte sau ca un produs cartezian de mulţimi.

Regula sumei ne spune că numărul de moduri de alegere a unui element din cel puţin una din două mulţimi *disjuncte* este suma cardinalilor mulţimilor. Altfel spus, dacă A şi B sunt două mulţimi finite fără nici un element comun, atunci $|A \cup B| = |A| + |B|$, ceea ce rezultă din ecuaţia (5.3). De exemplu, fiecare poziţie de pe un număr de maşină este o literă sau o cifră. Numărul de posibilităţi, pentru fiecare poziţie, este $26 + 10 = 36$ deoarece sunt 26 de alegeri pentru o literă şi 10 alegeri pentru o cifră.

Regula produsului spune că numărul de moduri în care putem alege o pereche ordonată este numărul de alegeri pentru primul element înmulţit cu numărul de alegeri pentru al doilea element. Adică, dacă A şi B sunt două mulţimi finite, atunci $|A \times B| = |A| \cdot |B|$, care este, pur şi simplu, ecuaţia (5.4). De exemplu, dacă un producător de îngheţată oferă 28 de arome de îngheţată şi 4 tipuri de glazuri, numărul total de sortimente cu un glob de îngheţată şi o glazură este $28 \cdot 4 = 112$.

Şiruri

Un ***şir*** peste o mulţime finită S este o secvenţă de elemente din S. De exemplu, există 8 şiruri binare de lungime 3:

$$000, 001, 010, 011, 100, 101, 110, 111.$$

Vom numi, uneori, ***k-şir*** un şir de lungime k. Un ***subşir*** s' al unui şir s este o secvenţă ordonată de elemente consecutive ale lui s. Un ***k-subşir*** al unui şir este un subşir de lungime k. De exemplu, 010 este un 3 subşir al lui 01101001 (3-subşirul care începe din poziţia 4), dar 111 nu este un subşir al lui 01101001.

Un k-şir peste o mulţime S poate fi privit ca un element al produsului cartezian S^k de k-tuple; astfel există $|S|^k$ şiruri de lungime k. De exemplu, numărul de k-şiruri binare este 2^k. Intuitiv, pentru a construi k-şiruri peste o mulţime de n elemente (n-mulţime), avem n moduri de alegere pentru primul element; pentru fiecare dintre aceste alegeri, avem n moduri de a alege al doilea element; şi aşa mai departe de k ori. Această construcţie ne conduce la concluzia că numărul de k-şiruri este produsul a k factori identici $n \cdot n \cdots n = n^k$.

Permutări

O ***permutare***a unei mulţimi finite S este o secvenţă ordonată a tuturor elementelor lui S, cu fiecare element apărând exact o singură dată. De exemplu, dacă $S = \{a, b, c\}$, există 6 permutări ale lui S:

$$abc, acb, bac, bca, cab, cba.$$

Există $n!$ permutări ale unei mulţimi având n elemente, deoarece primul element al secvenţei poate fi ales în n moduri, al doilea în $n-1$ moduri, al treilea în $n-2$ moduri ş.a.m.d.

O ***k-permutare***[1] a lui S este o secvenţă de k elemente ale lui S în care nici un element nu apare mai mult de o singură dată. (Astfel o permutare ordinară este o n-permutare a unei n-mulţimi.) Cele douăsprezece 2-permutări ale mulţimii $\{a, b, c, d\}$ sunt

$$ab, ac, ad, ba, bc, bd, ca, cb, cd, da, db, dc.$$

Numărul de k-permutări ale unei n-mulţimi este

$$n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}, \tag{6.1}$$

deoarece există n moduri de a alege primul element, $n-1$ moduri de a alege al doilea element şi tot aşa, până se selectează k elemente, ultimul fiind o selecţie de $n-k+1$ elemente.

Combinări

O ***k-combinare*** a unei n-mulţimi S este, pur şi simplu, o k-submulţime a lui S. Există şase 2-combinări ale unei 4-mulţimi $\{a, b, c, d\}$:

$$ab, ac, ad, bc, bd, cd.$$

[1] sau ***aranjament*** – n.t.

(Aici am utilizat prescurtarea ab pentru a nota mulţimea $\{a,b\}$ ş.a.m.d.) Putem construi o k-combinare a unei n-mulţimi alegând k elemente diferite (distincte) din n-mulţime.

Numărul de k-combinări ale unei n-mulţimi poate fi exprimat cu ajutorul numărului de k-permutări ale unei n-mulţimi. Pentru orice k-combinare, există exact $k!$ permutări ale elementelor sale, fiecare fiind o k-permutare distinctă a n-mulţimii. Astfel, numărul de k-combinări ale unei n-mulţimi este numărul de k-permutări împărţit cu $k!$; din ecuaţia (6.1), această valoare este

$$\frac{n!}{k!(n-k)!}. \tag{6.2}$$

Pentru $k=0$, această formulă ne spune că numărul de moduri în care putem alege 0 elemente dintr-o n-mulţime este 1 (nu 0), deoarece $0! = 1$.

Coeficienţi binomiali

Vom utiliza notaţia $\binom{n}{k}$ (se citeşte "combinări de n luate câte k") pentru a desemna numărul de k-combinări ale unei n-mulţimi. Din ecuaţia (6.2) avem

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}. \tag{6.3}$$

Această formulă este simetrică în k şi $n-k$:

$$\binom{n}{k} = \binom{n}{n-k}. \tag{6.4}$$

Aceste numere sunt cunoscute şi sub numele de ***coeficienţi binomiali*** deoarece apar în ***dezvoltarea binomială*** (formula binomului lui Newton):

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}. \tag{6.5}$$

Un caz special al dezvoltării binomiale este acela în care $x = y = 1$:

$$2^n = \sum_{k=0}^{n} \binom{n}{k}. \tag{6.6}$$

Această formulă corespunde numărării celor 2^n n-şiruri binare după numărul de cifre 1 pe care ele îl conţin: există $\binom{n}{k}$ şiruri binare ce conţin exact k de 1 deoarece există $\binom{n}{k}$ moduri de a alege k din cele n poziţii în care vom pune cifre 1.

Există multe identităţi referitoare la coeficienţii binomiali. Exerciţiile de la sfârşitul acestei secţiuni vă oferă posibilitatea să demonstraţi câteva.

Margini binomiale

Uneori, avem nevoie să delimităm mărimea unui coeficient binomial. Pentru $1 \le k \le n$, avem marginea inferioară

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1} = \left(\frac{n}{k}\right)\left(\frac{n-1}{k-1}\right)\cdots\left(\frac{n-k+1}{1}\right) \ge \left(\frac{n}{k}\right)^k. \tag{6.7}$$

Utilizând inegalitatea $k! \geq (k/e)^k$, dedusă din formula lui Stirling (2.12), obţinem marginea superioară

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1} \leq \frac{n^k}{k!} \leq \qquad (6.8)$$

$$\leq \left(\frac{en}{k}\right)^k. \qquad (6.9)$$

Pentru orice $0 \leq k \leq n$, putem utiliza inducţia (vezi exerciţiul 6.1-12) pentru a demonstra delimitarea

$$\binom{n}{k} \leq \frac{n^n}{k^k(n-k)^{n-k}}, \qquad (6.10)$$

unde, pentru comoditate, am presupus că $0^0 = 1$. Pentru $k = \lambda n$, unde $0 \leq \lambda \leq 1$, această margine se poate rescrie sub forma

$$\binom{n}{\lambda n} \leq \frac{n^n}{(\lambda n)^{\lambda n}((1-\lambda)n)^{(1-\lambda)n}} = \left(\left(\frac{1}{\lambda}\right)^{\lambda}\left(\frac{1}{1-\lambda}\right)^{1-\lambda}\right)^n \qquad (6.11)$$

$$= 2^{nH(\lambda)}, \qquad (6.12)$$

unde

$$H(\lambda) = -\lambda \lg \lambda - (1-\lambda)\lg(1-\lambda) \qquad (6.13)$$

este ***funcţia entropie (binară)*** şi unde, pentru comoditate, am presupus că $0 \lg 0 = 0$, astfel că $H(0) = H(1) = 0$.

Exerciţii

6.1-1 Câte k-subşiruri are un n-şir? (Considerăm k-subşirurile identice ce apar pe poziţii diferite ca fiind diferite.) Câte subşiruri are în total un n-şir?

6.1-2 O ***funcţie booleană*** având n intrări şi m ieşiri este o funcţie de la $\{\text{ADEVĂRAT, FALS}\}^n$ la $\{\text{ADEVĂRAT, FALS}\}^m$. Câte funcţii booleene având n intrări şi o ieşire există? Câte funcţii booleene având n intrări şi m ieşiri există?

6.1-3 În câte moduri se pot aşeza n profesori în jurul unei mese circulare la o conferinţă? Considerăm că două aşezări sunt identice dacă una poate fi rotită pentru a o obţine pe cealaltă.

6.1-4 În câte moduri pot fi alese trei numere distincte din mulţimea $\{1, 2, \ldots, 100\}$ astfel ca suma lor să fie pară?

6.1-5 Demonstraţi identitatea:

$$\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1} \qquad (6.14)$$

pentru $0 < k \leq n$.

6.1-6 Demonstraţi identitatea:

$$\binom{n}{k} = \frac{n}{n-k}\binom{n-1}{k}$$

pentru $0 \le k < n$.

6.1-7 Pentru a alege k obiecte din n, puteţi marca unul dintre obiecte şi să vedeţi dacă a fost ales obiectul marcat. Utilizaţi această abordare pentru a demonstra că

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}.$$

6.1-8 Utilizând rezultatul din exerciţiul 6.1-7, construiţi o tabelă a coeficienţilor binomiali, $\binom{n}{k}$ pentru $n = 0, 1, \ldots, 6$ şi $0 \le k \le n$, cu $\binom{0}{0}$ în vârf, $\binom{1}{0}$ şi $\binom{1}{1}$ pe linia următoare şi aşa mai departe. O astfel de tabelă de coeficienţi binomiali se numeşte ***triunghiul lui Pascal***.

6.1-9 Demonstraţi că

$$\sum_{i=1}^{n} i = \binom{n+1}{2}.$$

6.1-10 Arătaţi că, pentru orice $n \ge 0$ şi $0 \le k \le n$, maximul valorii lui $\binom{n}{k}$ se realizează când $k = \lfloor n/2 \rfloor$ sau $k = \lceil n/2 \rceil$.

6.1-11 $\star$ Argumentaţi că, pentru orice $n \ge 0$, $j \ge 0$, $k \ge 0$ şi $j + k \le n$,

$$\binom{n}{j+k} \le \binom{n}{j}\binom{n-j}{k}. \tag{6.15}$$

Daţi atât o demonstraţie algebrică cât şi una bazată pe o metodă de alegere a celor $j+k$ elemente din n. Daţi un exemplu în care egalitatea să nu aibă loc.

6.1-12 $\star$ Utilizaţi inducţia după $k \le n/2$ pentru a demonstra inegalitatea (6.10) şi utilizaţi ecuaţia (6.4) pentru a o extinde pentru orice $k \le n$.

6.1-13 $\star$ Utilizaţi aproximaţia lui Stirling pentru a demonstra că

$$\binom{2n}{n} = \frac{2^{2n}}{\sqrt{\pi n}}\left(1 + O(1/n)\right). \tag{6.16}$$

6.1-14 $\star$ Derivând funcţia entropie $H(\lambda)$, arătaţi că ea îşi atinge valoarea maximă în $\lambda = 1/2$. Cât este $H(1/2)$?

6.2. Probabilitate

Probabilitatea este un instrument esenţial pentru proiectarea şi analiza algoritmilor pur probabilişti şi a algoritmilor care se bazează pe generarea de numere aleatoare. În această secţiune, se trec în revistă bazele teoriei probabilităţilor.

