Lei 2,45

• XII • Matematică

GH. MIHOC
N. MICU

$$\begin{pmatrix} |x_1-m| & |x_2-m| & \cdots & |x_n-m| \\ p_1 & p_2 & \cdots & p_n \end{pmatrix}$$

XII

Matematică

manual pentru clasa a XII-a

Elemente de teoria probabilităţilor şi statistică matematică

Editura didactică și pedagogică
București, 1982

Gh. Mihoc
N. Micu

XII Matematică

manual pentru clasa a XII-a

Elemente de teoria probabilităţilor şi statistică matematică

Editura didactică şi pedagogică
Bucureşti

Manualul a fost revizuit pe baza programei școlare aprobate de Ministerul Educației și Învățămîntului cu nr. 39490/1978.

Referenți: dr. docent *M. Iosifescu*
cercetător *R. Gologanu*

Redactor: prof. *Ioan Șt. Mușat*
Tehnoredactor: *Ana Țimpău*
Grafician copertă: *N Sîrbu*

Scurt istoric

Bazele teoriei probabilităților au fost puse în secolul XVII de matematicienii B. Pascal (1623—1662) și P. Fermat (1601—1665). Un pasionat jucător de zaruri, cavalerul de Méré, susținea în discuțiile sale cu Pascal că jocurile de noroc uneori conduc la rezultate care contrazic matematica. Astfel, afirma el, a arunca un zar de 4 ori pentru a obține o dată fața șase, este același lucru cu a arunca de 24 ori cîte două zaruri pentru a obține o dublă de șase. Dacă aruncăm un zar avem 6 rezultate posibile (fețele: 1, 2, ..., 6) și facem 4 încercări. Avem raportul $\frac{4}{6}=\frac{2}{3}$. Dacă aruncăm două zaruri avem 36 cazuri posibile (perechile de fețe: (1, 1), (1, 2), ..., (6, 6) și 24 de încercări. Deci același raport $\frac{24}{36}=\frac{2}{3}$. Cu toate acestea, cavalerul de Méré a observat că jucînd în modul al doilea (cu două zaruri aruncate de 24 ori), pierde față de adversarul său, dacă acesta alege primul mod (aruncarea unui singur zar de 6 ori), ceea ce credea el, contrazice regulile matematice. Pascal și Fermat au arătat însă că probabilitatea de cîștig la jocul cu un singur zar este 0,518, iar la jocul cu două zaruri 0,492. Deși diferența dintre cele două probabilități este mică, totuși, la un număr mare de partide, jucătorul cu probabilitatea de cîștig 0,518 cîștigă în fața jucătorului cu probabilitatea de cîștig 0,492. Deci practica jocului confirmă justețea raționamentului matematic, contrar credinței lui de Méré.

O altă problemă, devenită de asemenea celebră prin faptul că a condus la nașterea unei noi discipline matematice, a constat în împărțirea mizei la un joc, care este întrerupt înainte de a fi desemnat un cîștigător. La un joc, la care participă doi parteneri în condiții egale, este învingător cel care cîștigă trei partide. După trei partide jucate, jocul se întrerupe, primul jucător avînd două partide cîștigate, iar al doilea numai una. Cum trebuie să se împartă miza? Cavalerul de Méré susținea că trebuie să se împartă proporțional cu numărul partidelor cîștigate de fiecare jucător, adică cu numerele 2 și 1. Pascal, Fermat și C. Huygens (1629—1695), care a contribuit și el al apariția teoriei probabilităților, au demonstrat pe căi diferite că miza trebuie împărțită proporțional cu numerele 3 și 1. Pare curios că niște savanți, cărora știința le datorează descoperiri fundamentale, se ocupau de rezolvarea unor probleme neînsemnate puse de practica jocurilor de noroc, dar ei erau convinși de importanța descoperirii lor din punct de vedere filozofic. În scrisoarea în limba latină adresată Academiei de Științe a Franței prin care Pascal anunța rezultatul cercetărilor sale, el arată că a reușit să concilieze incertitudinile hazardului cu demonstrațiile matematice.

Mai tîrziu în opera postumă „Ars conjectandi" (1713) a unui alt mare matematician J. Bernoulli (1654—1705) se stabilește pentru prima oară că noua teorie matematică este fundamentală pentru studiul fenomenelor de masă. Printr-o teoremă celebră, intitulată de el „teorema numerelor mari", J. Bernoulli stabilește relația matematică dintre frecvență și probabilitate după un număr mare de probe. Această teoremă constituie fundamentul statisticii matematice și justifică aplicarea teoriei probabilităților în alte domenii. N. Bernoulli (1687—1759), editorul operei „Ars conjectandi", a aplicat cu succes teoria probabilităţilor în ştiinţele moral politice şi în demografie, iar D. Bernoulli (1700—1782) a fost primul care a aplicat-o la studiul teoriei cinetice a gazelor şi a studiat probleme premergătoare teoriei deciziei de astăzi. N. Bernoulli şi D. Bernoulli au fost nepoţii lui J. Bernoulli.

Un alt matematician, care a adus contribuţii importante în teoria probabilităţilor, a fost A. de Moivre (1667—1754). El a găsit normală legea de probabilităţi, atribuită mai tîrziu pe nedrept altor oameni de ştiinţă.

Dar cel care pe drept cuvînt trebuie să fie considerat ca fondator al teoriei moderne a probabilităţilor este P. S. Laplace (1749—1827). În tratatul său „Teoria analitică a probabilităţilor" (1813), el expune în mod riguros propoziţiile de bază ale teoriei probabilităţilor, enunţă şi rezolvă în anumite cazuri teorema limită centrală, fundamentală în teoria erorilor, şi aplică în mod ştiinţific calculul probabilităţilor în demografie, astronomie şi în alte domenii.

Printre marii matematicieni, care au adus contribuţii în teoria probabilităţilor în secolul XIX, cităm pe K. F. Gauss (1777—1855), J. Bertrand (1822—1900), H. Poincaré (1854—1912). Trebuie semnalat, de asemenea, şi aportul şcolii ruse de probabilităţi, întemeiată de P. S. Cebîşev (1821—1894), avînd ca reprezentanţi străluciţi pe A. M. Liapunov (1857—1918) şi A. Markov (1856—1922), autorul unor procese stochastice de mare importanţă în ştiinţa de astăzi.

În secolul nostru s-a realizat axiomatizarea teoriei probabilităţilor. Au adus contribuţii în această direcţie, în ordinea vechimii: E. Borel, F. P. Cantelli, R. Mises, A. N. Kolmogorov, O. Onicescu, Bruno de Finetti, V. Glivenko, A. Renyi şi alţi matematicieni de seamă.

Epoca noastră cunoaşte o dezvoltare considerabilă a acestei teorii, care este aplicată, aproape fără excepţie, în toate domeniile de activitate (fizică, chimie, biologie, tehnică, astronomie, medicină, economie, sociologie, istorie, arheologie, psihologie, lingvistică ş.a.).

Dealtminteri, aplicaţiile teoriei probabilităţilor au mers mînă în mînă cu dezvoltarea ei teoretică. Încă la sfîrşitul secolului XVII au apărut primele calcule de asigurări, iar astronomul E. Halley (1656—1742) a construit prima tabelă de mortalitate a unei populaţii umane. Statistica a căpătat o mare dezvoltare, teoretică şi practică. Întemeietorii statisticii ca ştiinţă trebuie să fie consideraţi F. Galton (1822—1911), K. Pearson (1857—1936), R. Fisher (1890—1962).

Din teoria probabilităţilor s-au desprins, în ultimele decenii, noi discipline ştiinţifice, importante prin aplicaţiile lor: teoria informaţiei, teoria fiabilităţii, programarea matematică, teoria deciziei etc.

În ţara noastră există o însemnată şcoală de teoria probabilităţilor, care printre alte realizări, a introdus în ştiinţă teoria proceselor cu legături complete, aplicate în psihologie şi economie.

1 Cîmp de probabilitate finit

§ 1. Evenimente. Operaţii cu evenimente

1. Evenimente

Să considerăm experienţa aruncării unui zar. Este, evident, vorba de o experienţă aleatoare*, adică de o experienţă al cărei rezultat nu poate fi anticipat cu certitudine, el depinzînd de o serie de factori întîmplători.

Această experienţă se poate repeta de un număr oarecare de ori. Fiecare repetare a experienţei se numeşte *probă*.

Experienţa considerată are o mulţime E de cazuri sau rezultate posibile:

$$E = \{1, 2, 3, 4, 5, 6\}.$$

Legate de această experienţă, putem considera diverse *evenimente*:
A: „apariţia unui număr par",
B: „apariţia unui număr impar",
C: „apariţia unui număr $\leqslant 3$",
D: „apariţia numărului 5" etc.

Fiecare probă atrage după sine fie realizarea, fie nerealizarea oricărui eveniment. Astfel, dacă la o aruncare a zarului apare faţa 4, atunci evenimentul A s-a *realizat*, iar evenimentele B, C, D nu s-au realizat.

* Cuvîntul „aleator" are sens de întîmplător şi provine de la latinescul „alea" (zar).

Este clar că fiecărui eveniment îi corespunde o mulțime de *cazuri favorabile*, care este o submulțime a lui E. Aceasta este mulțimea de cazuri, care realizează evenimentul considerat.

Astfel: evenimentului A îi corespunde submulțimea $\{2, 4, 6\}$ a lui E;
evenimentului B îi corespunde submulțimea $\{1, 3, 5\}$ a lui E;
evenimentului C îi corespunde submulțimea $\{1, 2, 3\}$;
evenimentului D îi corespunde submulțimea $\{5\}$.

Se observă că un eveniment oarecare și submulțimea lui E, asociată evenimentului, se determină reciproc și de aceea nu vom face distincție între ele.

Vom considera deci, fiecare eveniment legat de experiența considerată ca fiind o submulțime a lui E. Astfel, vom scrie:

$$A = \{2, 4, 6\}; \quad B = \{1, 3, 5\}; \quad C = \{1, 2, 3\}; \quad D = \{5\}.$$

Evenimentele care au un singur caz favorabil se numesc *evenimente elementare*. Prin abuz de limbaj, vom numi la fel și elementele mulțimii E.

Mulțimea evenimentelor legate de o *experiență cu un număr finit de cazuri posibile* se identifică, deci, cu familia $\mathcal{P}(E)$ a tuturor submulțimilor mulțimii E, a tuturor cazurilor posibile ale experienței.

2. Eveniment sigur. Eveniment imposibil

Reținem, deci, că orice eveniment este o submulțime a lui E și, reciproc, orice submulțime a lui E este un eveniment.

Dar printre submulțimile lui E se găsesc mulțimea vidă $\emptyset$ și întreaga mulțime E (mulțimea totală). Acestea corespund *evenimentului imposibil* și, respectiv, *evenimentului sigur*.

Evenimentul sigur este acela care se realizează, cu certitudine, la orice probă.

Astfel, la aruncarea unui zar, evenimentul sigur este „apariția uneia din fețele 1, 2, 3, 4, 5, 6". Toate cazurile posibile ale experienței sînt cazuri favorabile acestui eveniment.

Evenimentul imposibil nu se poate realiza în nici o efectuare a experienței. El nu are nici un caz favorabil (spunem că mulțimea cazurilor sale favorabile este vidă).

3. Eveniment implicat de alt eveniment

Spunem că evenimentul A *implică* evenimentul B, dacă realizarea lui A atrage după sine realizarea lui B.

Aceasta înseamnă că orice caz care realizează pe A, realizează și pe B, adică mulțimea cazurilor favorabile lui A este inclusă în mulțimea cazurilor favorabile lui B.

La aruncarea unui zar, dacă $A = \{1, 2, 3\}$, $B = \{1, 2, 3, 4\}$ se vede că A implică B, iar ca relație între mulțimi $A \subset B$.

Se verifică ușor că $A \subset A$, $A \subset E$. Evenimentul imposibil implică orice eveniment ($\emptyset \subset A$).

4. Operații cu evenimente

Fiind date două evenimente A, B, legate de o experiență, „A sau B" este evenimentul a cărui realizare înseamnă realizarea a *cel puțin* unuia din ele.

La aruncarea unui zar, dacă $A = \{1, 2, 3\}$, $B = \{2, 3, 4\}$ atunci

$$„A \text{ sau } B" = \{1, 2, 3, 4\} = A \cup B.$$

Reținem: în loc de „A sau B" scriem $A \cup B$.

Se poate vorbi de „A sau B sau C", „A sau B sau C sau D" etc., respectiv, de $A \cup B \cup C$, $A \cup B \cup C \cup D$ etc.