Vom defini probabilitatea în termeni de ***spaţiu de selecţie*** S, care este o mulţime ale cărei elemente se numesc ***evenimente elementare***. Fiecare eveniment elementar poate fi privit ca un rezultat posibil al unui experiment. Pentru experimentul aruncării a două monede distincte şi distinctibile, putem privi spaţiul de selecţie ca fiind mulţimea tuturor 2-şirurilor peste $\{\text{S}, \text{B}\}$:

$$S = \{\text{SS}, \text{SB}, \text{BS}, \text{BB}\}.$$

Un ***eveniment*** este o submulţime[2] a spaţiului de selecţie S. De exemplu, în experimentul aruncării a două monede, evenimentul de a obţine un ban sau o stemă este $\{\text{SB},\text{BS}\}$. Evenimentul S se numeşte ***eveniment sigur***, iar evenimentul $\emptyset$ se numeşte ***eveniment nul*** sau ***eveniment imposibil***. Spunem că două evenimente sunt ***mutual exclusive*** sau ***incompatibile*** dacă $A \cap B = \emptyset$. Vom trata, uneori, un eveniment elementar $s \in S$ ca fiind evenimentul $\{s\}$. Prin definiţie, toate evenimentele elementare sunt mutual exclusive.

Axiomele probabilităţii

O ***distribuţie de probabilitate*** (probabilitate) $\Pr\{\}$ pe spaţiul de selecţie S este o aplicaţie de la mulţimea evenimentelor lui S la mulţimea numerelor reale care satisface următoarele ***axiome ale probabilităţii***:

1. $\Pr\{A\} \geq 0$ pentru orice eveniment A.
2. $\Pr\{S\} = 1$.
3. $\Pr\{A \cup B\} = \Pr\{A\} + \Pr\{B\}$ pentru oricare două evenimente mutual exclusive A şi B. Mai general, pentru orice secvenţă (finită sau infinit numărabilă) de evenimente $A_1, A_2, \ldots$ care sunt două câte două mutual exclusive

$$\Pr\left\{\bigcup_i A_i\right\} = \sum_i \Pr\{A_i\}.$$

Vom numi $\Pr\{A\}$ ***probabilitatea*** evenimentului A. De notat că axioma 2 este o cerinţă de normalizare: nu este nimic fundamental în alegerea lui 1 ca probabilitate a evenimentului sigur, exceptând faptul că este mai naturală şi mai convenabilă.

[2] Pentru o distribuţie de probabilitate generală, pot exista anumite submulţimi ale lui S care nu sunt considerate evenimente. Această situaţie apare de obicei când spaţiul de selecţie este infinit nenumărabil. Cerinţa principală este ca mulţimea evenimentelor din spaţiul de selecţie să fie închisă la operaţiile de complementare şi la intersecţia unui număr finit sau numărabil de evenimente. Cele mai multe dintre distribuţiile pe care le vom întâlni sunt definite peste spaţii de selecţie finite sau numărabile şi vom considera toate submulţimile spaţiului de selecţie ca fiind evenimente. O excepţie notabilă va fi distribuţia de probabilitate uniformă continuă, pe care o vom prezenta pe scurt.

Anumite rezultate se pot deduce imediat din aceste axiome şi din rezultatele de bază ale teoriei mulţimilor (vezi secţiunea 5.1). Evenimentul nul are probabilitatea $\Pr\{\emptyset\} = 0$. Dacă $A \subseteq B$, atunci $\Pr\{A\} \leq \Pr\{B\}$. Utilizând $\bar{A}$ pentru a nota evenimentul $S - A$ (***complementul*** sau ***evenimentul contrar*** lui A), avem $\Pr\{\bar{A}\} = 1 - \Pr\{A\}$. Pentru oricare două evenimente A şi B,

$$\begin{aligned} \Pr\{A \cup B\} &= \Pr\{A\} + \Pr\{B\} - \Pr\{A \cap B\} & (6.17)\\ &\leq \Pr\{A\} + \Pr\{B\}. & (6.18)\end{aligned}$$

În exemplul nostru cu aruncarea monedei presupunem că fiecare eveniment elementar are probabilitatea $1/4$. Probabilitatea de a obţine cel puţin o stemă este

$$\Pr\{\text{SS, SB, BS}\} = \Pr\{\text{SS}\} + \Pr\{\text{SB}\} + \Pr\{\text{BS}\} = 3/4.$$

Altfel, deoarece probabilitatea de a obţine mai puţin de o stemă este $\Pr\{\text{BB}\} = 1/4$, probabilitatea de a obţine cel puţin o stemă este $1 - 1/4 = 3/4$.

Distribuţii discrete de probabilitate

O distribuţie de probabilitate este ***discretă*** dacă este definită peste un spaţiu de selecţie finit sau numărabil. Fie S spaţiul de selecţie. Atunci pentru orice eveniment A,

$$\Pr\{A\} = \sum_{s \in A} \Pr\{s\},$$

deoarece evenimentele elementare, mai concret cele din A, sunt mutual exclusive. Dacă S este finită şi fiecare eveniment elementar $s \in S$ are probabilitatea

$$\Pr\{s\} = 1/|S|,$$

atunci avem ***distribuţia de probabilitate uniformă***[3] pe S. Într-un astfel de caz, experimentul este descris adesea ca "alegerea unui element al lui S la întâmplare".

De exemplu, să considerăm procesul de aruncare a unei ***monede perfecte*** (corecte, ideale), pentru care probabilitatea de a obţine stema este aceeaşi cu probabilitatea de a obţine banul, adică $1/2$. Dacă aruncăm moneda de n ori, avem distribuţia de probabilitate uniformă definită peste spaţiul de selecţie $S = \{\text{S, B}\}^n$, o mulţime de dimensiune (cardinal) 2^n. Fiecare element din S poate fi reprezentat ca un şir de lungime n peste $\{\text{S, B}\}$ şi fiecare apare cu probabilitatea $1/2^n$. Evenimentul

$$A = \{\text{stema apare de exact } k \text{ ori şi banul apare de } n - k \text{ ori}\}$$

este o submulţime a lui S, de dimensiune (cardinal) $|A| = \binom{n}{k}$ deoarece există $\binom{n}{k}$ şiruri de lungime n peste $\{\text{S, B}\}$ ce conţin exact k S-uri. Astfel probabilitatea lui A este $\Pr\{A\} = \binom{n}{k}/2^n$.

[3] distribuţia uniformă discretă pe S – n.t.

Distribuţia de probabilitate uniformă continuă

Distribuţia de probabilitate uniformă continuă este un exemplu de distribuţie de probabilitate în care nu toate submulţimile spaţiului de selecţie sunt considerate evenimente. Distribuţia de probabilitate uniformă continuă este definită pe un interval închis $[a, b]$ al axei reale, cu $a < b$. Intuitiv, dorim ca toate punctele intervalului $[a, b]$ să fie "echiprobabile". Există, totuşi, o infinitate nenumărabilă de puncte, aşa că, dacă atribuim tuturor punctelor aceeaşi probabilitate finită, pozitivă, nu putem satisface axiomele 2 şi 3. Din acest motiv, vom asocia o probabilitate numai anumitor submulţimi ale lui S, astfel ca axiomele să fie satisfăcute pentru astfel de evenimente.

Pentru orice interval închis $[c, d]$, unde $a \leq c \leq d \leq b$, ***distribuţia de probabilitate uniformă continuă*** defineşte probabilitatea evenimentului $[c, d]$ ca fiind

$$\Pr\{[c, d]\} = \frac{d - c}{b - a}.$$

De notat că, pentru orice punct $x = [x, x]$, probabilitatea lui x este 0. Dacă înlăturăm capetele unui interval $[c, d]$, obţinem un interval deschis (c, d). Deoarece $[c, d] = [c, c] \cup (c, d) \cup [d, d]$, pe baza axiomei 3 avem $\Pr\{[c, d]\} = \Pr\{(c, d)\}$. În general, mulţimea evenimentelor pentru distribuţia de probabilitate uniformă continuă este constituită din toate submulţimile lui $[a, b]$ care pot fi obţinute printr-o reuniune finită de intervale deschise sau închise.

Probabilitate condiţionată şi independenţă

Uneori avem, cu anticipaţie, unele cunoştinţe parţiale despre un experiment. De exemplu, să presupunem că un prieten a aruncat două monede perfecte şi că v-a spus că cel puţin una din ele a arătat stema. Care este probabilitatea să obţinem două steme? Informaţia dată elimină posibilitatea de a avea de două ori banul. Cele trei evenimente care au rămas sunt echiprobabile, deci deducem că fiecare din ele apare cu probabilitatea $1/3$. Deoarece numai unul dintre acestea înseamnă două steme, răspunsul la întrebarea noastră este $1/3$.

Probabilitatea condiţionată formalizează noţiunea de a avea cunoştinţe parţiale despre rezultatul unui experiment. ***Probabilitea condiţionată*** a evenimentului A, ştiind că un alt eveniment B a apărut, se defineşte prin

$$\Pr\{A|B\} = \frac{\Pr\{A \cap B\}}{\Pr\{B\}}, \tag{6.19}$$

ori de câte ori $\Pr\{B\} \neq 0$. (Notaţia $\Pr\{A|B\}$ se va citi "probabilitatea lui A condiţionată de B".) Intuitiv, deoarece se dă că evenimentul B a apărut, evenimentul ca A să apară, de asemenea este $A \cap B$. Adică, $A \cap B$ este mulţimea rezultatelor în care apar atât A cât şi B[4]. Deoarece rezultatul este unul din evenimentele elementare ale lui B, normalizăm probabilităţile tuturor evenimentelor elementare din B, împărţindu-le la $\Pr\{B\}$, astfel ca suma lor să fie 1. Probabilitatea lui A condiţionată de B este, de aceea, raportul dintre probabilitatea evenimentului $A \cap B$ şi probabilitatea lui B. În exemplul de mai sus, A este evenimentul ca ambele monede să dea stema, iar B este evenimentul ca să se obţină cel puţin o stemă. Astfel, $\Pr\{A|B\} = (1/4)/(3/4) = 1/3$.

Două evenimente sunt ***independente*** dacă

$$\Pr\{A \cap B\} = \Pr\{A\} \Pr\{B\},$$

[4]Elementele lor sau rezultatele lor – n.t.

ceea ce este echivalent, dacă $\Pr\{B\} \neq 0$, cu condiţia

$$\Pr\{A|B\} = \Pr\{A\}.$$

De exemplu, să presupunem că se aruncă două monede perfecte şi că rezultatele sunt independente. Atunci probabilitatea de a obţine două steme este $(1/2)(1/2) = 1/4$. Presupunem acum că un eveniment este ca prima monedă să dea o stemă şi cel de-al doilea este ca monedele să dea rezultate diferite. Fiecare dintre aceste evenimente apare cu probabilitatea $1/2$, iar probabilitatea ca să apară ambele evenimente este $1/4$; astfel, conform definiţiei independenţei, evenimentele sunt independente – chiar dacă ne-am putea gândi că ambele depind de prima monedă. În fine, să presupunem că monedele sunt legate, astfel încât pe ambele să cadă fie stema, fie banul şi că ambele posibilităţi sunt egal probabile. Atunci probabilitatea ca fiecare monedă să dea stema este $1/2$, dar probabilitatea ca ambele să dea stema este $1/2 \neq (1/2)(1/2)$. În consecinţă, evenimentul ca o monedă să dea stema şi evenimentul ca cealaltă să dea stema nu sunt independente.

O colecţie de evenimente $A_1, A_2, \ldots, A_n$ se numeşte ***independentă pe perechi*** (sau spunem că evenimentele sunt ***două câte două independente*** – n.t.) dacă

$$\Pr\{A_i \cap A_j\} = \Pr\{A_i\} \Pr\{A_j\}$$

pentru orice $1 \leq i < j \leq n$. Spunem că evenimentele $A_1, \ldots, A_n$ sunt (***mutual***) ***independente*** (sau ***independente în totalitate***) dacă fiecare k-submulţime $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$ a colecţiei, unde $2 \leq k \leq n$ şi $1 \leq i_1 < i_2 < \ldots < i_k \leq n$ satisface

$$\Pr\{A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}\} = \Pr\{A_{i_1}\} \Pr\{A_{i_2}\} \ldots \Pr\{A_{i_k}\}.$$

De exemplu, să presupunem că aruncăm două monede perfecte. Fie A_1 evenimentul ca prima monedă să dea stema, fie A_2 evenimentul ca a doua monedă să dea stema şi A_3 evenimentul ca cele două monede să fie diferite. Avem

$$\begin{aligned}
\Pr\{A_1\} &= 1/2, \\
\Pr\{A_2\} &= 1/2, \\
\Pr\{A_3\} &= 1/2, \\
\Pr\{A_1 \cap A_2\} &= 1/4, \\
\Pr\{A_1 \cap A_3\} &= 1/4, \\
\Pr\{A_2 \cap A_3\} &= 1/4, \\
\Pr\{A_1 \cap A_2 \cap A_3\} &= 0.
\end{aligned}$$

Deoarece pentru $1 \leq i < j \leq 3$, avem $\Pr\{A_i \cap A_j\} = \Pr\{A_i\} \Pr\{A_j\} = 1/4$, evenimentele A_1, A_2 şi A_3 sunt două câte două independente. Ele nu sunt mutual independente, deoarece $\Pr\{A_1 \cap A_2 \cap A_3\} = 0$ şi $\Pr\{A_1\} \Pr\{A_2\} \Pr\{A_3\} = 1/8 \neq 0$.

Teorema lui Bayes

Din definiţia probabilităţii condiţionate (6.19) rezultă că, pentru două evenimente A şi B, fiecare cu probabilitate nenulă, avem

$$\Pr\{A \cap B\} = \Pr\{B\} \Pr\{A|B\} = Pr\{A\} \Pr\{B|A\} \tag{6.20}$$

Exprimând din ecuaţiile de mai sus $\Pr\{A|B\}$, obţinem

$$\Pr\{A|B\} = \frac{\Pr\{A\} \Pr\{B|A\}}{\Pr\{B\}}, \tag{6.21}$$

relaţie care este cunoscută sub numele de ***teorema lui Bayes***. Numărătorul $\Pr\{B\}$ este o constantă de normalizare pe care o putem rescrie după cum urmează. Deoarece $B = (B \cap A) \cup (B \cap \bar{A})$ şi $B \cap A$ şi $B \cap \bar{A}$ sunt mutual exclusive,

$$\Pr\{B\} = \Pr\{B \cap A\} + \Pr\{B \cap \bar{A}\} = \Pr\{A\}\Pr\{B|A\} + \Pr\{\bar{A}\}\Pr\{B|\bar{A}\}.$$

Înlocuind în ecuaţia (6.21), obţinem o formă echivalentă a teoremei lui Bayes:

$$\Pr\{A|B\} = \frac{\Pr\{A\}\Pr\{B|A\}}{\Pr\{A\}\Pr\{B|A\} + \Pr\{\bar{A}\}\Pr\{B|\bar{A}\}}.$$

Teorema lui Bayes poate uşura calculul probabilităţilor condiţionate. De exemplu, să presupunem că avem o monedă perfectă şi o monedă falsă care cade întotdeauna pe stemă. Efectuăm un experiment ce constă din trei evenimente independente: una dintre cele două monede este aleasă la întâmplare, aruncată o dată şi apoi aruncată din nou. Să presupunem că moneda aleasă a dat stema la ambele aruncări. Care este probabilitatea ca ea să fie falsă?

Vom rezolva această problemă utilizând teorema lui Bayes. Fie A evenimentul ca să fie aleasă moneda falsă şi fie B evenimentul ca moneda să ne dea de două ori stema. Dorim să determinăm $\Pr\{A|B\}$. Avem $\Pr\{A\} = 1/2$, $\Pr\{B|A\} = 1$, $\Pr\{\bar{A}\} = 1/2$ şi $\Pr\{B|\bar{A}\} = 1/4$; deci

$$\Pr\{A|B\} = \frac{(1/2)\cdot 1}{(1/2)\cdot 1 + (1/2)\cdot(1/4)} = 4/5.$$

Exerciţii

6.2-1 Demonstraţi ***inegalitatea lui Boole***: pentru orice secvenţă finită sau infinit numărabilă de evenimente $A_1, A_2, \ldots$

$$\Pr\{A_1 \cup A_2 \cup \ldots\} \le \Pr\{A_1\} + \Pr\{A_2\} + \cdots. \tag{6.22}$$

6.2-2 Profesorul Rosencrantz aruncă o monedă perfectă. Profesorul Guildenstern aruncă două monede perfecte. Care este probabilitatea ca profesorul Rosencrantz să obţină mai multe steme decât profesorul Guildenstern?

6.2-3 Un pachet de 10 cărţi, fiecare conţinând un număr distinct de la 1 la 10, este amestecat bine. Se extrag trei cărţi din pachet, câte una la un moment dat. Care este probabilitatea ca cele trei cărţi selectate să fie sortate (crescător)?

6.2-4 $\star$ Se dă o monedă falsă, care, atunci când este aruncată, produce stema cu o probabilitate p (necunoscută), unde $0 < p < 1$. Arătaţi cum poate fi simulată "o aruncare de monedă perfectă" examinând aruncări multiple. (*Indicaţie:* aruncaţi moneda de două ori şi apoi fie returnaţi rezultatul, fie repetaţi experimentul.) Demonstraţi că răspunsul este corect.

6.2-5 $\star$ Descrieţi o procedură are acceptă, la intrare, doi întregi a şi b, astfel încât $0 < a < b$ şi, utilizând aruncări de monede perfecte produceţi, la ieşire, stema cu probabilitatea a/b şi banul cu probabilitatea $(b-a)/b$. Daţi o margine a numărului mediu de aruncări de monede, care poate fi un polinom în $\lg b$.

6.2-6 Demonstraţi că

$$\Pr\{A|B\} + \Pr\{\bar{A}|B\} = 1.$$

6.2-7 Arătaţi că, pentru orice colecţie de evenimente $A_1, A_2, \ldots, A_n$,

$$\Pr\{A_1 \cap A_2 \cap \ldots \cap A_n\} = \\ \Pr\{A_1\} \cdot \Pr\{A_2|A_1\} \cdot \Pr\{A_3|A_1 \cap A_2\} \cdots \Pr\{A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}\}.$$

6.2-8 $\star$ Arătaţi cum se poate construi o mulţime de n evenimente care să fie două câte două independente, dar orice submulţime de $k > 2$ dintre ele *să nu fie* mulţime de evenimente mutual independente.

6.2-9 $\star$ Două evenimente A şi B sunt ***independente condiţionat*** dacă

$$\Pr\{A \cap B|C\} = \Pr\{A|C\} \cdot \Pr\{B|C\},$$

unde C este dat. Daţi un exemplu simplu, dar netrivial, de două evenimente care nu sunt independente, dar care sunt independente condiţionat, fiind dat un al treilea eveniment.

6.2-10 $\star$ Sunteţi concurent într-un joc-spectacol în care un premiu este ascuns în spatele uneia dintre trei cortine. Veţi câştiga premiul dacă alegeţi cortina corectă. După ce aţi ales cortina, dar înainte ca ea să fie ridicată, prezentatorul ridică una dintre celelalte cortine, arătând o scenă goală şi vă întreabă dacă doriţi să vă schimbaţi opţiunea curentă pentru cortina rămasă. Care va fi şansa dumneavoastră dacă vă schimbaţi opţiunea?

6.2-11 $\star$ Un director al unei închisori a ales aleator un prizonier din trei pentru a fi eliberat. Ceilalţi doi vor fi executaţi. Gardianul ştie care va fi liber, dar îi este interzis să dea oricărui prizonier informaţii referitoare la starea acestuia. Să numim cei trei prizonieri X, Y şi Z. Prizonierul X întreabă gardianul, în particular, care dintre Y şi Z vor fi executaţi, argumentând că, dacă cel puţin unul dintre cei doi trebuie să moară, gardianul nu-i dă nici o informaţie despre starea sa. Gardianul îi spune lui X că Y va fi executat. Prizonierul X se simte mai fericit acum, deoarece sau el sau Z va fi liber, ceea ce înseamnă că probabilitatea ca el să fie liber este acum $1/2$. Are dreptate, sau şansa sa este tot $1/3$? Explicaţi.

6.3. Variabile aleatoare discrete

O ***variabilă aleatoare*** (***discretă***) X este o funcţie de la un spaţiu de selecţie S, finit sau infinit numărabil, la mulţimea numerelor reale. Ea asociază un număr real fiecărui rezultat al unui experiment, ceea ce ne permite să lucrăm cu distribuţia de probabilitate indusă de mulţimea de numere rezultată. Variabilele aleatoare pot fi definite şi pentru spaţii de selecţie nenumărabile, dar se ajunge la probleme tehnice a căror rezolvare nu este necesară pentru scopurile noastre. De aceea, vom presupune că variabilele aleatoare sunt discrete.

Pentru o variabilă aleatoare X şi un număr real x, definim evenimentul $X = x$ ca fiind $\{s \in S : X(s) = x\}$; astfel

$$\Pr\{X = x\} = \sum_{\{s \in S : X(s) = x\}} \Pr\{s\}.$$

Funcţia

$$f(x) = \Pr\{X = x\}$$

este ***densitatea de probabilitate***[5] a variabilei aleatoare X. Din axiomele probabilităţii, $\Pr\{X = x\} \geq 0$ şi $\sum_x \Pr\{X = x\} = 1$.

De exemplu, să considerăm experimentul aruncării unei perechi de zaruri obişnuite, cu 6 feţe. Există 36 de evenimente elementare posibile în spaţiul de selecţie. Presupunem că distribuţia de probabilitate este uniformă, astfel că fiecare eveniment elementar $s \in S$ are aceeaşi probabilitate: $\Pr\{s\} = 1/36$. Definim variabila aleatoare X ca fiind *maximul* celor două valori de pe feţele zarurilor. Avem $\Pr\{X = 3\} = 5/36$, deoarece X atribuie valoarea 3 la 5 din cele 36 de evenimente elementare posibile şi anume (1,3), (2,3), (3,3), (3,2) şi (3,1).

Se obişnuieşte ca mai multe variabile aleatoare să fie definite pe acelaşi spaţiu de selecţie. Dacă X şi Y sunt variabile aleatoare, funcţia

$$f(x, y) = \Pr\{X = x \text{ şi } Y = y\}$$

este ***funcţia densitate de probabilitate asociată*** lui X şi Y. Pentru o valoare fixată y,

$$\Pr\{Y = y\} = \sum_x \Pr\{X = x \text{ şi } Y = y\}$$

şi la fel, pentru o valoare fixată x

$$\Pr\{X = x\} = \sum_y \Pr\{X = x \text{ şi } Y = y\}.$$

Utilizând definiţia (6.19) a probabilităţii condiţionate, avem

$$\Pr\{X = x | Y = y\} = \frac{\Pr\{X = x \text{ şi } Y = y\}}{\Pr\{Y = y\}}.$$

Spunem că două variabile aleatoare X şi Y sunt ***independente*** dacă, pentru orice x şi y, evenimentele $X = x$ şi $Y = y$ sunt independente, sau echivalent, dacă, pentru orice x şi y, avem $\Pr\{X = x \text{ şi } Y = y\} = \Pr\{X = x\} \Pr\{Y = y\}$.

Dându-se o mulţime de variabile aleatoare peste acelaşi spaţiu de selecţie, se pot defini noi variabile aleatoare, cum ar fi suma, produsul sau alte funcţii de variabilele originale.

Valoarea medie a unei variabile aleatoare

Cea mai simplă şi mai utilă descriere a distribuţiei unei variabile aleatoare este "media" valorilor pe care le ia. ***Valoarea medie*** (sau sinonim ***speranţa*** sau ***media***) a unei variabile aleatoare discrete X este

$$\mathrm{E}[X] = \sum_x x \Pr\{X = x\}, \tag{6.23}$$

care este bine definită dacă suma este finită sau absolut convergentă. Uneori media se notează cu μ_X, sau când variabila aleatoare este subînţeleasă din context cu μ.

Să considerăm un joc în care se aruncă două monede perfecte. Câştigaţi 3\$ pentru fiecare stemă şi pierdeţi 2\$ pentru fiecare ban. Valoarea medie a variabilei aleatoare X ce reprezintă câştigul dumneavoastră este

$$\mathrm{E}[X] = 6 \cdot \Pr\{2\text{S}\} + 1 \cdot \Pr\{1\text{S},1\text{B}\} - 4 \cdot \Pr\{2\text{B}\} = 6(1/4) + 1(1/2) - 4(1/4) = 1.$$

[5]termenul corect din punctul de vedere al teoriei probabilitaţilor este ***funcţie de masă*** sau ***funcţie de frecvenţă*** – n.t.