„A și B" este evenimentul a cărui realizare înseamnă realizarea ambelor evenimente A, B.

La aruncarea zarului, dacă $A = \{1, 2, 3\}$, $B = \{2, 3, 4\}$, atunci ambele evenimente se realizează dacă apare una din fețele 2 sau 3, adică:

$$„A \text{ și } B" = \{2, 3\} = A \cap B.$$

Reținem: în loc de „A și B" vom scrie $A \cap B$, în loc de „A și B și C" vom scrie $A \cap B \cap C$ etc.

$\bar{A}$ = „non A" este evenimentul a cărui realizare constă în nerealizarea evenimentului A. Mulțimea cazurilor favorabile lui „non A" este formată din toate cazurile nefavorabile lui A.

La aruncarea unui zar, dacă $A = \{1, 2, 3\}$, atunci

$$\bar{A} = \text{„non } A\text{“} = \{4, 5, 6\} = \complement A.$$

Reținem : în loc de „non A“ sau „A nu se realizează“ vom scrie $\complement A$.

Se verifică ușor relațiile : $\bar{\bar{A}} = A$; $\bar{E} = \emptyset$; $\bar{\emptyset} = E$.

5. Evenimente incompatibile. Evenimente compatibile

Evenimentele A, B sînt incompatibile *dacă nu se pot realiza împreună în nici o efectuare a experienței.* Aceasta înseamnă că realizarea unuia din cele două evenimente atrage după sine nerealizarea celuilalt, adică A implică „non B“ și B implică „non A“.

Conform celor spuse mai sus, vom scrie $A \subset \complement B$, $B \subset \complement A$.

Se observă că A, B sînt incompatibile, atunci cînd nu au nici un caz favorabil comun.

De asemenea, este clar că A, B sînt evenimente incompatibile dacă și numai dacă realizarea lui „A și B“ este imposibilă („A și B“ = imposibil se scrie, conform celor spuse mai înainte, $A \cap B = \emptyset$).

Se verifică ușor, folosind limbajul evenimentelor sau al mulțmilor, echivalența relațiilor :

$$A \cap B = \emptyset, \ A \subset \complement B, \ B \subset \complement A.$$

Evenimentele A, B sînt compatibile *dacă se pot realiza împreună în aceeași probă, deci dacă au cel puțin un caz favorabil comun.* Pe scurt, A, B sînt compatibile dacă nu sînt incompatibile. La aruncarea zarului, dacă :

$$S = \{1, 2, 3\} ; \ B = \{2, 3, 4\} ; \ C = \{4, 5\}$$

atunci :

— A, B sînt compatibile. Ele se realizează împreună dacă apare una din fețele 2 sau 3.

— B, C sînt compatibile.

— A, C sînt incompatibile. Oricare ar fi rezultatul experienței, ele nu se realizează simultan.

Cu notațiile din teoria mulțimilor putem scrie

$$A \cap B = \{2, 3\} ; \ B \cap C = \{4\} ; \ A \cap B = \emptyset.$$

Se poate vorbi despre incompatibilitatea (sau compatibilitatea) unui număr oarecare de evenimente.

6. Dualitate de limbaj

Din cele spuse mai înainte reiese că orice eveniment legat de o experiență, cu un număr finit de cazuri posibile, poate fi interpretat ca o submulțime a unei mulțimi E (mulțimea cazurilor posibile ale experienței).

De aici a decurs următoarea dualitate de limbaj :

Limbajul evenimentelor	*Limbajul mulțimilor*
Eveniment.	Submulțime a lui E.
Eveniment sigur.	Mulțimea *(totală)* E.
Eveniment imposibil.	Mulțimea vidă $\emptyset$.
A implică B.	$A \subset B$.
A sau B.	$A \cup B$.
A și B.	$A \cap B$.
non A (eveniment contrar lui A).	$\complement A$.
A, B incompatibile.	$A \cap B = \emptyset$.
Eveniment elementar.	$\{e\}$ sau e cu $e \in E$.

● **Aplicație.** *O urnă conține bile albe și bile negre. Se extrag succesiv din urnă două bile.* Cu ajutorul evenimentelor :

A = prima bilă extrasă este albă,

B = a doua bilă extrasă este albă,

să se scrie evenimentele :

C = prima bilă este neagră,

D = cel puțin o bilă este albă,

F = ambele bile sînt negre,

G = o bilă și numai una este albă.

Rezolvare. Evenimentul „prima bilă este neagră“ este contrariul evenimentului „prima bilă este albă“.

$$C = \complement A$$

— Evenimentul „cel puțin o bilă este albă" înseamnă „prima bilă este albă" *sau* „a doua bilă este albă":

$$D = A \cup B.$$

— Evenimentul „ambele bile sînt negre" se poate exprima astfel: „prima bilă este neagră" *și* „a doua bilă este neagră".

Deci:
$$F = \complement A \cap \complement B.$$

Se observă, de asemenea, că F este contrariul lui D:

$$F = \complement(A \cup B).$$

— Evenimentul „o bilă și numai una este albă" se poate scrie: („prima bilă este albă" *și* „a doua bilă este neagră" *sau* „prima bilă este neagră" *și* „a doua bilă este albă").

$$G = (A \cap \complement B) \cap (\complement A \cup B).$$

Menționăm că pentru oricare două evenimente A, B, $A \cap \complement B$ se notează $A - B$, iar $(A - B) \cup (B - A)$ se notează $A \triangle B$ și se numește diferența simetrică a lui A și B. Cu aceste observații putem scrie:

$$G = A \triangle B.$$

§ 2. Probabilitate

1. Frecvența

Fie A un eveniment asociat unei experiențe. Dacă într-o serie de n probe, evenimentul A s-a realizat de n_A ori, atunci numim **frecvența relativă** a evenimentului A numărul $f(A) = \frac{n_A}{n}$. (Acest număr depinde de n și ar fi trebuit notat $f_n(A)$, dar vom folosi și forma $f(A)$.)

Se verifică ușor unele proprietăți ale frecvenței relative:

1° $0 \leqslant f(A) \leqslant 1$ pentru orice eveniment A,

2° $f(E) = 1$ (E = evenimentul sigur),

3° $f(A \cup B) = f(A) + f(B)$ dacă evenimentele A, B sînt incompatibile.

4° $f(A - B) = f(A) - f(B)$ dacă B implică A,

5° $f(A - B) = f(A) - f(A \cap B)$,

6° $f(A \cup B) = f(A) + f(B) - f(A \cap B)$,

7° $f(\complement A) = 1 - f(A)$.

Să verificăm, spre exemplu, relațiile 5° și 6°.

Dacă în cele n probe, A s-a realizat de n_A ori și $A \cap B$ de $n_{A\cap B}$ ori, atunci $f(A) = \frac{n_A}{n}$; $f(A \cap B) = \frac{n_{A\cap B}}{n}$ · De cîte ori s-a realizat $(A - B) = A \cap \complement B$, înseamnă de cîte ori s-a realizat evenimentul constînd în realizarea lui A și nerealizarea lui B. Pentru a găsi acest număr, trebuie să scădem din numărul probelor care au realizat $A (= n_A)$ pe ce al probelor care au realizat și pe $B (= n_{A\cap B})$.

Rezultă $n_{A-B} = n_A - n_{A\cap B}$ și deci:

$$f(A - B) = \frac{n_A - n_{A\cap B}}{n} = \frac{n_A}{n} - \frac{n_{A\cap B}}{n} = f(A) - f(A \cap B).$$

Dacă A, B sînt incompatibile, atunci $n_A + n_B$ este numărul de probe care realizează $A \cup B$ (rezultă 3°). Dacă cele două evenimente sînt compatibile, atunci din $n_A + n_B$ trebuie să scădem $n_{A\cap B}$ (probele care au realizat și pe A și pe B au fost numărate de două ori în suma $n_A + n_B$: o dată printre cele care au realizat pe A și o dată printre cele care au realizat pe B). Deci $n_{A\cup B} = n_A + n_B - n_{A\cap B}$ sau

$$f(A \cup B) = f(A) + f(B) - f(A \cap B).$$

Multe experiențe prezintă fenomenul de *regularitate statistică*, adică dacă m și n sînt numere mari și A este un eveniment legat de o astfel de experiență, atunci $f_n(A)$ și $f_m(A)$ nu diferă mult între ele și această diferență este cu atît mai mică cu cît cele două numere m, n sînt mai mari. Aceasta ne sugerează că frecvențele oscilează în jurul unei anumite valori și că apropierea de această valoare este cu atît mai mare cu cît numărul n este mai mare. În general, nu putem cunoaște această valoare, deoarece nu putem repeta decît de un număr finit de ori experiența.

Exemplu. S-a efectuat următoarea experiență: *s-au introdus într-o urnă două bile — una albă și una neagră — s-au făcut extrageri succesive, punîndu-se după fiecare extragere bila înapoi în urnă. După* 2 000 *de extrageri, bila albă a apărut de* 994 *ori: continuîndu-se experiența, după* 4 000 *extrageri bila albă a apărut de* 2 008 *ori. Deci* $f_{2\,000} = 0{,}497$; $f_{4\,000} = 0{,}502$. *Dacă am continua experiența, am găsi pentru frecvență valori care sînt foarte apropiate de* 0,5.

Noțiunea de frecvență relativă este suportul empiric al noțiunii de probabilitate.

2. Probabilitate

A defini o probabilitate în raport cu o experiență, avînd un număr finit de cazuri posibile, revine la a asocia fiecărui eveniment A, legat de respectiva experiență, un număr $P(A)$ numit probabilitatea evenimentului A. În plus, este natural să cerem ca P să aibă proprietățile frecvenței relative, printre care sînt și proprietățile 1°—7°.

Se constată că proprietățile 1°—3° sînt suficiente, din ele deducîndu-se toate celelalte. Cum orice eveniment poate fi considerat ca o submulțime a unei mulțimi E, o probabilitate P face să corespundă fiecărei submulțimi a lui E un număr real, această corespondență satisfăcînd proprietățile 1°, 2°, 3°.

Definiție. **Aplicația $P: \mathcal{P}(E) \to R$ este o probabilitate dacă**

a) $0 \leqslant P(A) \leqslant 1$,

b) $P(E) = 1$,

c) $P(A \cup B) = P(A) + P(B)$ *dacă* $A \cap B = \varnothing$.

Proprietatea *c)* se extinde imediat

$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n)$

dacă $A_1, \ldots, A_n$ sînt disjuncte două cîte două.

Dacă P este o probabilitate pe $\mathcal{P}(E)$, atunci

d) $P(A - B) = P(A) - P(B)$ dacă $B \subset A$,

e) $P(A - B) = P(A) - P(A \cap B)$,

f) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$,

g) $P(\complement A) = 1 - P(A)$; $P(\varnothing) = 1 - P(E) = 0$.

Demonstrație. d) Dacă $B \subset A$, atunci se arată ușor (folosind limbajul evenimentelor sau al mulțimilor) că :

$$A = B \cup (A - B)$$

și de aici

$P(A) = P(B) + P(A - B)$ (deoarece B și $A - B$ sînt disjuncte),

e) Avem : $A - B = A - (A \cap B)$ (exercițiu !) și $A \cap B \subset A$. Conform cu *d)*,

$$P(A - B) = P(A) - P(A \cap B),$$

f) $$A \cup B = A \cup (B - A) \text{ (exercițiu)}$$

și de aici

$$P(A \cup B) = P(A) + P(B - A) = P(A) + P(B) - P(A \cap B),$$

g) $$A \cup \complement A = E \Rightarrow P(A) + P(\complement A) = 1.$$

Observație importantă

Dacă cunoaștem probabilitatea evenimentelor elementare, cunoaștem probabilitatea oricărui eveniment.

Într-adevăr, dacă $E = \{e_1, e_2, \ldots, e_n\}$ și $A = \{e_{i_1}, e_{i_2}, \ldots, e_{i_k}\}$, atunci,

$$A = \{e_{i_1}\} \cup \{e_{i_2}\} \cup \ldots \cup \{e_{i_k}\}$$

și după axioma *b)* : $P(A) = P(e_{i_1}) + P(e_{i_2}) + \ldots + P(e_{i_k})$,

în particular

$$P(E) = P(e_1) + P(e_2) + \ldots + P(e_n) = 1.$$

3. Cazul evenimentelor elementare echiprobabile

Dacă $E = \{e_1, e_2, \ldots, e_n\}$, am văzut că

$$P(e_1) + P(e_2) + \ldots + P(e_n) = 1.$$

Dacă $\{e_1\}, \{e_2\}, \ldots, \{e_n\}$ au aceeași probabilitate (sînt echiprobabile)

$$P(e_1) = P(e_2) = \ldots = P(e_n)$$

atunci $P(e_i) = \frac{1}{n}$ $(i = 1, 2, \ldots, n)$.