Media sumei a două variabile aleatoare este suma mediilor, adică

$$\mathrm{E}[X+Y] = \mathrm{E}[X] + \mathrm{E}[Y], \tag{6.24}$$

ori de câte ori $\mathrm{E}[X]$ şi $\mathrm{E}[Y]$ sunt definite. Această proprietate se extinde la sume finite şi absolut convergente de medii.

Dacă X este o variabilă aleatoare, orice funcţie $g(x)$ defineşte o nouă variabilă aleatoare $g(X)$. Dacă valoarea medie a lui $g(X)$ este definită, atunci

$$\mathrm{E}[g(X)] = \sum_x g(x) \Pr\{X = x\}.$$

Luând $g(x) = ax$, avem pentru orice constantă

$$\mathrm{E}[aX] = a\mathrm{E}[X]. \tag{6.25}$$

În consecinţă, mediile sunt liniare: pentru oricare două variabile aleatoare X şi Y şi orice constantă a,

$$\mathrm{E}[aX+Y] = a\mathrm{E}[X] + \mathrm{E}[Y]. \tag{6.26}$$

Când două variabile aleatoare X şi Y sunt independente şi media fiecăreia este definită, avem

$$\begin{aligned}\mathrm{E}[XY] &= \sum_x \sum_y xy \Pr\{X = x \text{ şi } Y = y\} = \sum_x \sum_y xy \Pr\{X = x\} \Pr\{Y = y\} = \\ &= \left(\sum_x x \Pr\{X = x\}\right) \left(\sum_y y \Pr\{Y = y\}\right) = \mathrm{E}[X]\mathrm{E}[Y].\end{aligned}$$

În general, când n variabile aleatoare $X_1, X_2, \ldots, X_n$ sunt mutual independente

$$\mathrm{E}[X_1 X_2 \ldots X_n] = \mathrm{E}[X_1]\mathrm{E}[X_2] \ldots \mathrm{E}[X_n]. \tag{6.27}$$

Atunci când o variabilă aleatoare X ia valori din mulţimea numerelor naturale $\mathbb{N} = \{0, 1, 2, \ldots\}$, există o formulă convenabilă pentru media sa

$$\mathrm{E}[X] = \sum_{i=0}^{\infty} i \Pr\{X = i\} = \sum_{i=0}^{\infty} i \left(\Pr\{X \geq i\} - \Pr\{X \geq i+1\}\right) = \sum_{i=1}^{\infty} \Pr\{X \geq i\}, \tag{6.28}$$

deoarece fiecare termen $\Pr\{X \geq i\}$ este adunat de i ori şi scăzut de $i-1$ ori (exceptând $\Pr\{X \geq 0\}$ care este adunat de 0 ori şi nu este scăzut deloc).

Dispersie şi abatere medie pătratică

Dispersia unei variabile aleatoare X cu media $\mathrm{E}[X]$ este

$$\begin{aligned}\mathrm{Var}[X] &= E\left[(X - \mathrm{E}[X])^2\right] = E\left[X^2 - 2X\mathrm{E}[X] + E^2[X]\right] = \\ &= \mathrm{E}[X^2] - 2\mathrm{E}[X\mathrm{E}[X]] + E^2[X] = \mathrm{E}[X^2] - 2E^2[X] + E^2[X] = \\ &= \mathrm{E}[X^2] - E^2[X].\end{aligned} \tag{6.29}$$

Justificarea pentru egalităţile $\mathrm{E}[E^2[X]] = E^2[X]$ şi $\mathrm{E}[X\mathrm{E}[X]] = E^2[X]$ este aceea că $\mathrm{E}[X]$ nu este o variabilă aleatoare, ci pur şi simplu un număr real, ceea ce înseamnă că se aplică ecuaţia (6.25) (cu $a = \mathrm{E}[X]$). Ecuaţia (6.29) poate fi rescrisă pentru a obţine o expresie a mediei pătratului unei variabile aleatoare:

$$\mathrm{E}[X^2] = \mathrm{Var}[X] + E^2[X]. \tag{6.30}$$

Dispersia unei variabile aleatoare X şi dispersia lui aX sunt legate prin

$$\mathrm{Var}[aX] = a^2\mathrm{Var}[X].$$

Când X şi Y sunt variabile aleatoare independente,

$$\mathrm{Var}[X + Y] = \mathrm{Var}[X] + \mathrm{Var}[Y].$$

În general, dacă n variabile aleatoare $X_1, X_2, \ldots, X_n$ sunt două câte două independente, atunci

$$Var\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathrm{Var}[X_i]. \tag{6.31}$$

Abaterea medie pătratică (***deviaţia standard***) a variabilei aleatoare X este rădăcina pătrată pozitivă a dispersiei lui X. Abaterea medie pătratică a unei variabile aleatoare X se notează uneori cu σ_X sau cu σ, când variabila aleatoare este subînţeleasă din context. Cu această notaţie, dispersia lui X se notează cu σ^2.

Exerciţii

6.3-1 Se aruncă două zaruri obişnuite cu 6 feţe. Care este media sumei celor două valori arătate de zaruri? Care este media maximului celor două valori arătate?

6.3-2 Un tablou $A[1..n]$ conţine n numere distincte ordonate aleator, fiecare permutare a celor n numere având aceeaşi probabilitate de apariţie. Care este media indicelui elementului maxim al tabloului? Care este media indicelui elementului minim al tabloului?

6.3-3 La un joc de carnaval se pun trei zaruri într-o ceaşcă. Un jucător poate paria 1$ pe orice număr de la 1 la 6. Ceaşca este scuturată, zarurile sunt aruncate, iar plata se face în modul următor. Dacă numărul jucătorului nu apare pe nici unul din zaruri, el îşi pierde dolarul. Altfel, dacă numărul său apare pe exact k din cele trei zaruri, $k = 1, 2, 3$, el îşi păstrează dolarul şi mai câştigă k dolari. Care este câştigul mediu când jocul se joacă o singură dată?

6.3-4 $\star$ Fie X şi Y variabile aleatoare independente. Arătaţi că $f(X)$ şi $g(X)$ sunt independente pentru orice alegere a funcţiilor f şi g.

6.3-5 $\star$ Fie X o variabilă aleatoare nenegativă şi presupunem că $\mathrm{E}[X]$ este bine definită. Demonstraţi ***inegalitatea lui Markov***:

$$\Pr\{X \geq t\} \leq \mathrm{E}[X]/t \tag{6.32}$$

pentru orice $t > 0$.

6.3-6 $\star$ Fie S un spaţiu de selecţie şi X şi X' variabile aleatoare astfel încât $X(s) \geq X'(s)$ pentru orice $s \in S$. Demonstraţi că pentru orice constantă reală t,

$$\Pr\{X \geq t\} \geq \Pr\{X' \geq t\}.$$

6.3-7 Ce este mai mare: media pătratului unei variabile aleatoare sau pătratul mediei?

6.3-8 Arătaţi că, pentru orice variabilă aleatoare X ce ia numai valorile 0 şi 1, avem $\mathrm{Var}[X] = \mathrm{E}[X]\mathrm{E}[1 - X]$.

6.3-9 Demonstraţi că $\mathrm{Var}[aX] = a^2\mathrm{Var}[X]$ folosind definiţia (6.29) a dispersiei.

6.4. Distribuţia geometrică şi distribuţia binomială

Aruncarea unei monede este un exemplu de ***probă bernoulliană***, care este definită ca un experiment cu numai două rezultate posibile: ***succes***, care apare cu probabilitatea p şi ***eşec***, care apare cu probabilitatea $q = 1 - p$. Când vorbim de ***probe bernoulliene*** în mod colectiv, înseamnă că probele sunt mutual independente şi, dacă nu se specifică altceva, că fiecare are aceeaşi probabilitate de succes p. Două distribuţii importante derivă din probele bernoulliene: distribuţia geometrică şi cea binomială.

Distribuţia geometrică

Presupunem că avem o secvenţă de probe bernoulliene, fiecare cu probabilitatea de succes p şi probabilitatea de eşec $q = 1 - p$. Câte încercări apar înainte de a obţine un succes? Fie variabila aleatoare X numărul de încercări necesare pentru a obţine un succes. Atunci X ia valori în domeniul $\{1, 2, \ldots\}$ şi pentru $k \geq 1$

$$\Pr\{X = k\} = q^{k-1}p, \tag{6.33}$$

deoarece avem $k - 1$ eşecuri înaintea unui succes. O distribuţie de probabilitate ce satisface ecuaţia (6.33) se numeşte ***distribuţie geometrică***. Figura 6.1 ilustrează o astfel de distribuţie.

Presupunând că $p < 1$, media distribuţiei geometrice poate fi calculată utilizând identitatea (3.6):

$$\mathrm{E}[X] = \sum_{k=1}^{\infty} kq^{k-1}p = \frac{p}{q}\sum_{k=0}^{\infty} kq^k = \frac{p}{q} \cdot \frac{q}{(1-q)^2} = 1/p. \tag{6.34}$$

Astfel, în medie, este nevoie de $1/p$ încercări pentru a obţine un succes, rezultat intuitiv. Dispersia, care se poate calcula în acelaşi mod, este

$$\mathrm{Var}[X] = q/p^2. \tag{6.35}$$

De exemplu, să presupunem că aruncăm, repetat, două zaruri până când obţinem fie suma şapte, fie suma unsprezece. Din cele 36 de rezultate posibile, 6 ne dau suma şapte şi 2 suma unsprezece. Astfel probabilitatea de succes este $p = 8/36 = 2/9$ şi trebuie să aruncăm zarurile, în medie, de $1/p = 9/2 = 4.5$ ori pentru a obţine suma şapte sau unsprezece.

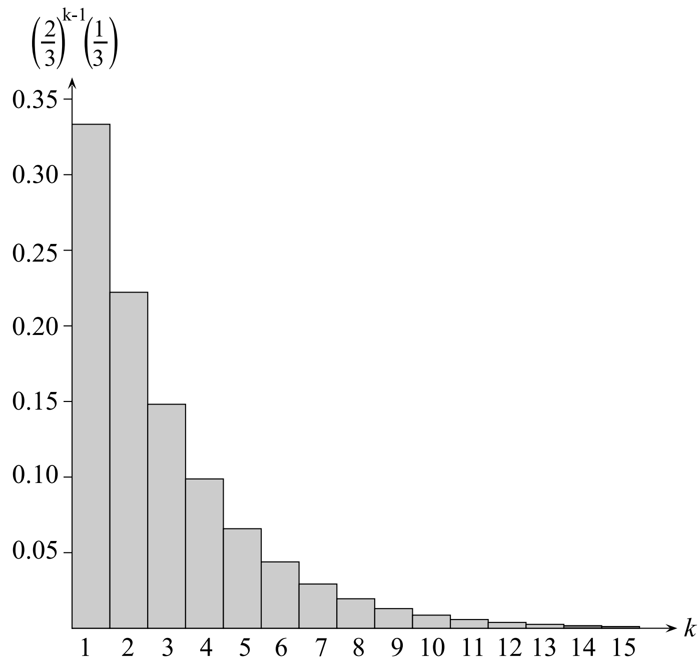

Figura 6.1 O distribuţie geometrică cu probabilitatea de succes $p = 1/3$ şi probabilitatea de eşec $q = 1 - p$. Media distribuţiei este $1/p = 3$.