Dacă $A = \{e_{i_1}, e_{i_2}, \ldots, e_{i_k}\}$ atunci

$$P(A) = P(e_{i_1}) + P(e_{i_2}) + \ldots + P(e_{i_k}) = \frac{k}{n}.$$

Rezultă că, în acest caz, *probabilitatea unui eveniment A este egală cu raportul dintre numărul cazurilor favorabile lui A și numărul total al cazurilor posibile ale experienței.*

În multe din aplicații se consideră satisfăcută condiția de echiprobabilitate a cazurilor posibile ale experienței. Astfel :

— Aruncarea unui zar „perfect" este o experiență cu 6 cazuri posibile echiprobabile.

— Extragerea unei bile dintr-o urnă conținînd n bile este o experiență cu n cazuri posibile echiprobabile.

— Aruncarea a două zaruri este o experiență cu 36 de cazuri posibile echiprobabile.

— Aruncarea de n ori a unei monede este o experiență cu 2^n cazuri posibile echiprobabile.

● Aplicație. *O urnă conține 3 bile albe și 4 bile negre, iar o altă urnă conține 4 bile albe și 5 bile negre. Din fiecare urnă se extrage cîte o bilă. Se consideră evenimentele:*

A = bila extrasă din prima urnă este albă,

B = bila extrasă din a doua urnă este albă.

Să se calculeze $P(A \cap B)$, $P(A \cup B)$, $P(A - B)$, $P(\complement A)$.

Rezolvare. Evident $P(A) = \frac{3}{7}$; $P(B) = \frac{4}{9}$.

Experiența constînd în extragerea celor două bile are $7 \cdot 9 = 63$ cazuri (egal) posibile. Evenimentul $A \cap B$ (prima bilă este albă și a doua bilă este albă) are $3 \cdot 4 = 12$ cazuri favorabile. Deci

$$P(A \cap B) = \frac{12}{63} = \frac{4}{21}.$$

Pentru a calcula $P(A \cup B)$, folosim formula:

$$P(A \cup B) = P(A) + P(B) - P(A) \cap B)$$

și rezultă

$$P(A \cup B) = \frac{3}{7} + \frac{4}{9} - \frac{4}{21} = \frac{43}{63}.$$

De asemenea,

$$P(A - B) = P(A) - P(A) \cap B) = \frac{3}{7} - \frac{4}{21} = \frac{5}{21}.$$

$$P(\complement A) = 1 - P(A) = \frac{4}{7}.$$

4. Cîmp finit de probabilitate

Fiind dată o mulțime finită E și o aplicație $\mathcal{P}(E) \to R$, satisfăcînd axiomele a), b), c), spunem că s-a dat un cîmp finit de probabilitate.

În unele cazuri, cînd mulțimea E nu este finită, din motive pe care nu le vom expune aici, mulțimea evenimentelor nu se consideră familia tuturor părților lui E, ci numai o subfamilie $\mathcal{K}$ a acesteia. Dar P fiind o funcție care face ca fiecărui eveniment (element

al lui $\mathcal{K}$) să-i corespundă un număr (probabilitatea sa), pentru a putea scrie formulele $P(A \cup B) = P(A) + P(B) - P(A \cap B)$; $P(\complement A) = 1 - P(A)$ etc. trebuie să ne asigurăm că dacă $A, B \in \mathcal{K}$, atunci $A \cup B$, $A \cap B \in \mathcal{K}$, $\complement A \in \mathcal{K}$. Deoarece $\mathcal{K}$ este domeniul de definiție al funcției P, expresiile $P(A)$, $P(B)$, $P(\complement A)$, $P(A \cup B)$, $P(A \cap B)$, $P(E)$ etc. au sens numai dacă A, B, $\complement A$, $A \cup B$, $A \cap B$, E etc. aparțin lui $\mathcal{K}$.

Probleme

Să se demonstreze:

1. $\complement(A \cup B) = \complement A \cap \complement B$.
2. $\complement(A \cap B) = \complement A \cup \complement B$.

Să se generalizeze cele două relații:

3. Un zar are fețele 1, 2 vopsite roșu, fețele 3, 4 vopsite galben, fețele 5, 6 vopsite albastru. Se aruncă acest zar și se notează:

A : evenimentul obținerii unei fețe roșii;

B : evenimentul obținerii unei fețe galbene;

C : evenimentul obținerii unei fețe albastre;

D : evenimentul obținerii unui număr par;

F : evenimentul obținerii unui număr impar.

$\{k\}$ — evenimentul obținerii feței cu k puncte ($k = 1, 2, 3, 4, 5, 6$). Să se arate că:

a) $\overline{A \cup B} = \overline{A} \cap \overline{B} = C$; ($\overline{A} = \complement A$);

b) $\overline{D} = F$;

c) $A \cap D = \{2\}$; $B \cap D = \{4\}$; $C \cap D = \{6\}$;
$A \cap F = \{1\}$; $B \cap F = \{3\}$; $C \cap F = \{5\}$;

d) $(A \cup B) \cap D = (A \cap D) \cup (B \cap D) = \{2\} \cup \{4\}$.

4. O țintă este formată din 10 cercuri concentrice de raze r_k ($k = 1, 2, \ldots, 10$) $r_1 < r_2 < \ldots < r_{10}$. Evenimentul A_k constă în nimerirea cercului de rază r_k. Să se spună ce înseamnă evenimentele:

a) $A = A_6 - A_5$;

b) $B = A_7 - A_1$;

c) $C = \overline{A}_k$.

Să se arate că:

d) $A_k \subset A_{k+1}$ ($k = 1, 2, \ldots, 9$).

5. O persoană urmează să dea patru telefoane la patru numere diferite. Fiecare număr este format o singură dată. Notăm cu A_i evenimentul că la chemarea i nu primește răspuns. Cum se scriu evenimentele:

a) primește răspuns la toate chemările;

b) la cel mult o chemare nu primește răspuns;

c) la cel puțin o chemare nu primește răspuns;

d) la o singură chemare nu primește răspuns;

e) nu primește răspuns la prima chemare și la încă una din celelalte trei chemări, iar la celelalte două chemări primește răspuns.

6. O urnă conține n bile numerotate cu 1, 2, ..., n. Notăm cu $\{Mk\}$ evenimentul ca la o extragere să obținem o bilă numerotată cu un multiplu de k. Să se arate că :

a) $\{M\,21\} = \{M\,3\} \cap \{M\,7\}$;

b) $\{M\,10\} \cap \{M\,6\} = \{M\,30\}$;

c) dacă $\{Ma\} \subset \{Mb\}$, atunci b este un divizor al lui a și reciproc ;

d) $\{Ma\} \cap \{Mb\} = \{M[a, b]\}$,

unde $[a, b]$ este cel mai mic multiplu comun al numerelor a și b.

7. O urnă conține n bile numerotate 1, 2, ..., n. Din această urnă se extrage o bilă. Evenimentele :

A : bila extrasă dă un pătrat perfect,

B : bila extrasă dă un număr care mărit cu 1 este multiplu de 3,

sînt compatibile ? Să se calculeze $P(A \cup B)$.

8. O urnă conține 10 bile albe și 6 bile negre. Din această urnă se extrag 2 bile, nepunîndu-se înapoi prima bilă extrasă.

Se cere :

a) probabilitatea ca cele 2 bile să fie albe ;

b) probabilitatea ca cele 2 bile să fie negre ;

c) probabilitatea ca prima bilă să fie albă și a doua neagră ;

d) probabilitatea ca prima bilă să fie neagră și a doua albă ;

e) probabilitatea ca bilele să fie de aceeași culoare ;

f) probabilitatea ca bilele să fie de culori diferite.

9. Într-un fișier sînt 10 000 de fișe. Care este probabilitatea ca numărul primei fișe extrase să conțină cifra 5 ?

10. Din 100 mere, 10 sînt stricate. Care este probabilitatea ca luînd 5 mere la întîmplare să luăm și mere stricate ?

11. Din mulțimea numerelor de 7 cifre diferite ce se pot forma cu cifrele 1, 2, 3, 4, 5, 6, 7, se ia la întîmplare un număr. Care este probabilitatea ca numărul ales să conțină pe 1 și 2 ca cifre consecutive în ordine crescătoare ?

12. Se aruncă 6 zaruri. Care este probabilitatea obținerii tuturor numerelor de la 1 la 6 ? Dar probabilitatea să apară cel puțin o dată fața 5 ?

13. Se aruncă o monedă pînă cînd obținem deasupra fața cu stema. Care este probabilitatea să facem cel mult 3 aruncări ?

14. O urnă conține α bile albe și β bile negre. Din această urnă este extrasă o bilă de culoare necunoscută. Care este probabilitatea ca la o nouă extragere să obținem o bilă albă ?

15. O urnă conține 2 bile albe și 3 negre, iar o altă urnă conține 3 bile albe și 4 negre. Dar fiecare urnă este extrasă cîte o bilă.

a) Care este probabilitatea obținerii a 2 bile albe ?

b) Care este probabilitatea obținerii a 2 bile negre ?

c) Care este probabilitatea obținerii a cel puțin unei bile albe ?

d) Care este probabilitatea obținerii a 2 bile de aceeași culoare ?

16. Păstrînd condițiile și notațiile problemei 6, să se arate că

$$a)\ P\{(Ma)\} = \frac{\left[\frac{n}{a}\right]}{n},$$

unde prin $[\alpha]$ notăm partea întreagă a numărului α;

b) dacă n este multiplu de a, atunci probabilitatea obținerii unui număr care nu se divide la a este $\frac{a-1}{a}$.

17. Se amestecă un pachet de 10 cărți de joc numerotate 1, 2, ..., 10. Care este probabilitatea ca prima carte din pachet să fie 1 ? *b)* Dar probabilitatea ca primele două cărți să fie 1 și 2 în această ordine ?

18. Să se demonstreze formula :

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = \Sigma P(A_i) - \Sigma P(A_{i_1} \cup A_{i_2}) + \ldots$$

$$\ldots + (-1)^{k+1} \Sigma P(A_{i_1} \cap \ldots \cap A_{i_k}) + \ldots + (-1)^{n+1} P(A_1 \cap \ldots \cap A_n).$$

19. O urnă conține n bile numerotate 1, 2, ..., n. Se extrag, una cîte una, toate cele n bile din urnă. Spunem că avem o concordanță la extragerea k, dacă la această extragere am obținut bila cu numărul k. Care este probabilitatea obținerii a cel puțin unei concordanțe ?

20. Păstrînd condițiile și notațiile problemei 6 să se arate că :

a) $P(\{M3\} \cup \{M7\}) = P(\{M3\}) + P(\{M7\}) - P(\{M21\})$;

b) $P(\{M6\} \cup \{M10\}) = P(\{M6\}) + P(\{M10\}) - P(\{M30\})$;

c) $P(\{Ma\} \cup \{Mb\}) = P(\{Ma\}) + P(\{Mb\}) - P(\{M[a, b]\})$.

21. Cu notațiile problemei 6 să se arate că :

a) $P(\{M2\} \cup \{M3\} \cup \{M5\}) = P(\{M2\}) + P(\{M3\}) + P(\{M5\}) - P(\{M6\}) - P(\{M10\}) - P(\{M15\}) + P(\{M30\})$;

b) $P(\{M6\} \cup \{M10\} \cup \{M15\}) = P(\{M6\}) + P(\{M10\}) + P(\{M15\}) - 2P(\{M30\})$;

c) $P(\{M30\} \cup \{M42\} \cup \{M70\} \cup \{M105\}) = P(\{M30\}) + P(\{M42\}) + P(\{M70\}) + P(\{M105\}) - 3P(\{M210\})$;

d) $P(\{Ma\} \cup \{Mb\} \cup \{Mc\}) = P(\{Ma\}) + P(\{Mb\}) + P(\{Mc\}) - P(\{M[a, b]\}) - P(\{M[a, c]\}) - P(\{M[b, c]\}) + P(\{M[a, b, c]\})$.

Să se generalizeze ultima relație.

22. Cu notațiile din problema 6 să se arate că dacă a, b, c sînt numere naturale prime între ele două cîte două atunci

a) $P(\{Ma\} \cup \{Mb\} \cup \{Mc\}) = P(\{Ma\}) + P(\{Mb\}) + P(\{Mc\}) - P(\{Mab\}) - P(\{Mac\}) - P(\{Mbc\}) + P(\{Mabc\})$.