Distribuţia binomială

Câte succese apar în n probe bernouliene, unde un succes apare cu probabilitatea p şi un eşec cu probabilitatea $q = 1 - p$? Definim variabila aleatoare X ca fiind numărul de succese în n probe. Atunci X ia valori în domeniul $\{0, 1, \ldots, n\}$ şi pentru $k = 0, \ldots, n$,

$$\Pr\{X = k\} = \binom{n}{k} p^k q^{n-k}, \qquad (6.36)$$

deoarece sunt $\binom{n}{k}$ moduri de a alege care k probe din totalul celor n sunt succese şi probabilitatea de apariţie a fiecăreia dintre ele este $p^k q^{n-k}$. O distribuţie de probabilitate ce satisface ecuaţia (6.36) se numeşte ***distribuţie binomială***. Pentru comoditate, definim familia distribuţiilor binomiale utilizând notaţia

$$b(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}. \qquad (6.37)$$

Figura 6.2 ilustrează o distribuţie binomială. Numele de "binomial" vine de la faptul că (6.37) este al k-lea termen al dezvoltării lui $(p + q)^n$. În consecinţă, deoarece $p + q = 1$,

$$\sum_{k=0}^{n} b(k; n, p) = 1, \qquad (6.38)$$

aşa cum cere axioma 2 de probabilitate.

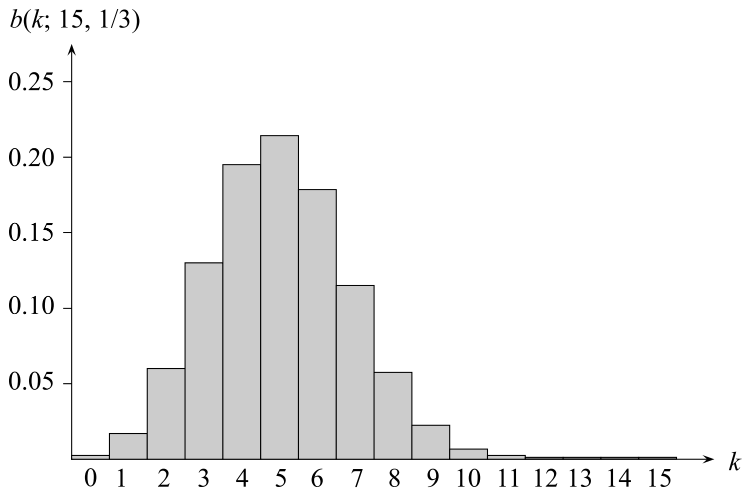

Figura 6.2 Distribuţia binomială $b(k;15,1/3)$ ce rezultă din $n = 15$ probe bernoulliene, fiecare cu probabilitatea de succes $p = 1/3$. Media distribuţiei este $np = 5$.

Putem calcula media unei variabile aleatoare având o distribuţie binomială din ecuaţiile (6.14) şi (6.38). Fie X o variabilă aleatoare ce urmează distribuţia binomială $b(k;n,p)$ şi fie $q = 1 - p$. Din definiţia mediei avem

$$\begin{aligned} \mathrm{E}[X] &= \sum_{k=0}^{n} kb(k;n,p) = \sum_{k=1}^{n} k\binom{n}{k}p^k q^{n-k} = np\sum_{k=1}^{n}\binom{n-1}{k-1}p^{k-1}q^{n-k} = \\ &= np\sum_{k=0}^{n-1}\binom{n-1}{k}p^k q^{(n-1)-k} = np\sum_{k=0}^{n-1} b(k;n-1,p) = np. \end{aligned} \tag{6.39}$$

Utilizând liniaritatea mediei, putem obţine acelaşi rezultat cu substanţial mai puţină algebră. Fie X_i o variabilă aleatoare ce descrie numărul de succese la a i-a probă. Atunci, $\mathrm{E}[X_i] = p\cdot 1 + q\cdot 0 = p$ şi, din liniaritatea mediei (6.26), numărul mediu de succese din n încercări este

$$\mathrm{E}[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n}\mathrm{E}[X_i] = \sum_{i=1}^{n} p = np.$$

Aceeaşi abordare poate fi folosită la calculul dispersiei. Utilizând ecuaţia (6.29), avem $\mathrm{Var}[X_i] = \mathrm{E}[X_i^2] - E^2[X_i]$. Deoarece X_i ia numai valorile 0 şi 1, avem $\mathrm{E}[X_i^2] = \mathrm{E}[X_i] = p$ şi deci

$$\mathrm{Var}[X_i] = p - p^2 = pq. \tag{6.40}$$

Pentru a calcula dispersia lui X, vom exploata avantajul independenţei celor n încercări; astfel din ecuaţia (6.31)

$$\mathrm{Var}[X] = \mathrm{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n}\mathrm{Var}[X_i] = \sum_{i=1}^{n} pq = npq. \tag{6.41}$$

Conform figurii 6.2, distribuţia binomială $b(k;n,p)$ creşte pe măsură ce k variază de la 0 la n, până când atinge media np şi apoi scade. Putem demonstra că distribuţia se comportă întotdeauna în acest mod, examinând raportul termenilor succesivi

$$\begin{aligned}\frac{b(k;n,p)}{b(k-1;n,p)} &= \frac{\binom{n}{k}p^k q^{n-k}}{\binom{n}{k-1}p^{k-1}q^{n-k+1}} = \frac{n!(k-1)!(n-k+1)!p}{k!(n-k)!n!q} = \\ &= \frac{(n-k+1)p}{kq} = 1 + \frac{(n+1)p-k}{kq}. \end{aligned} \tag{6.42}$$

Acest raport este mai mare decât 1 când $(n+1)p-k$ este pozitiv. În consecinţă, $b(k;n,p) > b(k-1;n,p)$ pentru $k < (n+1)p$ (distribuţia creşte) şi $b(k;n,p) < b(k-1;n,p)$ pentru $k > (n+1)p$ (distribuţia descreşte). Dacă $k = (n+1)p$ este întreg, atunci $b(k;n,p) = b(k-1;n,p)$ şi distribuţia are două maxime: în $k = (n+1)p$ şi în $k-1 = (n+1)p-1 = np-q$. Altfel, distribuţia îşi atinge maximul în întregul k unic, situat în domeniul $np-q < k < (n+1)p$.

Lema următoare ne dă o margine superioară a distribuţiei binomiale.

Lema 6.1 Fie $n \geq 0$, $0 < p < 1$, $q = 1-p$ şi $0 \leq k \leq n$. Atunci

$$b(k;n,p) \leq \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k}.$$

Demonstraţie. Utilizând ecuaţia (6.10) avem

$$b(k;n,p) = \binom{n}{k}p^k q^{n-k} \leq \left(\frac{n}{k}\right)^k \left(\frac{n}{n-k}\right)^{n-k} p^k q^{n-k} = \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k}.$$

■

Exerciţii

6.4-1 Verificaţi a doua axiomă a probabilităţii pentru distribuţia geometrică.

6.4-2 De câte ori trebuie aruncate, în medie, 6 monede perfecte înainte de a obţine 3 steme şi 3 bani?

6.4-3 Arătaţi că $b(k;n,p) = b(n-k;n,q)$, unde $q = 1-p$.

6.4-4 Arătaţi că valoarea maximă a distribuţiei binomiale $b(k;n,p)$ este aproximativ $1/\sqrt{2\pi npq}$, unde $q = 1-p$.

6.4-5 $\star$ Arătaţi că probabilitatea de a nu avea nici un succes din n probe bernoulliene, fiecare cu probabilitatea $p = 1/n$, este aproximativ $1/e$. Arătaţi că probabilitatea de a avea exact un succes este, de asemenea, aproximativ $1/e$.

6.4-6 $\star$ Profesorul Rosencrantz aruncă o monedă perfectă de n ori şi la fel face şi profesorul Guildenstern. Arătaţi că probabilitatea ca ei să obţină acelaşi număr de steme este egală cu $\binom{2n}{n}/4^n$. (***Indicaţie:*** pentru profesorul Rosencrantz considerăm stema succes; pentru profesorul Guildenstern considerăm banul succes). Utilizaţi raţionamentul pentru a verifica identitatea

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}.$$

6.4-7 $\star$ Arătaţi că pentru $0 \le k \le n$,

$$b(k; n, 1/2) \le 2^{nH(k/n)-n},$$

unde $H(x)$ este funcţia entropie binară (6.13).

6.4-8 $\star$ Considerăm n probe bernoulliene, unde, pentru $i = 1, 2, \ldots, n$, a i-a probă are probabilitatea de succes p_i şi fie X variabila aleatoare ce desemnează numărul total de succese. Fie $p \ge p_i$ pentru orice $i = 1, 2, \ldots, n$. Demonstraţi că, pentru $1 \le k \le n$,

$$\Pr\{X < k\} \le \sum_{i=0}^{k-1} b(i; n, p).$$

6.4-9 Fie X o variabilă aleatoare ce reprezintă numărul total de succese într-o mulţime A de n probe bernoulliene, unde a i-a probă are probabilitatea de succes p_i, şi fie X' variabila aleatoare ce reprezintă numărul total de succese dintr-o mulţime A' de n probe bernoulliene, unde a i-a probă are o probabilitate de succes $p'_i \ge p_i$. Demonstraţi că, pentru $0 \le k \le n$,

$$\Pr\{X' \ge k\} \ge \Pr\{X \ge k\}.$$

(*Indicaţie:* Arătaţi cum se pot obţine probele bernoulliene din A' printr-un experiment ce implică probele din A şi utilizaţi rezultatul exerciţiului 6.3-6.)

6.5. Cozile distribuţiei binomiale

Probabilitatea de a avea cel puţin sau cel mult k succese în n probe bernoulliene, fiecare cu probabilitatea de succes p, este, adesea, de mai mare interes decât probabilitatea de a avea exact k succese. În această secţiune vom investiga ***cozile*** distribuţiei binomiale, adică cele două regiuni ale distribuţiei $b(k; n, p)$ care sunt îndepărtate de media np.[6] Vom demonstra câteva margini importante (ale sumei tuturor termenilor) dintr-o coadă. Vom demonstra, mai întâi, o margine pentru coada dreaptă a distribuţiei $b(k; n, p)$. Marginile pentru coada stângă pot fi determinate din cele pentru coada dreaptă schimbând între ele rolurile succeselor şi eşecurilor.

Teorema 6.2 Considerăm o secvenţă de n probe bernoulliene, unde succesul apare cu probabilitatea p. Fie X o variabilă aleatoare ce desemnează numărul total de succese. Atunci pentru $0 \le k \le n$, probabilitatea să avem cel puţin k succese este

$$\Pr\{X \ge k\} = \sum_{i=k}^{n} b(i; n, p) \le \binom{n}{k} p^k.$$

Demonstraţie. Vom utiliza inegalitatea (6.15)

$$\binom{n}{k+i} \le \binom{n}{k}\binom{n-k}{i}.$$

[6] Regiuni de forma $\{X \ge k\}$ sau $\{X \le k\}$ – n.t.

Avem

$$\begin{aligned}
\Pr\{X \geq k\} &= \sum_{i=k}^{n} b(i;n,p) = \sum_{i=0}^{n-k} b(k+i;n,p) = \sum_{i=0}^{n-k} \binom{n}{k+i} p^{k+i}(1-p)^{n-(k+i)} \\
&\leq \sum_{i=0}^{n-k} \binom{n}{k}\binom{n-k}{i} p^{k+i}(1-p)^{n-(k+i)} = \binom{n}{k} p^k \sum_{i=0}^{n-k} \binom{n-k}{i} p^i (1-p)^{(n-k)-i} \\
&= \binom{n}{k} p^k \sum_{i=0}^{n-k} b(i;n-k,p) = \binom{n}{k} p^k,
\end{aligned}$$

căci $\sum_{i=0}^{n-k} b(i;n-k,p) = 1$, din ecuaţia (6.38). ■

Corolarul următor reformulează teorema pentru coada stângă a distribuţiei binomiale. În general, vom lăsa în seama cititorului trecerea de la marginea unei cozi la alta.

Corolarul 6.3 Considerăm o secvenţă de n probe bernoulliene, unde succesul apare cu probabilitatea p. Dacă X este o variabilă aleatoare ce reprezintă numărul total de succese, atunci, pentru $0 \leq k \leq n$, probabilitatea să avem cel mult k succese este

$$\Pr\{X \leq k\} = \sum_{i=0}^{k} b(i;n,p) \leq \binom{n}{n-k}(1-p)^{n-k} = \binom{n}{k}(1-p)^{n-k}.$$

■

Marginea următoare se concentrează asupra cozii stângi a distribuţiei binomiale. Cu cât ne îndepărtăm de medie, numărul de succese din coada stângă scade exponenţial, aşa cum ne arată teorema următoare.

Teorema 6.4 Considerăm o secvenţă de n probe bernoulliene, în care succesul apare cu probabilitatea p şi eşecul cu probabilitatea $q = 1-p$. Fie X o variabilă aleatoare ce reprezintă numărul de succese. Atunci, pentru $0 < k < np$, probabilitatea a mai puţin de k succese este

$$\Pr\{X < k\} = \sum_{i=0}^{k-1} b(i;n,p) < \frac{kq}{np-k} b(k;n,p).$$

Demonstraţie. Vom majora seria $\sum_{i=0}^{k-1} b(i;n,p)$ cu o serie geometrică utilizând tehnica din secţiunea 3.2. Pentru $i = 1, 2, \ldots, k$ din ecuaţia (6.42) avem

$$\frac{b(i-1;n,p)}{b(i;n,p)} = \frac{iq}{(n-i+1)p} < \left(\frac{i}{n-i}\right)\left(\frac{q}{p}\right) \leq \left(\frac{k}{n-k}\right)\left(\frac{q}{p}\right).$$

Dacă luăm

$$x = \left(\frac{k}{n-k}\right)\left(\frac{q}{p}\right) < 1,$$

rezultă că

$$b(i-1;n,p) < x b(i;n,p),$$

pentru $0 < i \leq k$. Iterând, obţinem

$$b(i;n,p) < x^{k-i}b(k;n,p)$$

pentru $0 \leq i < k$ şi deci

$$\sum_{i=0}^{k-1} b(i;n,p) \quad < \quad \sum_{i=0}^{k-1} x^{k-i}b(k;n,p) < b(k;n,p)\sum_{i=1}^{\infty} x^i = \frac{x}{1-x}b(k;n,p) = \frac{kq}{np-k}b(k;n,p).$$

■

Când $k \leq np/2$, avem $kq/(np-k) \leq 1$, ceea ce înseamnă că $b(k;n,p)$ majorează suma tuturor termenilor mai mici decât k. De exemplu, să presupunem că aruncăm n monede perfecte. Teorema 6.4 ne spune că, luând $p = 1/2$ şi $k = n/4$, probabilitatea de a obţine mai puţin de $n/4$ steme este mai mică decât probabilitatea de a obţine $n/4$ steme. Mai mult, pentru orice $r \geq 4$, probabilitatea de a obţine mai puţin de n/r steme este mai mică decât probabilitatea de a obţine exact n/r steme. Teorema 6.4 poate fi, de asemenea, utilă în combinaţie cu margini superioare ale distribuţiei binomiale cum ar fi lema 6.1. O margine a cozii drepte se poate determina la fel.

Corolarul 6.5 Considerăm o secvenţă de n probe bernoulliene, în care succesul apare cu probabilitatea p. Fie X o variabilă aleatoare ce reprezintă numărul total de succese. Atunci, pentru $np < k < n$, probabilitatea de a avea mai mult de k succese este

$$\Pr\{X > k\} = \sum_{i=k+1}^{n} b(i;n,p) < \frac{(n-k)p}{k-np}b(k;n,p).$$

■

Teorema următoare consideră n probe bernoulliene, fiecare cu o probabilitate p_i de succes, pentru $i = 1, 2, \ldots, n$. Aşa cum ne arată corolarul următor, putem utiliza teorema pentru a furniza o margine superioară a cozii drepte a distribuţiei binomiale, punând la fiecare încercare $p = p_i$.

Teorema 6.6 Considerăm o secvenţă de n probe bernoulliene, unde, la a i-a probă, pentru $i = 1, 2, \ldots, n$, succesul apare cu probabilitatea p_i, iar eşecul cu probabilitatea $q_i = 1 - p_i$. Fie X o variabilă aleatoare ce reprezintă numărul total de succese şi fie $\mu = E[X]$. Atunci, pentru $r > \mu$,

$$\Pr\{X - \mu \geq r\} \leq \left(\frac{\mu e}{r}\right)^r.$$

Demonstraţie. Deoarece, pentru orice $\alpha > 0$, funcţia $e^{\alpha x}$ este strict crescătoare în x, avem

$$\Pr\{X - \mu \geq r\} = \Pr\{e^{\alpha(X-\mu)} \geq e^{\alpha r}\},$$

unde α va fi determinat mai târziu. Utilizând inegalitatea lui Markov (6.32) obţinem

$$\Pr\{X - \mu \geq r\} \leq E\left[e^{\alpha(X-\mu)}\right]e^{-\alpha r}. \tag{6.43}$$

Partea cea mai consistentă a demonstraţiei o constituie mărginirea lui $E\left[e^{\alpha(X-\mu)}\right]$ şi substituirea lui α în inegalitea (6.43) cu o valoare potrivită. Mai întâi evaluăm $E\left[e^{\alpha(X-\mu)}\right]$. Pentru $i =$

$1, 2, \ldots, n$, fie X_i o variabilă aleatoare care are valoarea 1, dacă a i-a probă bernoulliană este un succes şi 0, dacă este un eşec. Astfel

$$X = \sum_{i=1}^{n} X_i$$

şi

$$X - \mu = \sum_{i=1}^{n} (X_i - p_i).$$

Înlocuind pentru $X - \mu$, obţinem

$$E\left[e^{\alpha(X-\mu)}\right] = E\left[\prod_{i=1}^{n} e^{\alpha(X_i - p_i)}\right] = \prod_{i=1}^{n} E\left[e^{\alpha(X_i-p_i)}\right],$$

care rezultă din (6.27) deoarece independenţa mutuală a variabilelor aleatoare X_i implică independenţa mutuală a variabilelor aleatoare $e^{\alpha(X_i-p_i)}$ (vezi exerciţiul 6.3-4). Din definiţia mediei

$$\begin{aligned} E\left[e^{\alpha(X_i-p_i)}\right] &= e^{\alpha(1-p_i)}p_i + e^{\alpha(0-p_i)}q_i = p_i e^{\alpha q_i} + q_i e^{-\alpha p_i} \leq \\ &\leq p_i e^{\alpha} + 1 \leq \exp(p_i e^{\alpha}), \end{aligned} \tag{6.44}$$

unde $\exp(x)$ reprezintă funcţia exponenţială: $\exp(x) = e^x$. (Inegalitatea (6.44) rezultă din inegalităţile $\alpha > 0$, $q_i \leq 1$, $e^{\alpha q_i} \leq e^{\alpha}$ şi $e^{-\alpha p_i} \leq 1$, iar ultima linie rezultă din inegalitatea (2.7)). În consecinţă,

$$E\left[e^{\alpha(X-\mu)}\right] \leq \prod_{i=1}^{n} \exp(p_i e^{\alpha}) = \exp(\mu e^{\alpha})$$

căci $\mu = \sum_{i=1}^{n} p_i$. Deci, din inegalitatea (6.43), rezultă că

$$\Pr\{X - \mu \geq r\} \leq \exp(\mu e^{\alpha} - \alpha r). \tag{6.45}$$

Alegând $\alpha = \ln(r/\mu)$ (vezi exerciţiul 6.5-6), obţinem

$$\Pr\{X - \mu \geq r\} \leq \exp\left(\mu e^{\ln(r/\mu)} - r\ln(r/\mu)\right) = \exp(r - r\ln(r/\mu)) = \frac{e^r}{(r/\mu)^r} = \left(\frac{\mu e}{r}\right)^r.$$

■

Când se aplică probelor bernoulliene, în care la fiecare încercare avem aceeaşi probabilitate de succes, teorema 6.6 ne dă următorul corolar de mărginire a cozii drepte a unei distribuţii binomiale.

Corolarul 6.7 Considerăm o secvenţă de n probe bernoulliene, unde, la fiecare probă, succesul apare cu probabilitatea p, iar eşecul cu probabilitatea $q = 1 - p$. Atunci, pentru $r > np$,

$$\Pr\{X - np \geq r\} = \sum_{k=\lceil np+r \rceil}^{n} b(k; n, p) \leq \left(\frac{npe}{r}\right)^r.$$

Demonstraţie. Pentru o distribuţie binomială, (6.39) implică $\mu = E[X] = np$. ■

Exerciţii

6.5-1 $\star$ Ce este mai puţin probabil: să nu obţinem nici o stemă când aruncăm o monedă perfectă de n ori sau să obţinem mai puţin de n steme când aruncăm moneda de $4n$ ori?

6.5-2 $\star$ Arătaţi că

$$\sum_{i=0}^{k-1} \binom{n}{i} a^i < (a+1)^n \frac{k}{na - k(a+1)} b(k; n, a/(a+1)),$$

pentru orice $a > 0$ şi orice k, astfel încât $0 < k < n$.

6.5-3 $\star$ Demonstraţi că, dacă $0 < k < np$, unde $0 < p < 1$ şi $q = 1 - p$, atunci

$$\sum_{i=0}^{k-1} p^i q^{n-i} < \frac{kq}{np-k} \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k}.$$

6.5-4 $\star$ Arătaţi că ipotezele teoremei 6.6 implică

$$\Pr\{\mu - X \geq r\} \leq \left(\frac{(n-\mu)e}{r}\right)^r.$$

La fel, arătaţi că ipotezele corolarului 6.7 implică

$$\Pr\{np - X \geq r\} \leq \left(\frac{nqe}{r}\right)^r.$$

6.5-5 $\star$ Considerăm o secvenţă de n probe bernoulliene, unde la a i-a probă, pentru $i = 1, 2, \ldots, n$ succesul apare cu probabilitatea p_i, iar eşecul cu probabilitatea $q_i = 1 - p_i$. Fie X o variabilă aleatoare ce reprezintă numărul total de succese şi fie $\mu = \mathrm{E}[X]$. Arătaţi că, pentru $r \geq 0$,

$$\Pr\{X - \mu \geq r\} \leq e^{-r^2/2n}.$$

(*Indicaţie:* Demonstraţi că $p_i e^{\alpha q_i} + q_i e^{1-\alpha p_i} \leq e^{-\alpha^2/2}$. Apoi urmaţi mersul demonstraţiei teoremei 6.6, utilizând această inegalitate în locul inegalităţii (6.44).)

6.5-6 $\star$ Arătaţi că, alegând $\alpha = \ln(r/\mu)$, se minimizează membrul drept al inegalităţii (6.45).

6.6. Analiză probabilistică

Această secţiune utilizează trei exemple pentru a ilustra analiza probabilistică. Primul determină probabilitatea ca, într-o cameră în care sunt k persoane, să existe o anumită pereche care să-şi sărbătorească ziua de naştere în aceeaşi zi. Al doilea exemplu ilustrează aruncarea aleatoare a bilelor în cutii. Al treilea investighează secvenţele de steme consecutive la aruncarea monedei.

6.6.1. Paradoxul zilei de naştere

Un exemplu bun pentru ilustrarea raţionamentului probabilist este clasicul ***paradox al zilei de naştere***. Câţi oameni trebuie să fie într-o încăpere pentru a avea o bună şansă să găsim doi născuţi în aceeaşi zi a anului? Răspunsul este că trebuie să fie surprinzător de puţini. Paradoxul este acela că, de fapt, trebuie mult mai puţini decât numărul de zile dintr-un an, aşa cum se va vedea. Pentru a răspunde la întrebare, să numerotăm oamenii din cameră cu întregii $1, 2, \ldots, k$, unde k este numărul de oameni din încăpere. Vom ignora anii bisecţi şi vom presupune că toţi anii au $n = 365$ de zile. Pentru $i = 1, 2, \ldots, k$, fie b_i ziua din an în care cade ziua de naştere a lui i, unde $1 \leq b_i \leq n$. Presupunem că zilele de naştere sunt uniform distribuite în cele n zile ale anului, aşa că $\Pr\{b_i = r\} = 1/n$, pentru $i = 1, 2, \ldots, k$ şi $r = 1, 2, \ldots, n$.