Să se generalizeze.

II Probabilități condiționate. Independență

În cele ce urmează ne vom situa permanent într-un cîmp finit de probabilitate.

§ 1. Probabilitate condiționată

1. Definiția și semnificația noțiunii de probabilitate condiționată

Să presupunem că avem două urne: U_1 conținînd 5 bile albe și 6 bile negre și U_2 conținînd 6 bile albe și 7 bile negre. Din una din aceste urne (nu știm din care) se extrage o bilă. Care este probabilitatea ca bila să fie albă? Nu putem răspunde dintr-o dată la această întrebare. Dar dacă căpătăm informația: extragerea s-a făcut din urna U_1, atunci putem spune că probabilitatea ca bila să fie albă este $\frac{5}{11}$. Să reținem deci că $\frac{5}{11}$ nu este probabilitatea ca bila să fie albă, ci probabilitatea ca bila să fie albă, știind că extragerea s-a făcut din urna U_1.

Dacă notăm evenimentele:

A: bila extrasă este albă,

B: extragerea se face din U_1,

C: extragerea se face din U_2,

vom scrie:

$$P_B(A) = \frac{5}{11}.$$

La fel:
$$P_C(A) = \frac{6}{13}.$$

Spunem că *probabilitatea evenimentului A condiționată de evenimentul B* este egală cu $\frac{5}{11}$.

Să considerăm experiența aruncării unui zar. Aceasta are 6 cazuri posibile. Dacă $A = \{1, 2, 3\}$; $B = \{2, 3, 4, 5, 6\}$, atunci $P(A) = \frac{1}{2}$, $P(B) = \frac{5}{6}$.

Să presupunem acum că acoperim fețele 2, 3, 4, 5, 6 cu un strat subțire de vopsea roșie. Dacă în urma aruncării zarului s-a obținut o față roșie, atunci știm că s-a realizat B, dar nu știm ce față a apărut. Ne putem întreba care este probabilitatea ca A să se realizeze după ce am căpătat informația: B este realizat.

În acest moment nu mai avem 6 cazuri posibile, ci numai 5 (cazurile favorabile lui B au devenit cazuri posibile). Din aceste 5 cazuri posibile 2 sînt favorabile lui A. Rezultă $P_B(A) = \frac{2}{5}$.

Să repetăm acest raționament pentru o experiență avînd n cazuri posibile echiprobabile. Dacă B este un eveniment cu m cazuri favorabile, atunci $P(B) = \frac{m}{n}$. Dacă dintre cele m cazuri favorabile lui B, p sînt favorabile unui eveniment A, atunci $P(A \cap B) = \frac{p}{n}$.

În momentul cînd știm că B s-a realizat, mai avem m cazuri posibile. Dintre acestea p sînt favorabile lui A:

$$P_B(A) = \frac{p}{m} = \frac{\frac{p}{n}}{\frac{m}{n}} = \frac{P(A \cap B)}{P(B)}.$$

Raționamentul se poate face și pentru frecvențe relative. Să presupunem că A și B sînt evenimente legate de o experiență $\mathcal{E}$ și că se face de n ori experiența. Să presupunem că evenimentul B s-a realizat de n_B ori, iar ambele evenimente s-au produs de $n_{A \cap B}$ ori.

Aceasta înseamnă că

$$f(B) = \frac{n_B}{n};\ f(A \cap B) = \frac{n_{A \cap B}}{n}.$$

În plus, putem considera că frecvența (absolută) a realizărilor lui A, printre cazurile cînd s-a realizat B, este $n_{A\cap B}$, iar frecvența relativă a lui A condiționată de B, este :

$$f_B(A) = \frac{n_{A\cap B}}{n_B} = \frac{\frac{n_{A\cap B}}{n}}{\frac{n_B}{n}} = \frac{f(A\cap B)}{f(B)}.$$

Aceste observații, ca și altele, pe care nu le prezentăm aici, ne conduc la alegerea egalității :

$$P_B(A) = \frac{P(A\cap B)}{P(B)}, (P(B) \neq 0)$$

ca definiție a probabilității condiționate (și în cazul în care A și B nu sînt legate de o experiență cu un număr finit de cazuri echiprobabile).

2. Probabilitatea intersecției de n evenimente

Să scriem din nou relația de definiție a probabilității evenimentului A_2 condiționată de evenimentul A_1 $(P(A_1) \neq 0)$:

$$P_{A_1}(A_2) = \frac{P(A_1\cap A_2)}{P(A_1)},$$

care se mai poate înscrie :

$$P(A_1 \cap A_2) = P(A_1)\cdot P_{A_1}(A_2).$$

Această relație se poate generaliza :

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)\cdot P_{A_1}(A_2)\cdot P_{A_1\cap A_2}(A_3) \ldots P_{A_1\cap A_2 \cap \ldots \cap A_{n-1}}(A_n).$$

Această relație rezultă înmulțind membru cu membru relațiile :

$$P(A_1) = P(A_1)$$

$$P_{A_1}(A_2) = \frac{P(A_1 \cap A_2)}{P(A_1)},$$

$$P_{A_1\cap A_2}(A_3) = \frac{P(A_1\cap A_2 \cap A_3)}{P(A_1\cap A_2)},$$

. . . .

$$P_{A_1\cap A_2 \cdots \cap A_{n-1}}(A_n) = \frac{P(A_1\cap A_2 \cap \cdots \cap A_n)}{P(A_1\cap A_2\cap \cdots \cap A_{n-1})}.$$

Formula demonstrată este utilă la calcularea probabilității unei intersecții de evenimente, atunci cînd probabilitățile condiționate care apar se calculează cu ușurință fără a folosi definiția. Iată cîteva exemple :

1) Se aruncă două zaruri, unul alb și altul roșu. Este o experiență cu 36 cazuri posibile.

Fie A și B evenimentele :

A : la primul zar apare fața 5 ;

B : numărul total de puncte obținute la cele două zaruri este >9.

$$\text{Avem : } P(A) = \frac{1}{6};\ P(B) = \frac{1}{6}.$$

Într-adevăr cazurile posibile ale experienței sînt perechi de numere (primul număr scris este cel apărut la zarul alb, iar cel de-al doilea numărul apărut la zarul roșu) :

(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)
(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	$\boxed{(4, 6)}$
$\underline{(5, 1)}$	$\underline{(5, 2)}$	$\underline{(5, 3)}$	$\underline{(5, 4)}$	$\underline{\boxed{(5, 5)}}$	$\underline{\boxed{(5, 6)}}$
(6, 1)	(6, 2)	(6, 3)	$\boxed{(6, 4)}$	$\boxed{(6, 5)}$	$\boxed{(6, 6)}$

Cazurile favorabile lui A sînt cele subliniate, iar cele favorabile lui B sînt încadrate. Se observă că două cazuri sînt subliniate și încadrate, adică sînt două cazuri favorabile, atît lui A, cît și lui B (deci lui $A \cap B$). Acestea sînt (5, 5), (5, 6) : $P(A \cap B) = \frac{1}{18}$.

Dacă ne interesează $P_A(B)$, adică probabilitatea ca suma celor două numere obținute să fie >9 știind că primul zar a dat 5, după definiție :

$$P_A(B) = \frac{P(A\cap B)}{P(A)} = \frac{\frac{1}{18}}{\frac{1}{6}} = \frac{1}{3}.$$

La fel

$$P_B(A) = \frac{P(A\cap B)}{P(A)} = \frac{1}{3}.$$

Dar aceste probabilități condiționate se pot calcula mai rapid, printr-un raționament mai direct, fără folosirea definiției probabilității condiționate.

În momentul cînd s-a spus că A s-a realizat, mai avem la dispoziție 6 cazuri posibile (cele care pot realiza pe A) : (5, 1) ; (5, 2) ; (5, 3) ; (5, 4) ; (5, 5) ; (5, 6). Dintre acestea, 2 sînt favorabile lui B : (5, 5) ; (5, 6). Deci

$$P_A(B) = \frac{2}{6} = \frac{1}{3}.$$

2) *Avem 3 urne : U_1 (3 bile albe, 4 bile negre), U_2 (4 bile albe, 5 bile negre), U_3 (5 bile albe, 6 bile negre). Din una din aceste urne se extrage o bilă.*

Să considerăm evenimentele :

A : bila extrasă este albă,

A_1 : extragerea se face din U_1,

A_2 : extragerea se face din U_2,

A_3 : extragerea se face din U_3.

Ne dăm seama imediat, fără a folosi definiția, că

$$P_{A_1}(A) = \frac{3}{7}; \quad P_{A_2}(A) = \frac{4}{9}; \quad P_{A_3}(A) = \frac{5}{11}.$$

3) *O urnă conține* a *bile albe și* b *bile negre. Se extrag succesiv două bile* (fără întoarcerea bilei extrase în urnă).

Să considerăm evenimentele :

A : prima bilă extrasă este albă,

B : a doua bilă extrasă este albă.

Eets clar că $P_A(B) = \dfrac{a-1}{a+b-1}$ (cînd s-a afirmat că A s-a realizat, deci că prima bilă a fost albă, în urnă au rămas $a + b - 1$ bile, dintre care $a - 1$ albe).

Dacă se extrag 3 bile și

A : prima bilă este albă,

B : a doua bilă este albă,

C : a treia bilă este neagră,

$$P_A(B) = \frac{a-1}{a+b-1}; \quad P_{A\cap B}(C) = \frac{b}{a+b-2}.$$

4) *O urnă conține 6 bile albe și 5 bile negre. Se extrag succesiv 3 bile* (fără întoarcerea bilei extrase). *Care este probabilitatea ca prima bilă să fie albă, iar celelalte două negre ?*

Rezolvare. A_1 : prima bilă este albă ; A_2 : a doua bilă este neagră ; A_3 : a treia bilă este neagră

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)\cdot P_{A_1}(A_2)\cdot P_{A_1\cap A_2}(A_3) = \frac{6}{11}\cdot\frac{5}{10}\cdot\frac{4}{9}.$$

3. Evenimente independente

Să reluăm experiența aruncării unui zar. Mulțimea cazurilor posibile este $E = \{1, 2, 3, 4, 5, 6\}$. Dacă $A = \{1, 2, 3\}$; $B = \{2, 3, 4, 5, 6\}$, atunci $P(A) = \frac{1}{2}$; $P(B) = \frac{5}{6}$; $P(\complement A) = \frac{1}{2}$; $P(\complement B) = \frac{1}{6}$. $P_B(A) = \frac{2}{5}$; $P_{\complement B}(A) = 1$; $P_A(B) = \frac{2}{3}$; $P_{\complement A}(B) = 1$ etc.

Ce rezultă din aceste egalități ? Probabilitatea lui A este $\frac{1}{2}$. În momentul cînd căpătăm informația că s-a realizat B, probabilitatea lui A se modifică, devenind $\frac{2}{5}$, iar dacă B nu s-a realizat, devine 1.

La fel, probabilitatea lui B este $\frac{5}{6}$, iar dacă A s-a realizat devine $\frac{2}{3}$ etc. *Concluzie :* fiecare din evenimentele A, B își modifică probabilitatea în funcție de realizarea sau nerealizarea celuilalt. Este natural să spunem că A și B sînt *evenimente dependente*. Să considerăm acum un nou eveniment. $D = \{2, 3, 4, 5\}$. Avem :

$$P_D(A) = \frac{2}{4} = \frac{1}{2}; \quad P_{\complement D}(A) = \frac{1}{2}; \quad P(D) = \frac{4}{6} = \frac{2}{3};$$

$$P_A(D) = \frac{2}{3}; \quad P_{\complement A}(D) = \frac{2}{3} \text{ etc.}$$

Concluzie : evenimentele A, D sînt astfel că niciunul din ele nu-și modifică probabilitatea de a se realiza sau de a nu se realiza, în funcție de realizarea sau nerealizarea celuilalt. Este natural să numim evenimentele A, D independente. S-ar părea că pentru a fi asigurată independența a două evenimente este nevoie să fie satisfăcute multe relații. În realitate, pentru ca două evenimente A, B să fie independente este necesar și suficient ca

$$P(A \cap B) = P(A)\cdot P(B). \tag{1}$$

Într-adevăr, să vedem ce înseamnă că probabilitatea $P(A)$ a evenimentului A, rămîne neschimbată, dacă facem ipoteza că B s-a realizat :

$$P(A) = P_B(A) \Leftrightarrow P(A) = \frac{P(A \cap B)}{P(B)} \Leftrightarrow P(A \cap B) = P(A)\cdot P(B)$$

La fel :

$$P(B) = P_A(B) \Leftrightarrow P(A \cap B) = P(A)\cdot P(B), \; P(A) \neq 0)^*$$

---

* Relația $P(A \cap B) = P(A)\cdot P(B)$ este mai generală ca $P(A) = \dfrac{P(A \cap B)}{P(B)}$ sau $P(B) = \dfrac{P(A \cap B)}{P(A)}$, pentru că nu mai necesită condiția $P(B) \neq 0$, respectiv $P(A) \neq 0$.

adică relația (1) ne asigură că nici probabilitatea $P(B)$ nu se schimbă cînd facem ipoteza că s-a realizat A.