Probabilitatea ca două persoane i şi j să aibă aceeaşi zi de naştere depinde de faptul că selecţia aleatoare este dependentă sau independentă. Dacă zilele de naştere sunt independente, atunci probabilitatea ca ziua de naştere a lui i şi ziua de naştere a lui j să cadă amândouă în ziua r este

$$\Pr\{b_i = r \text{ şi } b_j = r\} = \Pr\{b_i = r\} \Pr\{b_j = r\} = 1/n^2.$$

Astfel, probabilitatea ca ambele să cadă în aceeaşi zi este

$$\Pr\{b_i = b_j\} = \sum_{r=1}^{n} \Pr\{b_i = r \text{ şi } b_j = r\} = \sum_{r=1}^{n} \frac{1}{n^2} = 1/n.$$

Intutiv, o dată ales b_i, probabilitatea ca b_j să fie ales acelaşi, este $1/n$. Astfel, probabilitatea ca i şi j să aibă aceeaşi zi de naştere este egală cu probabilitatea ca una dintre zilele de naştere să cadă într-o zi dată. De notat, totuşi, că această coincidenţă depinde de presupunerea că zilele de naştere sunt independente.

Putem analiza probabilitatea ca cel puţin 2 din k oameni să aibă aceeaşi zi de naştere, analizând evenimentul complementar. Probabilitatea ca cel puţin două zile de naştere să coincidă este 1 minus probabilitatea ca toate zilele de naştere să fie diferite. Evenimentul ca toate cele k persoane să aibă zile de naştere distincte este

$$B_k = \bigcap_{i=1}^{k-1} A_i,$$

unde A_i este evenimentul ca ziua de naştere a persoanei $(i + 1)$ să fie diferită de ziua de naştere a persoanei j, pentru orice $j \leq i$, adică,

$$A_i = \{b_{i+1} \neq b_j : j = 1, 2, \ldots, i\}.$$

Deoarece putem scrie $B_k = A_{k-1} \cap B_{k-1}$, din ecuaţia (6.20) obţinem recurenţa

$$\Pr\{B_k\} = \Pr\{B_{k-1}\} \Pr\{A_{k-1} | B_{k-1}\}, \tag{6.46}$$

unde luăm $\Pr\{B_1\} = 1$ drept condiţie iniţială. Cu alte cuvinte, probabilitatea ca b_1, b_2, $\ldots$, b_k să fie zile de naştere distincte este probabilitatea ca b_1, b_2, $\ldots$, b_{k-1} să fie distincte înmulţită cu probabilitatea ca $b_k \neq b_i$, pentru $i = 1, 2, \ldots, k - 1$, ştiind că $b_1, b_2, \ldots, b_{k-1}$ sunt distincte.

Dacă $b_1, b_2, \ldots, b_{k-1}$ sunt distincte, probabilitatea condiţionată ca $b_k \neq b_i$ pentru $i = 1, 2, \ldots, k-1$ este $(n-k+1)/n$, deoarece printre cele n zile există $n-(k-1)$ care nu sunt luate în considerare. Iterând recurenţa (6.46), obţinem

$$\begin{aligned}
\Pr\{B_k\} &= \Pr\{B_1\}\Pr\{A_1|B_1\}\Pr\{A_2|B_2\}\ldots\Pr\{A_{k-1}|B_{k-1}\} = \\
&= 1\left(\frac{n-1}{n}\right)\left(\frac{n-2}{n}\right)\ldots\left(\frac{n-k+1}{n}\right) = \\
&= 1\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{k-1}{n}\right).
\end{aligned}$$

Inegalitatea (2.7), $1+x \leq e^x$ ne dă

$$\Pr\{B_k\} \leq e^{-1/n}e^{-2/n}\ldots e^{-(k-1)n} = e^{-\sum_{i=1}^{k-1} i/n} = e^{-k(k-1)/2n} \leq 1/2$$

când $-k(k-1)/2n \leq \ln(1/2)$. Probabilitatea ca toate cele k zile de naştere să fie distincte este cel mult $1/2$ când $k(k-1) \geq 2n\ln 2$, adică pentru $k \geq (1+\sqrt{1+(8\ln 2)n})/2$, rezultat obţinut rezolvând inecuaţia pătratică. Pentru $n = 365$, trebuie să avem $k \geq 23$. Astfel, dacă în încăpere sunt cel puţin 23 de oameni, probabilitatea ca cel puţin doi să aibă aceeaşi zi de naştere este cel puţin $1/2$. Pe Marte, unde anul are 669 de zile, este nevoie de 31 de marţieni pentru a obţine aceeaşi probabilitate.

O altă metodă de analiză

Putem utiliza liniaritatea mediei (ecuaţia (6.26)) pentru o obţine o analiză mai simplă, dar aproximativă, a paradoxului zilei de naştere. Pentru fiecare pereche (i, j) dintre cele k persoane din încăpere, să definim variabilele aleatoare X_{ij}, pentru $1 \leq i < j \leq k$, prin

$$X_{ij} = \begin{cases} 1 & \text{dacă persoanele } i \text{ şi } j \text{ au aceeaşi zi de naştere} \\ 0 & \text{în caz contrar.} \end{cases}$$

Probabilitatea ca două persoane să aibă aceeaşi zi de naştere este $1/n$ şi astfel, conform definiţiei mediei (6.23),

$$E[X_{ij}] = 1 \cdot (1/n) + 0 \cdot (1-1/n) = 1/n.$$

Numărul mediu de perechi de indivizi având aceeaşi zi de naştere este, conform ecuaţiei (6.24), suma mediilor individuale ale tuturor perechilor, care este

$$\sum_{i=2}^{k}\sum_{j=1}^{i-1} E[X_{ij}] = \binom{k}{2}\frac{1}{n} = \frac{k(k-1)}{2n}.$$

Când $k(k-1) \geq 2n$, numărul mediu de zile de naştere este cel puţin 1. Astfel, dacă avem cel puţin $\sqrt{2n}$ persoane în încăpere, ne putem aştepta ca cel puţin două să aibă aceeaşi zi de naştere. Pentru $n = 365$, dacă $k = 28$, numărul mediu de persoane cu aceeaşi zi de naştere este $(28 \cdot 27)/(2 \cdot 365) \approx 1,0536$. Astfel, dacă avem cel puţin 28 de persoane, ne aşteptăm să găsim cel puţin o pereche având aceeaşi zi de naştere. Pe Marte, unde un an este de 669 de zile, avem nevoie de cel puţin 38 de marţieni.

Prima analiză a determinat numărul de persoane necesar pentru ca probabilitatea existenţei unei coincidenţe a zilelor de naştere să fie mai mare decât $1/2$, iar a doua analiză a determinat numărul de persoane necesar pentru ca numărul mediu de coincidenţe de zile de naştere să fie cel puţin 1. Deşi numărul de oameni determinat în fiecare din cele două situaţii diferă, ambele sunt asimptotic la fel: $\Theta(\sqrt{n})$.

6.6.2. Bile şi cutii

Considerăm procesul de aruncare aleatoare a unor bile identice în b cutii, numerotate cu $1, 2, \ldots, b$. Aruncările sunt independente şi, la fiecare aruncare, bila are aceeaşi probabilitate de a nimeri în oricare dintre cutii. Probabilitatea ca bila să aterizeze în oricare dintre cutiile date este $1/b$. Astfel, procesul de aruncare a bilelor este o secvenţă de probe bernoulliene cu o probabilitate de succes de $1/b$, unde succes înseamnă căderea bilei în cutia dată. Se poate formula o varietate de întrebări interesante referitoare la procesul de aruncare a bilelor.

Câte bile cad într-o cutie dată? Numărul de bile care cad într-o cutie dată urmează distribuţia binomială $b(k; n, 1/b)$. Dacă se aruncă n bile, numărul mediu de bile care cad în cutia dată este n/b.

Cât de multe bile trebuie aruncate, în medie, până când o cutie dată conţine o bilă? Numărul de aruncări, până când cutia dată conţine o bilă, urmează distribuţia geometrică cu probabilitatea $1/b$ şi astfel numărul mediu de aruncări, până se obţine un succes, este $1/(1/b) = b$.

Câte bile trebuie aruncate până când fiecare cutie conţine cel puţin o bilă? Să numim "lovitură" o aruncare în care o bilă cade într-o cutie goală. Dorim să ştim numărul mediu n de aruncări necesare pentru a obţine b lovituri.

Loviturile se pot utiliza pentru a partiţiona cele n aruncări în stadii. Al i-lea stadiu constă din aruncările de după a $(i-1)$-a lovitură până la a i-a lovitură. Primul stadiu constă din prima aruncare, deoarece ni se garantează că avem o lovitură atunci când toate cutiile sunt goale. Pentru fiecare aruncare din timpul celui de-al i-lea stadiu, există $i-1$ cutii ce conţin bile şi $b-i+1$ cutii goale. Astfel, pentru toate aruncările din stadiul i, probabilitatea de a obţine o lovitură este $(b-i+1)/b$.

Fie n_i numărul de aruncări la al i-lea stadiu. Astfel, numărul de aruncări necesare pentru a obţine b lovituri este $n = \sum_{i=1}^{b} n_i$. Fiecare variabilă aleatoare n_i are o distribuţie geometrică cu probabilitatea de succes $(b-i+1)/b$ şi de aceea

$$E[n_i] = \frac{b}{b-i+1}.$$

Din liniaritatea mediei, obţinem

$$E[n] = E\left[\sum_{i=1}^{b} n_i\right] = \sum_{i=1}^{b} E[n_i] = \sum_{i=1}^{b} \frac{b}{b-i+1} = b\sum_{i=1}^{b} \frac{1}{i} = b(\ln b + O(1)).$$

Ultima inegalitate rezultă din marginea (3.5) referitoare la seria armonică. Deci sunt necesare aproximativ $b \ln b$ aruncări înainte de a ne putea aştepta ca fiecare cutie să conţină o bilă.

6.6.3. Linii (secvenţe)

Să presupunem că aruncăm o monedă perfectă de n ori. Care este cea mai lungă linie (secvenţă) de steme consecutive pe care ne aşteptăm s-o obţinem? Răspunsul este $\Theta(\lg n)$, aşa cum ne arată analiza următoare.

Vom demonstra, întâi, că lungimea celei mai lungi secvenţe de steme este $O(\lg n)$. Fie A_{ik} evenimentul ca o secvenţă de steme de lungime cel puţin k să înceapă de la a i-a aruncare sau, mai precis, evenimentul ca cele k aruncări consecutive, $i, i+1, \ldots, i+k-1$ să ne dea numai steme, unde $1 \le k \le n$ şi $1 \le i \le n-k+1$. Pentru un eveniment dat A_{ik}, probabilitatea ca toate cele k aruncări să aibă ca rezultat stema, are o distribuţie geometrică cu $p = q = 1/2$:

$$\Pr\{A_{ik}\} = 1/2^k. \tag{6.47}$$

Pentru $k = 2\lceil \lg n \rceil$,

$$\Pr\left\{A_{i,2\lceil \lg n\rceil}\right\} = 1/2^{2\lceil \lg n\rceil} \le 1/2^{2\lg n} = 1/n^2,$$

şi, astfel, probabilitatea ca o secvenţă de steme de lungime cel puţin $\lceil 2\lg n\rceil$ să înceapă din poziţia i este chiar mică, considerând că sunt cel mult n poziţii (de fapt $n - 2\lceil \lg n\rceil + 1$) în care secvenţa ar putea începe. Probabilitatea ca o secvenţă de steme de lungime cel puţin $\lceil 2\lg n\rceil$ să înceapă oriunde este deci

$$\Pr\left\{\bigcup_{i=1}^{n-2\lceil \lg n\rceil+1} A_{i,2\lceil \lg n\rceil}\right\} \le \sum_{i=1}^{n} 1/n^2 = \frac{1}{n},$$

deoarece, conform inegalităţii lui Boole (6.22), probabilitatea unei reuniuni de evenimente este cel mult suma probabilită ţilor evenimentelor individuale. (De notat că inegalitatea lui Boole are loc şi pentru evenimente care nu sunt independente.)