În plus, o relație de tipul (1) rămîne valabilă dacă unul din evenimentele A, B (sau ambele) se înlocuiește cu $\complement A$, respectiv $\complement B$. Într-adevăr să presupunem satisfăcută relația (1). Avem:

$$P(A \cap \complement B) = P(A - B) = P(A) - P(A \cap B) = P(A) - P(A) \cdot P(B) =$$
$$= P(A) \cdot (1 - P(B)).$$
$$P(A \cap \complement B) = P(A) \cdot P(\complement B).$$

Prin definiție, evenimentele $A_1, A_2, \ldots, A_n$ sînt independente, dacă și numai dacă probabilitatea oricărei intersecții de evenimente diferite din cele n este egală cu produsul probabilităților evenimentelor intersectate. Astfel, A, B, C sînt independente dacă

$$P(A \cap B) = P(A) \cdot P(B); \quad P(A \cap C) = P(A) \cdot P(C); \quad P(B \cap C) =$$
$$= P(B) \cdot P(C)$$
$$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C).$$

În multe din problemele pe care le avem de rezolvat, ne dăm ușor seama că unele evenimente care apar sînt independente, în sensul că realizarea unuia (sau unora) dintre ele nu modifică probabilitatea de realizare a celorlalte. În această situație vom putea scrie că probabilitatea intersecției acelor evenimente este egală cu produsul probabilităților lor. Astfel, dacă o experiență $(\mathcal{E})$ constă în efectuarea a două experiențe $(\mathcal{E}_1)$ și $(\mathcal{E}_2)$, care nu-și influențează rezultatele, și dacă A și B sînt evenimentele legate respectiv de $\mathcal{E}_1$ și de $\mathcal{E}_2$, atunci aceste evenimente (ca evenimente ale lui $\mathcal{E}$) sînt independente.

— Aruncarea a două zaruri este o experiență $(\mathcal{E})$ care constă în aruncarea primului zar $(\mathcal{E}_1)$ și aruncarea celui de-al doilea zar $(\mathcal{E}_2)$. Este clar că cunoașterea rezultatului obținut la $(\mathcal{E}_1)$ nu modifică probabilitatea nici unui eveniment legat de $(\mathcal{E}_2)$.

— Extragerea unei bile din fiecare din urnele U_1, U_2, U_3, conținînd bile albe, negre și roșii este o experiență $(\mathcal{E})$ care constă în extragerea unei bile din $U_1(\mathcal{E}_1)$, extragerea unei bile din U_2 $(\mathcal{E}_2)$ și extragerea unei bile din $U_3(\mathcal{E}_3)$.

În acest caz evenimentele:

A : bila extrasă din U_1 este albă,
B : bila extrasă din U_2 este neagră,
C : bila extrasă din U_3 este albă,

sînt independente, ca oricare alte 3 evenimente legate respectiv de $\mathcal{E}_1$, $\mathcal{E}_2$, $\mathcal{E}_3$.

• **Aplicație.** *Se aruncă o monedă de 3 ori. Care este probabilitatea să obținem de fiecare dată „stema"?*

A_1 : „la prima aruncare se obține stema",
A_2 : „la a doua aruncare se obține stema",
A_3 : „la a treia aruncare se obține stema".

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2) \cdot P(A_3) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}.$$

Probleme

1. O urnă conține 3 bile albe și 4 bile negre. Din această urnă se extrage o bilă. În locul ei se introduce o bilă de cealaltă culoare și se face o nouă extragere.

a) Care este probabilitatea ca a doua bilă extrasă să fie neagră, știind că prima a fost albă? Dar știind că prima a fost neagră? Să se observe că dacă

A : „prima bilă extrasă este albă",
B : „prima bilă extrasă este neagră",
C : „a doua bilă extrasă este albă",
D : „a doua bilă extrasă este neagră".

atunci cele două întrebări se pot scrie: $P_A(D) = ?$; $P_B(D) = ?$.

b) Care este probabilitatea ca a doua bilă să fie neagră?

Indicație. Se observă că evenimentul $F =$ „a doua bilă este neagră" înseamnă «„prima este albă *și* a doua este neagră" *sau* „prima este neagră *și* a doua neagră"» sau cu notațiile de la punctul *a)*

$$F = (A \cap D) \cup (B \cap D)$$

și deci

$$P(F) = P(A \cap D) + P(B \cap D).$$

c) Care este probabilitatea să obținem bile de culori diferite în cele două extrageri?

Indicație. Se cere probabilitatea evenimentului F «„prima este albă și a doua neagră" *sau* „prima este neagră și a doua albă"», adică

$$F = (A \cap D) \cup (B \cap C).$$

2. O urnă conține 3 bile albe și 4 bile negre, iar o altă urnă conține 4 albe și 5 negre. Se extrage o bilă din prima urnă și se introduce în cea de-a doua, după care se face o extragere din a doua urnă.

Dacă

A : „prima este albă",

B : „a doua este albă".

a) să se calculeze : $P_A(B)$; $P_{\bar{A}}(B)$; $P_A(\bar{B})$; $P_{\bar{A}}(\bar{B})$; ($\bar{A} = \complement A$; $\bar{B} = \complement B$).

b) Care este probabilitatea ca a doua bilă să fie albă ?

c) Care este probabilitatea ca ambele bile extrase să fie albe ?

d) Care este probabilitatea ca bilele apărute în cele două extrageri să fie de aceeași culoare ?

3. O urnă conține 5 bile albe și 4 bile negre. Se fac 3 extrageri succesive din urnă, fără întoarcerea în urnă a bilei extrase.

a) Care este probabilitatea ca cele 3 bile extrase să fie de aceeași culoare ?

b) Care este probabilitatea ca două bile să fie albe și una să fie neagră ?

4. Să se rezolve problema 14) de la capitolul precedent folosind probabilități condiționate.

5. Să se arate că dacă P este o probabilitate de $\mathcal{P}(E)$ adică funcția $P : \mathcal{P}(E) \to R$ are proprietățile 1° $P(A) \geq 0$; $\forall A \in \mathcal{P}(E)$; 2° $P(E) = 1$; 3° $P(A \cup B) = P(A) + P(B)$, dacă $A \cap B = (\emptyset)$, atunci *a)* P_B este o probabilitate pe $\mathcal{P}(E)$, $(P(B) \neq 0)$; *b)* P_B este o probabilitate pe $\mathcal{P}(B)$*.

Indicație. a) Se arată că $P_B : \mathcal{P}(E) \to R$ are proprietățile 1°, 2°, 3° ; *b)* Se observă că $P_B(B) = 1$.

6. Se dau două urne identice în exterior : U_1 — conținînd 3 bile albe și 4 bile negre și U_2 — conținînd 4 bile albe și 5 bile negre. Din una din aceste urne se ia o bilă. Care este probabilitatea ca bila extrasă să fie albă ?

Indicație. A — „bila extrasă este albă", A_1 — „extragerea se face din U_1" ; A_2 — „extragerea se face din U_2". A obține o bilă albă (A) înseamnă că extragem din $U_1(A_1)$ și $(\cap)$ obținem o bilă albă (A) *sau* $(\cup)$ extragem $U_2(A_2)$ și $(\cap)$ obținem o bilă albă (A), adică :

$$A = (A_1 \cap A) \cup (A_2 \cap A).$$

Rezultă : $P(A) = P(A_1 \cap A) + P(A_2 \cap A)$, *($A_1$, A_2 sînt incompatibile) sau* $P(A) = P(A_1)\cdot P_{A1}(A) + P(A_2)\cdot P_{A2}(A)$.

Se consideră $P(A_1) = P(A_2) = \dfrac{1}{2}$.

7. Sîntem în condițiile problemei 6 de la sfîrșitul capitolului precedent. Să se arate că dacă n este multiplu de a și de b, iar a și b sînt prime între ele, atunci evenimentele $\{Ma\}$ și $\{Mb\}$ sînt independente.

* Restricția funcției $P_B : \mathcal{P}(E) \to R$ la $\mathcal{P}(B)$ este o probabilitate pe $\mathcal{P}(B)$. În $\mathcal{P}(B)$, complementarea se ia în raport cu B.

8. Urna U_1 conține 8 bile numerotate 1, 2, . . ., 8, iar urna U_2 conține 7 bile numerotate 1, 2, . . ., 7. Se efectuează experiența $\mathcal{E}$ constînd dintr-o extragere din urna U_1 (experiența $\mathcal{E}_1$) și o extragere din urna U_2 (experiența $\mathcal{E}_2$).

a) Cîte cazuri (egal) posibile are experiența $\mathcal{E}$?

Notăm cu A evenimentul legat de $\mathcal{E}_1$, „numărul extras din U_1 este par" și tot cu A evenimentul legat de $\mathcal{E}$ „numărul extras din U_1 este par". La fel B este evenimentul legat de $\mathcal{E}_2$ sau $\mathcal{E}$ „numărul extras (din U_2) este par".

b) Cîte cazuri favorabile are evenimentul A, ca eveniment legat de $\mathcal{E}_1$? Dar ca eveniment legat de $\mathcal{E}$? Cîte cazuri favorabile are evenimentul B ca eveniment legat de $\mathcal{E}_2$? Dar ca eveniment legat de $\mathcal{E}$?

c) Comparați probabilitatea lui A ca eveniment legat de $\mathcal{E}_1$ și apoi ca eveniment legat de $\mathcal{E}$. La fel pentru evenimentul B. Pare natural rezultatul obținut ?

d) Să se calculeze $P(A \cap B)$, (A, B evenimentele legate de $\mathcal{E}$). Să se arate că $P(A \cap B) = P(A)\cdot P(B)$.

e) Să se generalizeze pentru cazul cînd urnele U_1 și U_2 conțin n_1 și respectiv n_2 bile, iar A și B sînt două evenimente legate respectiv de $\mathcal{E}_1$ și $\mathcal{E}_2$.

9. Doi trăgători trag cîte un foc asupra unei ținte. Primul nimerește ținta cu probabilitatea $\dfrac{7}{9}$, iar al doilea cu probabilitatea $\dfrac{9}{11}$. Care este probabilitatea ca ținta să fie atinsă ?

10. 3 trăgători trag cîte un foc asupra unei ținte. Primul nimerește ținta cu probabilitatea $\dfrac{3}{4}$, al doilea cu probabilitatea $\dfrac{4}{5}$, iar al treilea cu probabilitatea $\dfrac{5}{6}$. Care este probabilitatea ca ținta să fie atinsă ?

11. Se aruncă de 3 ori o pereche de zaruri. Care este probabilitatea să obținem un total de 6 puncte la prima aruncare, 7 puncte la a doua aruncare și 8 puncte la a treia aruncare ? Dar probabilitatea să obținem 6 puncte la o aruncare, 7 puncte la o altă aruncare și 8 puncte la altă aruncare ?

12. Urna U_1 conține 3 bile albe și 4 bile negre, U_2 — 4 albe și 5 negre, U_3 — 5 albe și 6 negre. Din fiecare urnă se ia cîte o bilă. *a)* Care este probabilitatea ca bilele extrase să fie de aceeași culoare ? *b)* care este probabilitatea ca o bilă să fie albă și două să fie negre ?

13. Două persoane joacă un joc. Partida este cîștigată de cel care obține primul al treilea punct. Dacă scorul este 2 — 1 în favoarea unuia din jucători, care este probabilitatea ca el să cîștige partida ? Se admite că jucătorii sînt de forțe egale, adică orice punct pus în joc este adjudecat de unul din cei doi parteneri cu probabilitatea 1/2.

III Scheme de probabilitate

1. Schema binomială generalizată (Poisson)

Dacă $A_1, A_2, \ldots, A_n$ sînt evenimente independente, atunci probabilitatea să se realizeze k din cele n evenimente (și să nu se realizeze $n-k$) este coeficientul lui x^k din polinomul: $(p_1x+q_1)(p_2x+q_2)\ldots(p_nx+q_n)$, unde probabilitatea $P(A_i)=p_i$; $q_i=1-p_i$ $(i=1, 2, \ldots, n)$.

Cum scriem evenimentul A, a cărui realizare înseamnă realizarea a k din cele n evenimente? Pentru a se realiza A, trebuie să se realizeze k din evenimentele A_i, (fie $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$ aceste evenimente) și să nu se realizeze celelalte $n-k$: $A_{i_{k+1}}, \ldots, A_{i_n}$, adică trebuie să se realizeze unul din evenimentele de forma: $A_{i_1}\cap A_{i_2}\cap\ldots\cap A_{i_k}\cap \cap \bar{A}_{i_{k+1}}\cap\ldots\cap\bar{A}_{i_n}$.

Rezultă că A este reuniunea evenimentelor incompatibile de această formă:

$$A=\bigcup(A_{i_1}\cap A_{i_2}\cap\ldots\cap A_{i_k}\cap\bar{A}_{i_{k+1}}\cap\ldots\cap\bar{A}_{i_n})$$

unde $\{i_1, i_2, \ldots, i_k\}$ parcurge familia submulțimilor de k elemente ale mulțimii de indici $\{1, 2, \ldots, n\}$. Rezultă

$$P(A)=\Sigma p_{i_1}\cdot p_{i_2}\ldots p_{i_k}\cdot q_{i_{k+1}}\ldots q_{i_n}.$$

Se observă că suma din membrul drept al acestei egalități este egală cu coeficientul lui x^k din polinomul $(p_1x+q_1)(p_2x+q_2)\ldots(p_nx+q_n)$.

● **Aplicație.** *Se dau 3 urne: prima conține 2 bile albe și 3 bile negre, a doua conține 4 bile albe și o bilă neagră, iar a treia conține 3 bile albe și 2 bile negre. Din fiecare urnă se extrage cîte o bilă. Care este probabilitatea ca 2 bile să fie albe și una neagră?*

Rezolvare. Considerăm evenimentele independente:

A_1 = bila extrasă din prima urnă este albă;

A_2 = bila extrasă din a doua urnă este albă;

A_3 = bila extrasă din a treia urnă este albă.

Problema cere probabilitatea realizării a 2 din cele 3 evenimente. Sîntem în cazul schemei lui Poisson cu

$$n=3;\ k=2;\ p_1=P(A_1)=\frac{2}{5};\ p_2=P(A_2)=\frac{4}{5};$$

$$p_3=P(A_3)=\frac{3}{5}.$$

Probabilitatea căutată este coeficientul lui x^2 din polinomul

$$\left(\frac{2}{5}x+\frac{3}{5}\right)\left(\frac{4}{5}x+\frac{1}{5}\right)\left(\frac{3}{5}x+\frac{2}{5}\right),$$

adică $\frac{58}{125}$.

2. Schema binomială (Bernoulli)*

Dacă evenimentele $A_1, A_2, \ldots, A_n$ au aceeași probabilitate, $p_i=p$; $q_i=q=1-p$, $(i=1, 2, \ldots, n)$, atunci probabilitatea să se realizeze k din cele n evenimente este coeficientul lui x^k din polinomul $(px+q)^n$, adică este egală cu $C_n^kp^kq^{n-k}$.

Aruncarea unei monede de n ori este o experiență $\mathcal{E}$ care constă în prima aruncare $(\mathcal{E}_1)$, a doua aruncare $(\mathcal{E}_2)$ etc.

Dacă A_1: „apare stema la prima aruncare“

A_2: „apare stema la a doua aruncare“ etc.

atunci $A_1, A_2, \ldots, A_n$ sînt respectiv legate de $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n$, deci sînt independente și $p=P(A_1)=P(A_2)=\ldots=P(A_n)=\frac{1}{2}$, $\left(q=\frac{1}{2}\right)$.

În acest caz, probabilitatea ca în cele n aruncări să obținem de k ori stema este $C_n^k\left(\frac{1}{2}\right)^k\cdot\left(\frac{1}{2}\right)^{n-k}$.

În general, dacă A este un eveniment legat de o experiență și $P(A)=p$ și dacă repetăm de n ori experiența, atunci probabilitatea ca A să se

* J. Bernoulli (1654—1705) a studiat cel dintîi această schemă, care constituie modelul matematic al multor fenomene întîlnite în natură și în societate.

realizeze de k ori (prin abuz de limbaj, am spus A în loc de $A_1, A_2, \ldots, A_n$) este $C_n^k p^k q^{n-k}$, $(q = 1 - p)$.

Modelul matematic al schemei lui Bernoulli poate fi dat de o urnă în care avem a bile albe și b bile roșii și efectuăm din urnă n extrageri, punînd după fiecare extragere bila extrasă înapoi în urnă. Probabilitatea ca la o extragere să obținem o bilă albă este :

$$p = \frac{a}{a + b}.$$

Conform celor arătate mai sus, probabilitatea ca din n extracții să obținem k bile albe este

$$C_n^k p^k q^{n-k}.$$

● **Aplicație.** *Se aruncă două zaruri de 10 ori. Care este probabilitatea să apară de 4 ori suma 7 ?*

Rezolvare. La o efectuare a experienței evenimentul „apariția sumei 7" are probabilitatea $\frac{1}{6}$ (6 cazuri favorabile din 36 posibile). Deci :

$$p = \frac{1}{6};\quad q = \frac{5}{6};\quad n = 10;\quad k = 4. \text{ Răspuns : } C_{10}^4 \left(\frac{1}{6}\right)^4 \left(\frac{5}{6}\right)^6.$$

3. Schema hipergeometrică

Dintr-o urnă, în care sînt a bile albe și b bile roșii $(a + b = N)$, se extrag n bile, $n \leqslant N$, fără ca să se pună după fiecare extragere bila extrasă înapoi în urnă. Însemnăm prin α numărul de bile albe obținut în n extrageri. Evident, $\alpha \leqslant n$ și $\alpha \leqslant a$. Prin urmare : $\alpha \leqslant \min(a, n)$. La fel, numărul de bile roșii, obținut în n extrageri, egal cu $n - \alpha$, trebuie să îndeplinească condiția :

$$n - \alpha \leqslant b \text{ sau } \alpha \geqslant n - b.$$

Cum $\alpha \geqslant 0$, rezultă că $\alpha \geqslant \max(0, n - b)$.

Deci :

$$\max[0, n - b] \leqslant \alpha \leqslant \min[a, n].$$

Probabilitatea ca din n extrageri, efectuate în modul pe care l-am arătat, să obținem α bile albe este

$$P_n(\alpha) = \frac{C_a^\alpha C_b^{n-\alpha}}{C_N^n}.$$

Într-adevăr numărul cazurilor posibile este dat de numărul de combinări făcute cu numărul N de bile din urnă, luat în raport cu numărul n de bile extrase. Să calculăm acum numărul cazurilor favorabile, adică numărul grupurilor de α bile albe și $(n - \alpha)$ bile roșii.

Avem C_a^α grupuri de α bile albe pe care le putem forma cu cele a bile albe din urnă și $C_b^{n-\alpha}$ grupări de $(n - \alpha)$ bile roșii, pe care le putem forma cu cele b bile roșii existente în urnă. Avem deci :

$$C_a^\alpha C_b^{n-\alpha}$$

cazuri favorabile. Rezultă formula de mai sus și egalitatea

$$\sum_{\alpha = \max(0,\, n-b)}^{\min.(a,\, n)} \frac{C_a^\alpha C_b^{n-\alpha}}{C_N^n} = 1,$$

care propunem să fie demonstrată și direct, folosind teoria combinărilor.

Numele de schemă hipergeometrică provine de la faptul că valorile $P_n(\alpha)$ intervin în seria hipergeometrică a lui Gauss (1777—1855), unul din marii matematicieni ai lumii, care a publicat lucrări importante și în teoria probabilităților.

Probleme

1. Într-o cutie sînt 4 pachete a cîte 20 de țigări. În primul pachet este o țigară ruptă, în al doilea sînt 2 țigări rupte, în al treilea 3 țigări rupte, iar în al patrulea 4. Din fiecare pachet se ia cîte o țigară. *a)* Care este probabilitatea să iasă 3 țigări nerupte și una ruptă ? *b)* Dar probabilitatea să iasă cel puțin 3 țigări nerupte ?

2. Se aruncă 2 zaruri de 10 ori. Care este probabilitatea să obținem suma 7 exact de trei ori ?

3. Se aruncă o monedă de 8 ori. Care este probabilitatea să obținem de 4 ori stema și de 4 ori banul ? Dar probabilitatea să obținem de cel puțin 4 ori stema ?

4. Se aruncă un zar de 10 ori. Care este probabilitatea obținerii de 4 ori a unei fețe cu un număr mai mare de 4 puncte.

5. Se dau 4 urne : U_1 conține 3 bile albe și 4 negre, U_2 conține 2 bile albe și 5 negre, U_3 conține 6 bile albe și 2 negre, U_4 conține 4 bile albe și 3 bile negre. Din prima urnă se fac 3 extrageri, punîndu-se de fiecare dată bila înapoi în urnă, iar din celelalte 3 urne se face cîte o extragere.

Care este probabilitatea obținerii sau a 2 bile albe și una neagră din prima urnă, sau a 2 bile albe și una neagră din următoarele 3 urne ?

6. Să considerăm urnele U_1, U_2, U_3 avînd compozițiile U_1 — 5 bile albe și 5 negre; U_2 — 4 bile albe și 6 negre; U_3 — 4 bile albe și 5 negre. Din fiecare urnă se extrag cîte 5 bile, punîndu-se bila extrasă înapoi în urnă. Care este probabilitatea ca din 2 urne să obținem 2 bile albe și 3 bile negre, iar din a treia urnă să obținem altă combinație?

7. Să considerăm urnele din problema precedentă. Din fiecare urnă se extrage cîte o bilă, bila extrasă punîndu-se înapoi. Dacă se efectuează de 5 ori experiența, care este probabilitatea ca de 3 ori să obținem o bilă albă și de 2 ori o bilă neagră?

8. Un fumător cumpără două cutii de chibrituri. Apoi, de fiecare dată cînd are nevoie, scoate la întîmplare una sau alta din cutii. *a)* Care este probabilitatea ca în momentul în care constată că una din cutii este goală, cealaltă cutie să mai conțină k bețe, dacă inițial fiecare cutie a conținut cîte n bețe? (Problema lui Banach). *b)* Utilizînd rezultatul obținut, să se arate că:

$$C_{2n}^{n} + 2C_{2n-1}^{n} + 2^2C_{2n-2}^{n} + \ldots + 2^nC_n^n = 2^{2n}.$$

9. Dintr-o urnă în care sînt așezate toate numerele întregi de la 1 la 90 se extrag 6 numere. Care este probabilitatea ca să iasă trei din numerele:

$$5,\ 72,\ 27,\ 31,\ 16,\ 82\ ?$$

10. Un profesor pregătește pentru examenul oral al elevilor săi 20 de bilete, dintre care 10 sînt de algebră, 5 de geometrie și 5 de aritmetică. Un elev trage succesiv trei bilete, fără a pune înapoi biletul pe care l-a tras. Să se determine probabilitatea ca:

a) cele trei bilete să fie de algebră;

b) un singur bilet să fie de geometrie;

c) cel puțin un bilet să fie de geometrie.

IV Variabile aleatoare. Valori medii

§ 1. Definiția variabilei aleatoare. Exemple

În viața de toate zilele întîlnim la tot pasul mărimi care iau valori ce se schimbă sub influența unor factori întîmplători. Așa sînt, de exemplu, numărul de zile dintr-un an în care cade ploaia peste o anumită regiune, numărul băieților din 100 nou-născuți, numărul de puncte care apar la aruncarea unui zar, numărul de bile albe care apar în n extrageri dintr-o urnă care conține bile de diferite culori, printre care și bile albe, masa unui bob de mazăre luat dintr-o anumită recoltă, rezultatul obținut în urma măsurării unei mărimi fizice, viteza unei molecule de gaz etc. În capitolul de față ne interesează dintre aceste mărimi numai acelea care iau un număr finit de valori. Fiecare din mărimile de mai sus poate lua diferite valori în diversele efectuări ale experienței, chiar dacă toate condițiile rămîn neschimbate la fiecare efectuare a experienței. Modificarea valorilor are la bază factori întîmplători. De aceea vom numi aceste mărimi *variabile aleatoare (întîmplătoare)*.