Probabilitatea ca orice secvenţă să aibă lungimea cel puţin $2\lceil \lg n\rceil$ este cel mult $1/n$; deci probabilitatea ca cea mai lungă secvenţă să aibă lungimea mai mică decât $2\lceil \lg n\rceil$ este cel puţin $1 - 1/n$. Deoarece fiecare secvenţă are lungimea cel mult n, lungimea medie a celei mai lungi secvenţe este mărginită superior de

$$(\lceil 2\lg n\rceil)(1 - 1/n) + n(1/n) = O(\lg n).$$

Şansa ca o secvenţă de steme să depăşească $r\lceil \lg n\rceil$ se diminuează rapid o dată cu r. Pentru $r \ge 1$, probabilitatea ca o secvenţă de $r\lceil \lg n\rceil$ să înceapă în poziţia i este

$$\Pr\{A_{i,r\lceil \lg n\rceil}\} = 1/2^{r\lceil \lg n\rceil} \le 1/n^r.$$

Astfel, probabilitatea ca cea mai lungă secvenţă să aibă lungimea cel puţin $r\lceil \lg n\rceil$ este cel mult $n/n^r = 1/n^{r-1}$, sau echivalent, probabilitatea ca cea mai lungă secvenţă să aibă lungimea mai mică decât $r\lceil \lg n\rceil$ este cel puţin $1 - 1/n^{r-1}$.

De exemplu, pentru $n = 1000$ de aruncări de monedă, probabilitatea de a avea o secvenţă de cel puţin $2\lceil \lg n\rceil = 20$ de steme este de cel mult $1/n = 1/1000$. Şansa de a avea o secvenţă mai lungă de $3\lceil \lg n\rceil = 30$ de steme este de cel mult $1/n^2 = 1/1000000$.

Vom demonstra acum o margine inferioară complementară: lungimea celei mai lungi secvenţe de steme în n aruncări este $\Omega(\lg n)$. Pentru a demonstra această margine vom căuta o secvenţă de lungime $\lfloor \lg n\rfloor/2$. Din ecuaţia (6.47) avem

$$\Pr\{A_{i,\lfloor \lg n\rfloor/2}\} = 1/2^{\lfloor \lg n\rfloor/2} \ge 1/\sqrt{n}.$$

Probabilitatea ca o secvenţă de lungime cel puţin $\lfloor \lg n\rfloor/2$ să nu înceapă în poziţia i este deci cel mult $1 - 1/\sqrt{n}$. Putem partiţiona cele n aruncări ale monedei în cel puţin $\lfloor 2n/\lfloor \lg n\rfloor\rfloor$ grupuri de $\lfloor \lg n \rfloor/2$ aruncări consecutive de monede. Deoarece aceste grupuri sunt formate din aruncări mutual exclusive şi independente, probabilitatea ca nici unul dintre aceste grupuri să nu fie o secvenţă de lungime $\lfloor \lg n \rfloor/2$ este

$$(1 - 1/\sqrt{n})^{\lfloor 2n/\lfloor \lg n \rfloor \rfloor} \leq (1 - 1/\sqrt{n})^{2n/\lg n - 1} \leq e^{-(2n/\lg n - 1)/\sqrt{n}} = O(e^{-\lg n}) = O(1/n).$$

Pentru a justifica aceasta, am utilizat inegalitatea (2.7), $1 + x \leq e^x$. Astfel, probabilitatea ca cea mai lungă linie să depăşească $\lfloor \lg n \rfloor/2$ este cel puţin $1 - O(1/n)$. Deoarece cea mai lungă secvenţă are lungimea cel puţin 0, lungimea medie a celei mai lungi linii (secvenţe) este de cel puţin

$$(\lfloor \lg n \rfloor/2)(1 - O(1/n)) + 0 \cdot (1/n) = \Omega(\lg n).$$

Exerciţii

6.6-1 Presupunem că mai multe bile sunt aruncate în b cutii. Fiecare aruncare este independentă şi fiecare bilă are aceeaşi probabilitate de a ajunge în oricare dintre cutii. Care este numărul mediu de bile aruncate, înainte ca cel puţin una din cutii să conţină două bile?

6.6-2 $\star$ Pentru analiza paradoxului zilei de naştere, este important ca zilele de naştere să fie independente în totalitate (mutual independente) sau este suficient să fie independente două câte două? Justificaţi răspunsul.

6.6-3 $\star$ Câţi oameni trebuie invitaţi la o petrecere pentru a face verosimilă existenţa a ***trei*** dintre ei cu aceeaşi zi de naştere?

6.6-4 $\star$ Care este probabilitatea ca un k-şir peste o mulţime de lungime n să fie o k-permutare? Ce legătură este între această chestiune şi paradoxul zilei de naştere?

6.6-5 $\star$ Presupunem că n bile sunt aruncate în n cutii, fiecare aruncare este independentă şi fiecare bilă are aceeaşi probabilitate de a nimeri în oricare dintre cutii. Care este numărul mediu de cutii goale? Care este numărul mediu de cutii cu exact o bilă?

6.6-6 $\star$ Îmbunătăţiţi marginea inferioară a lungimii unei secvenţe arătând că, pentru n aruncări ale unei monede perfecte, probabilitatea ca să nu apară nici o linie de mai mult de $\lg n - 2 \lg \lg n$ steme consecutive este mai mică decât $1/n$.

Probleme

6-1 Bile şi cutii
În această problemă investigăm efectul diverselor ipoteze asupra numărului de moduri de a plasa n bile în b cutii distincte.

a. Să presupunem că cele n bile sunt distincte şi ordinea lor într-o cutie nu contează. Argumentaţi că numărul de moduri în care se pot aşeza bilele în cutii este b^n.

b. Să presupunem că bilele sunt distincte şi că bilele din fiecare cutie sunt ordonate. Demonstraţi că numărul de moduri în care se pot aşeza bilele în cutii este $(b+n-1)!/(b-1)!$. (*Indicaţie:* Consideraţi numărul de moduri de aranjare a n bile distincte şi $b-1$ beţişoare nedistinctibile într-un rând.)

c. Presupunem că bilele sunt identice şi că ordinea lor în cutie nu contează. Arătaţi că numărul de moduri în care se pot aşeza bilele în cutii este $\binom{b+n-1}{n}$. (*Indicaţie:* câte dintre aranjamentele din partea (b) se repetă dacă bilele devin identice?)

d. Presupunem că bilele sunt identice şi că nici o cutie nu conţine mai mult de o bilă. Arătaţi că numărul de moduri în care se pot aşeza bilele este $\binom{b}{n}$.

e. Presupunem că bilele sunt identice şi că nici o cutie nu poate rămâne goală. Arătaţi că numărul de moduri în care se pot aşeza bilele este $\binom{n-1}{b-1}$.

6-2 Analiza programului de determinare a maximului

Programul următor determină valoarea maximă dintr-un tablou neordonat $A[1..n]$.

1: $max \leftarrow -\infty$
2: **pentru** $i \leftarrow 1, n$ **execută**
3: $\quad \triangleright$ Compară $A[i]$ cu max
4: $\quad$ **dacă** $A[i] > max$ **atunci**
5: $\qquad max \leftarrow A[i]$

Dorim să determinăm numărul mediu de execuţii ale atribuirii din linia 5. Presupunem că numerele din A sunt o permutare aleatoare a n numere distincte.

a. Dacă un număr x este ales aleator dintr-o mulţime de i numere distincte, care este probabilitatea ca x să fie cel mai mare din mulţime?

b. Care este relaţia dintre $A[i]$ şi $A[j]$ când se execută linia 5 a programului, pentru $1 \le j \le i$?

c. Pentru orice i din domeniul $1 \le i \le n$, care este probabilitatea ca linia 5 să fie executată?

d. Fie $s_1, s_2, \ldots, s_n$ n variabile aleatoare, unde s_i reprezintă numărul de execuţii (0 sau 1) ale liniei 5 în timpul celei de-a i-a iteraţii a ciclului **pentru**. Ce este $\mathrm{E}[s_i]$?

e. Fie $s = s_1 + s_2 + \cdots + s_n$ numărul total de execuţii ale liniei 5 în timpul unei execuţii a programului. Arătaţi că $\mathrm{E}[s] = \Theta(\lg n)$.

6-3 Problema angajării

D-na profesoară Dixon vrea să angajeze un nou asistent cercetător. Ea a aranjat interviuri cu n solicitanţi şi vrea să-şi bazeze decizia doar pe calificarea lor. Din nefericire, regulile universităţii cer ca după fiecare interviu candidatul să fie respins sau să i se ofere postul solicitat. Profesoara Dixon a decis să adopte strategia de a selecta un întreg pozitiv $k < n$, să intervieveze şi să respingă primii k solicitanţi şi apoi să-l angajeze pe solicitantul care este mai pregătit decât toţi solicitanţii precedenţi. Dacă cel mai bun este între primii k intervievaţi, va fi angajat al n-lea solicitant. Arătaţi că d-na profesoară Dixon îşi maximizează şansa de a angaja cel mai bine pregătit solicitant, alegând k aproximativ egal cu n/e şi că şansa de a-l angaja pe cel mai bine pregătit este atunci aproximativ $1/e$.

6-4 Numărare probabilistică
Cu un contor de t biţi putem număra în mod obişnuit până la $2^t - 1$. Cu ***numărarea probabilistică*** a lui R. Morris putem număra până la valori mult mai mari, cu preţul unei pierderi a preciziei. Vom lua o valoare de contor i care va reprezenta numărul de n_i, pentru $i = 0, 1, \ldots, 2^t - 1$, unde n_i formează o secvenţă crescătoare de valori nenegative. Vom presupune că valoarea iniţială a contorului este 0, reprezentând numărul de valori $n_0 = 0$. Operaţia INCREMENTEAZĂ lucrează pe un contor ce conţine valoarea i în mod probabilist. Dacă $i = 2^t - 1$, se raportează o eroare de tip depăşire superioară. Altfel, contorul este incrementat cu 1 cu probabilitatea $1/(n_{i+1} - n_i)$ şi rămâne neschimbat cu probabilitatea $1 - 1/(n_{i+1} - n_i)$. Dacă selectăm $n_i = i$, pentru orice $i \geq 0$, atunci contorul este unul obişnuit. O situaţie mai interesantă apare dacă selectăm, de exemplu, $n_i = 2^{i-1}$, pentru $i > 0$, sau $n_i = F_i$ (al i-lea număr Fibonacci – vezi secţiunea 2.2). Pentru această problemă să presupunem că n_{2^t-1} este suficient de mare, astfel ca probabilitatea unei depăşiri superioare să fie neglijabilă.

a. Arătaţi că valoarea medie reprezentată de contor, după ce au fost efectuate n operaţii INCREMENTEAZĂ, este exact n.

b. Analiza dispersiei numărului reprezentat de contor depinde de secvenţa de numere n_i. Să considerăm un caz simplu: $n_i = 100i$, pentru orice $i \geq 0$. Estimaţi dispersia valorii reprezentate de registru după ce au fost efectuate n operaţii INCREMENTEAZĂ.

Note bibliografice

Primele metode generale de rezolvare a problemelor de probabilităţi au fost discutate în celebra corespondenţă dintre B. Pascal şi P. de Fermat, care a început în 1654 şi într-o carte a lui C. Huygens din 1657. Teoria riguroasă a probabilităţilor a început cu lucrarea lui J. Bernoulli din 1713 şi cu cea a lui A. de Moivre din 1730. Dezvoltări ulterioare ale teoriei au fost date de P. S. de Laplace, S.-D. Poisson şi C. F. Gauss.

Sumele de variabile aleatoare au fost studiate pentru prima dată de P. L. Chebyshev (Cebîşev) şi de A. A. Markov. Teoria probabilităţilor a fost axiomatizată de A. N. Kolmogorov în 1933. Marginile pentru cozile distribuţiilor au fost date de Chernoff [40] şi Hoeffding [99]. Lucrarea de bază referitoare la structurile aleatoare combinatorice a fost dată de P. Erdös.

Knuth [121] şi Liu [140] sunt referinţe bune pentru combinatorică elementară şi numărare. Cărţi standard ca Billingsley [28], Chung [41], Drake [57], Feller [66] şi Rozanov [171] dau o introducere comprehensivă în teoria probabilităţilor. Bollobás [30], Hofri [100] şi Spencer [179] conţin o mare varietate de tehnici probabilistice avansate.