Pentru cunoașterea unei variabile aleatoare trebuie să cunoaștem, în primul rînd, valorile pe care le poate lua. Dar cunoașterea acestor valori este departe de a fi suficientă. După cum am văzut, fiecare valoare este luată sub influența unor factori întîmplători. Deci, unele valori pot apărea mult mai des decît celelalte. Variabila aleatoare va fi mult mai bine precizată dacă vom cunoaște și probabilitatea cu care este luată fiecare valoare.

putem scrie

$$X + Y + Z\begin{pmatrix} x_1+y_1+z_1 & x_1+y_1+z_2 & \dots & x_i+y_j+z_k & \dots & x_m+y_n+z_s \\ p_{111} & p_{112} & \dots & p_{ijk} & \dots & p_{mns} \end{pmatrix}.$$

Cînd scriem tabloul de distribuţie al unei variabile aleatoare este bine să avem în vedere ca valorile din primul rînd să fie diferite două cîte două.

● **Aplicaţie.** *Probabilitatea extragerii unei bile albe dintr-o urnă este p. Din această urnă se fac 2 extrageri, punîndu-se înapoi bila după prima extragere. Se consideră variabilele X_1 şi X_2, prima reprezentînd numărul de bile albe ieşite la prima extragere şi a doua numărul de bile albe ieşite la a doua extragere. Să se scrie distribuţia sumei celor două variabile aleatoare.*

Evident, cele două variabile au distribuţiile:

$$X_1\begin{pmatrix} 1 & 0 \\ p & q \end{pmatrix}; \quad X_2\begin{pmatrix} 1 & 0 \\ p & q \end{pmatrix},$$

unde $q = 1 - p$. Conform definiţiei sumei variabilelor aleatoare putem scrie:

$$X_1 + X_2 \begin{pmatrix} 1+1 & 1+0 & 0+1 & 0+0 \\ p^2 & pq & qp & q^2 \end{pmatrix}.$$

Vom scrie

$$X_1 + X_2 \begin{pmatrix} 2 & 1 & 0 \\ p^2 & 2pq & q^2 \end{pmatrix}.$$

Variabila sumă reprezintă numărul de bile albe ieşite în 2 extrageri din urnă.

3. Produsul variabilelor aleatoare

Fiind date două variabile aleatoare X şi Y, vom numi *produsul* lor variabila XY, care ia valoarea x_iy_j, atunci cînd X ia valoarea x_i şi Y ia valoarea y_j.

Dacă X şi Y au distribuţiile

$$X\begin{pmatrix} x_1 & x_2 & \dots & x_m \\ p_1 & p_2 & \dots & p_m \end{pmatrix},$$

$$Y\begin{pmatrix} y_1 & y_2 & \dots & y_n \\ q_1 & q_2 & \dots & q_n \end{pmatrix},$$

XY are distribuţia

$$XY\begin{pmatrix} x_1y_1 & x_1y_2 & \dots x_iy_j & \dots & x_my_n \\ p_{11} & p_{12} & \dots & p_{ij} & \dots & p_{mn} \end{pmatrix},$$

unde p_{ij} este probabilitatea realizării simultane a egalităţilor.

$$X = x_i \quad Y = y_j.$$

Fiind date mai multe variabile aleatoare $X, Y, \dots, V$, vom numi produsul lor variabila, $X \cdot Y \dots V$, care ia valoarea $x_iy_j \dots v_k$, dacă $X, Y, \dots, V$ iau respectiv valorile $x_i, y_j, \dots, v_k$.

De exemplu, fiind date 3 variabile aleatoare

$$X\begin{pmatrix} x_1 & x_2 & \dots & x_m \\ p_1 & p_2 & \dots & p_m \end{pmatrix},$$

$$Y\begin{pmatrix} y_1 & y_2 & \dots & y_n \\ q_1 & q_2 & \dots & q_n \end{pmatrix},$$

$$Z\begin{pmatrix} z_1 & z_2 & \dots & z_s \\ r_1 & r_2 & \dots & r_s \end{pmatrix}.$$

putem scrie

$$XYZ\begin{pmatrix} x_1y_1z_1 & x_1y_1z_2 & \dots & x_iy_jz_k & \dots & x_my_nz_s \\ p_{111} & p_{112} & \dots & p_{ijk} & \dots & p_{mns} \end{pmatrix}.$$

unde p_{ijk} este probabilitatea realizării simultane a relaţiilor

$$X = x_i, \ Y = y_j, \ Z = z_k.$$

● **Aplicaţie.** *Să se scrie distribuţia produsului X_1X_2, unde X_1, X_2 sînt variabilele de la aplicaţia din paragraful precedent.*

Folosind direct definiţia produsului variabilelor aleatoare putem scrie

$$X_1X_2 \begin{pmatrix} 1\cdot 1 & 1\cdot 0 & 0\cdot 1 & 0\cdot 0 \\ p^2 & pq & qp & q^2 \end{pmatrix}.$$

Vom scrie deci

$$X_1X_2 \begin{pmatrix} 1 & 0 \\ p^2 & 2pq + q^2 \end{pmatrix}.$$

convențională valorile obținute, putem spune că aceste valori sînt:

$$\underbrace{x_1, x_1, \ldots, x_1}_{n_1} \underbrace{x_2, x_2, \ldots, x_2}_{n_2}, \ldots, \underbrace{x_m, x_m, \ldots, x_m}_{n_m}.$$

Media aritmetică a acestor numere este

$$\frac{n_1x_1 + n_2x_2 + \ldots + n_mx_m}{n} = \frac{n_1}{n} x_1 + \frac{n_2}{n} x_2 + \ldots + \frac{n_m}{n} x_m =$$

$$= f_1x_1 + f_2x_2 + \ldots + f_mx_m,$$

unde f_i este frecvența relativă corespunzătoare valorii x_i $(i=1, 2, \ldots, m)$. Această observație ne conduce în mod natural la următoarea

D e f i n i ț i e. Fiind dată o variabilă aleatoare

$$X \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ p_1 & p_2 & \ldots & p_n \end{pmatrix},$$

vom numi valoarea medie a acestei variabile numărul

$$M(X) = p_1x_1 + p_2x_2 + \ldots + p_nx_n = \sum_{i=1}^{n} p_ix_i.$$

Exemple

Variabilele X_1 și X_2 din aplicația de la punctul 2 din § 2, avînd distribuțiile

$$\begin{pmatrix} 1 & 0 \\ p & q \end{pmatrix},$$

au valoarea medie

$$1 \cdot p + 0 \cdot q = p.$$

$X_1 + X_2$ avînd distribuția

$$X_1 + X_2 \begin{pmatrix} 2 & 1 & 0 \\ p^2 & 2pq & q^2 \end{pmatrix},$$

are valoarea medie

$$2p^2 + 1 \cdot 2pq + 0 \cdot q^2 = 2p(p + q) = 2p.$$

Dacă luăm ca variabilă X numărul de puncte ieșite la aruncarea unui zar, distribuția variabilei X este

$$X \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix},$$

iar valoarea medie a acestei variabile este

$$M(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{7}{2}.$$

Să scoatem în evidență cîteva proprietăți ale valorii medii.

1) *Valoarea medie a unei constante este egală cu constanta.*

Distribuția unei variabile aleatoare care ia o singură valoare este de forma $\begin{pmatrix} a \\ 1 \end{pmatrix}$ și deci valoarea medie va fi egală cu $a \cdot 1 = a$.

2) *Dacă X este o variabilă aleatoare și a o constantă, atunci sînt adevărate relațiile:*

a) $M(a + X) = a + M(X)$,

b) $M(aX) = aM(X)$.

Într-adevăr, fie

$$X \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ p_1 & p_2 & \ldots & p_n \end{pmatrix}$$

distribuția variabilei X. Distribuția variabilei $a + X$ este

$$a + X \begin{pmatrix} a + x_1 & a + x_2 & \ldots & a + x_n \\ p_1 & p_2 & \ldots & p_n \end{pmatrix}$$

și media

$$M(a + X) = (a + x_1)p_1 + (a + x_2)p_2 + \ldots + (a + x_n)p_n =$$

$$= a(p_1 + p_2 + \ldots + p_n) + (p_1x_1 + p_2x_2 + \ldots + p_nx_n) = a + M(X).$$

Distribuția variabilei aX este

$$aX \begin{pmatrix} ax_1 & ax_2 & \ldots & ax_n \\ p_1 & p_2 & \ldots & p_n \end{pmatrix};$$

$$(X = x_i) = [(X = x_i) \cap (Y = y_1)] \cup [(X = x_i) \cap (Y = y_2)] \cup \ldots \cup [(X = x_i) \cap (Y = y_n)];$$

$$p_i = P(X = x_i) = P[(X = x_i) \cup (Y = y_1)] + P[(X = x_i) \cap (Y = y_2)] + \ldots + P([X = x_i) \cap (Y = y_n)]$$

sau

$$p_i = p_{i1} + p_{i2} + \ldots + p_{in}$$

Raționînd la fel, se poate arăta că

$$q_j = p_{1j} + p_{2j} + \ldots + p_{mj}.$$

Putem deci scrie

$$M(X + Y) = \sum_{i=1}^{m} x_i p_i + \sum_{j=1}^{n} y_i q_j = M(X) + M(Y).$$

Valoarea medie a unei sume de două variabile aleatoare este egală cu suma valorilor medii ale celor două variabile.

Proprietatea este adevărată pentru suma unui număr finit oarecare de variabile aleatoare. Demonstrația acestei afirmații se face prin recurență. Pentru aceasta este suficient să observăm că putem scrie

$$X + Y + Z = (X + Y) + Z.$$

Din această relație rezultă proprietatea pentru trei variabile aleatoare

$$M(X + Y + Z) = M(X + Y) + M(Z) = M(X) + M(Y) + M(Z).$$

În general, dacă știm că

$$M(X_1 + X_2 + \ldots + X_n) = M(X_1) + M(X_2) + \ldots + M(X_n),$$

din relația

$$X_1 + X_2 + \ldots + X_n + X_{n+1} = (X_1 + X_2 + \ldots + X_n) + X_{n+1},$$

rezultă

$$M(X_1 + X_2 + \ldots + X_n + X_{n+1}) = M(X_1 + X_2 + \ldots + X_n) + M(X_{n+1}) = M(X_1) + M(X_2) + \ldots + M(X_n) + M(X_{n+1}).$$

● **Aplicația 1.** *Să considerăm variabilele X_1 și X_2 de la aplicația de la punctul 2 din § 2. Am văzut că*

$$M(X_1) = p \qquad M(X_2) = p.$$

Rezultă că

$$M(X_1 + X_2) + M(X_1) + M(X_2) = 2p.$$

Acest rezultat l-am găsit și pe altă cale (la începutul acestui paragraf).

● **Aplicația 2.** *Se aruncă 4 zaruri. Să se calculeze valoarea medie a numărului de puncte obținute.*

Să notăm cu X_1, X_2, X_3, X_4 respectiv numărul de puncte obținute la primul, la al doilea, la al treilea și la al patrulea caz. Distribuția fiecăreia din aceste variabile aleatoare este

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$$

Dacă notăm cu X suma numerelor de puncte ieșite pe fiecare din cele 4 zaruri pentru a calcula valoarea medie a variabilei X este suficient să observăm că

$$X = X_1 + X_2 + X_3 + X_4.$$

De aici rezultă

$$M(X) = M(X_1) + M(X_2) + M(X_3) + M(X_4) = 14.$$

5) *Valoarea medie a unui produs finit de variabile aleatoare.*

Fiind date două variabile aleatoare

$$X \begin{pmatrix} x_1 & x_2 & \ldots & x_m \\ p_1 & p_2 & \ldots & p_m \end{pmatrix}, \qquad Y \begin{pmatrix} y_1 & y_2 & \ldots & y_n \\ q_1 & q_2 & \ldots & q_n \end{pmatrix},$$

valoarea medie a variabilei

$$XY \begin{pmatrix} x_1 y_1 & x_1 y_2 & \ldots & x_i y_i & \ldots & x_m y_n \\ p_{11} & p_{12} & \ldots & p_{ij} & \ldots & p_{mn} \end{pmatrix}$$

este

$$M(XY) = p_{11} x_1 y_1 + p_{12} x_1 y_2 + \ldots + p_{1n} x_1 y_n + p_{21} x_2 y_1 + p_{22} x_2 y_2 + \ldots + p_{2n} x_2 y_n + \ldots + p_{m1} x_m y_1 + p_{m2} x_m y_2 + \ldots + p_{mn} x_m y_n.$$

O variabilă aleatoare care ia numai valori pozitive are momente de orice ordin.

Fiind date o variabilă aleatoare X și o constantă α, *vom numi abaterea de la constanta α a variabilei X, variabila $X - \alpha$. Momentul de ordinul k raportat la constanta α a variabilei X, se definește ca valoarea medie a variabilei $(X - \alpha)^k$.* Pentru $\alpha = 0$, obținem momentul inițial de ordinul k, sau, pe scurt, momentul de ordinul k al variabilei X pe care l-am definit la începutul acestui paragraf. Pentru $\alpha = M(X) = m$, obținem momentul centrat de ordinul k al variabilei X. Variabila $X - M(X)$ se numește abaterea de la medie a variabilei X sau, pe scurt, abaterea variabilei X.

Valoarea medie a abaterii oricărei variabile aleatoare este nulă. Într-adevăr,

$$M(X - M(X)) = M(X) - M(X) = 0.$$

De multe ori, la o variabilă aleatoare ne interesează cît de mult se abat valorile variabilei de la valoarea medie. Trebuie să stabilim un indicator numeric al împrăștierii valorilor variabilei aleatoare în jurul valorii medii. Valoarea medie a abaterii de la medie nu poate caracteriza această împrăștiere, deoarece este nulă pentru orice variabilă aleatoare. Abaterile diferitelor valori, avînd semne diferite, se compensează reciproc. Este foarte firesc să caracterizăm împrăștierea variabilei X, prin valoarea medie a abaterii absolute $|X - M(X)|$, pe care o vom numi *abatere medie*. Dacă X are distribuția

$$X\begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix},$$

atunci distribuția abaterii absolute este

$$\begin{pmatrix} |x_1 - m| & |x_2 - m| & \dots & |x_n - m| \\ p_1 & p_2 & \dots & p_n \end{pmatrix},$$

unde $m = M(X)$, iar abaterea medie este

$$p_1|x_1 - m| + p_2|x_2 - m| + \dots + p_n|x_n - m|.$$

Se observă că semnul expresiilor $x_i - m$ nu influențează valoarea abaterii medii.

Putem lua ca indicator al împrăștierii variabilei X și $M[(X - m)^2]$ sau $M[(X - m)^4]$ etc. Folosirea abaterii medii este foarte incomodă în calcule. Foarte comodă este în schimb folosirea expresiei

$$M[(X - m)^2].$$

2. Dispersia

Vom numi *dispersie* a unei variabile aleatoare X momentul centrat de ordinul al doilea al acestei variabile. Vom scrie

$$\sigma^2 = D^2(X) = M[(X - m)^2],$$

unde $m = M(X)$.

Dacă X are distribuția

$$X\begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix},$$

atunci:

$$D^2(X) = p_1(x_1 - m)^2 + p_2(x_2 - m)^2 + \dots + p_n(x_n - m)^2 =$$
$$= p_1(x_1^2 - 2mx_1 + m^2) + p_2(x_2^2 - 2mx_2 + m^2) + \dots$$
$$\dots + p_n(x_n^2 - 2mx_n + m^2) = (p_1x_1^2 + p_2x_2^2 + \dots + p_nx_n^2) -$$
$$- 2m(p_1x_1 + p_2x_2 + \dots + p_nx_n) +$$
$$+ m^2(p_1 + p_2 + \dots + p_n) = M(X^2) - 2m^2 + m^2 = M(X^2) - [M(X)]^2.$$

Să scoatem în evidență cîteva proprietăți mai însemnate ale dispersiei.

a) *Dispersia unei constante este nulă*

$$D^2(a) = 0.$$

Este adevărată și reciproca acestei afirmații.

b) *Două variabile care diferă printr-o constantă au dispersii egale.*

Să considerăm variabila

$$X\begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}$$

3. În condițiile problemelor 1 și 2 să se facă tabloul pentru $S + 1$ și pentru $2N$.

4. Se dă variabila aleatoare

$$X \begin{pmatrix} 0 & 1 \\ 0{,}3 & 0{,}7 \end{pmatrix}.$$

Să se facă tabloul pentru X^n $(n \in N)$.

5. Se dau variabilele

$$X \begin{pmatrix} -1 & 0 & 1 \\ 0{,}3 & 0{,}3 & 0{,}4 \end{pmatrix}; \qquad Y \begin{pmatrix} -1 & 1 \\ 0{,}5 & 0{,}5 \end{pmatrix}.$$

Să se facă tabloul de distribuție pentru X^n și Y^n $(n \in N)$.

5. Distribuția variabilei X este

$$X \begin{pmatrix} 1 & 2 & 3 & 4 \\ p & \frac{7}{4}p & \frac{1}{3} & \frac{1}{6} \end{pmatrix}.$$

Care este probabilitatea ca X să ia o valoare $\leqslant 3$?

7. Să se scrie tabloul de distribuție al sumei variabilelor aleatoare independente

$$X \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}, \quad Y \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}.$$

8. Dacă variabilele de la problema 5 sînt independente, ce distribuție are suma pătratelor lor ?

9. Ce distribuție are suma variabilelor aleatoare independente

$$X \begin{pmatrix} -1 & 0 & 1 \\ p^2 & \frac{5}{3}p & \frac{1}{3} \end{pmatrix}, \quad Y \begin{pmatrix} -1 & 0 & 1 & 2 \\ q^2 & \frac{8}{5}q & \frac{1}{6} & \frac{1}{30} \end{pmatrix}.$$

Dar produsul lor ?

10. Se dau variabilelor aleatoare independente

$$X \begin{pmatrix} -1 & 0 & 1 \\ p + \frac{1}{6} & q + \frac{1}{3} & \frac{1}{3} \end{pmatrix}, \quad Y \begin{pmatrix} a & 0 & 1 \\ \frac{1}{3} & 2p - q & 12p^2 \end{pmatrix}, \; (a \neq 0 \,;\, a \neq 1)$$

Să se scrie distribuția variabilei $X + Y$. Pentru ce valoare a lui a avem

$$P(X + Y = 0) > \frac{2}{9} \,?$$

11. Variabila X are distribuția

$$X \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}.$$

Să se scrie repartițiile variabilelor $X + X^2$ și $X + X^3$.

12. Să se calculeze valoarea medie a variabilelor care apar în problemele 1, 2, 3, 4, 5, 6, 7, 8, 9.

13. Variabila X are distribuția

$$X \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0{,}3 & 0{,}4 & 0{,}2 & 0{,}1 \end{pmatrix}.$$

Să se calculeze $[M(X)]^2$, $M(X^2)$, $M(X - 1)$, $M(X^2 - 2X)$.

14. Să se calculeze valoarea medie a numărului de puncte ce se obțin la aruncarea a 2 zaruri. Să se rezolve problema folosind și rezultatul problemei 1.

15. Să se calculeze valoarea medie a numărului total de puncte care apar la aruncarea a 6 zaruri.

16. Să se calculeze valoarea medie a variabilelor :

$$X \begin{pmatrix} 0 & 2 & 3 & 4 & \dots & n \\ \frac{1}{2^{n-1}} & \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \dots & \frac{1}{2^{n-1}} \end{pmatrix}.$$

$$Y \begin{pmatrix} \frac{1}{1 \cdot 2} & \frac{1}{2 \cdot 3} & \dots & \frac{1}{n(n+1)} \\ \frac{1}{n} & \frac{1}{n} & \dots & \frac{1}{n} \end{pmatrix}.$$

17. Să se calculeze dispersiile variabilelor care apar la problemele 1, 2, 3, 4, 5, 6, 7.

18. Fie X o variabilă aleatoare cu media m și dispersia σ^2. Să se calculeze valoarea medie și dispersia variabilei

$$Y = \frac{X - m}{\sigma}.$$

19. Să se calculeze dispersia numărului total de puncte care apar la aruncarea a șase zaruri.

20. Se consideră variabila

$$X \begin{pmatrix} 1 & 2 & 3 & 4 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{6} & \frac{1}{12} \end{pmatrix}.$$

Exemple

a) Dacă ne interesează rezultatele obținute, la teză la matematică, de elevii din clasa a XII-a a unui liceu, atunci :

— mulțimea tuturor elevilor din clasa a XII-a a acelui liceu formează populația statistică ;

— fiecare elev din clasa a XII-a a acestui liceu este o unitate statistică ;

— nota la teză la matematică este caracteristica studiată.

b) Dacă ne interesează numărul locuitorilor din fiecare oraș al țării la o anumită dată, atunci :

— mulțimea tuturor orașelor țării la data respectivă formează populația statistică ;

— fiecare oraș constituie o unitate statistică ;

— numărul de locuitori la data respectivă este caracteristica studiată.

c) Dacă ne interesează diametrul unor piese de același fel fabricate într-o întreprindere dată, atunci :

— mulțimea pieselor fabricate de întreprindere este populația statistică ;

— o piesă constituie o unitate statistică ;

— diametrul piesei este caracteristica studiată.

d) Dacă ne interesează distribuția unui grup de copii după culoarea ochilor și culoarea părului, atunci :

— mulțimea copiilor grupului considerat formează populația statistică ;

— fiecare copil în parte din grupul respectiv este o unitate statistică ;

— culoarea ochilor și culoarea părului sînt caracteristicile care ne interesează.

Se pot da nenumărate alte exemple de mulțimi care pot constitui obiectul unei analize statistice : distribuția unui grup de persoane după talie, vîrstă și distribuția orașelor după numărul de oameni ai muncii, distribuția cardiacilor printre fumători etc.

Din înseși exemplele date rezultă existența a două feluri de caracteristici.

O caracteristică se numește *cantitativă* dacă se poate măsura. În caz contrar, caracteristica se numește *calitativă*.

Nota la teză, numărul de locuitori, diametrul piesei, vîrsta, talia, retribuția lunară etc. sînt exemple de caracteristici cantitative. Între aceste caracteristici distingem unele care pot lua numai valori întregi (numărul de locuitori ai unui oraș, numărul de copii dintr-o familie, numărul cîștigătorilor la loto într-un oraș etc.). Aceste caracteristici se numesc *discrete* sau *discontinue*.

O caracteristică care poate lua orice valoare dintr-un interval finit sau infinit se numește *continuă*. Este cazul taliei, greutății, lungimii firului de păr la oi etc.

Culoarea părului, culoarea ochilor, sexul, profesia etc. sînt exemple de caracteristici calitative.

2. Gruparea datelor

Să presupunem că s-a măsurat înălțimea unui grup de 120 de persoane. Rezultatele obținute (înălțimea în centimetri) sînt date în tabela 1, în ordinea în care ele au apărut.

Tabela 1

176	173	161	171	174	168	178	166	169	172
181	172	163	174	173	169	172	175	158	182
186	190	173	173	169	171	176	172	188	175
162	170	176	177	171	164	162	175	176	176
170	176	178	164	174	177	180	175	175	180
174	171	175	170	179	186	177	178	169	180
188	173	172	174	183	177	176	174	181	159
183	174	179	167	165	182	176	178	171	169
168	179	177	177	181	178	184	177	173	177
162	177	173	170	176	179	170	168	174	175
173	178	185	185	171	165	167	174	175	172
179	168	171	175	165	178	172	175	166	171

Este clar că sub această formă, tabela nu ne permite să tragem prea multe concluzii cu caracter mai general. De aceea, este necesar să facem o grupare a acestor date. O primă posibilitate de grupare este aceea din tabela 2.

În cazul seriilor statistice cu o singură caracteristică, pentru obținerea cu ușurință a unor concluzii generale asupra fenomenului studiat, tabelele cu două coloane, așa cum au fost ele prezentate mai sus, sînt suficiente, dacă numărul valorilor pe care le ia caracteristica este în jur de 20.

Cînd acest număr este depășit, citirea tabelei devine greoaie, fiind prea voluminoasă. Este cazul tabelei 1. În această situație se impune o nouă grupare a datelor. Revenind la tabela 1, împărțim mulțimea valorilor caracteristicii în clase, după cum reiese din tabela 6.

Tabela 6

Clase de valori în centimetri	Nr. persoane
<160	2
160—165	7
165—170	16
170—175	37
175—180	40
180—185	11
185—190	7
	120

La aceste tabele facem convenția ca extremitatea dreaptă a fiecărei clase (cu excepția, eventual a ultimei clase) să nu aparțină clasei. Astfel, clasa 165—170 cuprinde valorile x ale caracteristicii, pentru care $165 \leqslant \leqslant x < 170$.

În cazul pe care l-am prezentat, intervalele reprezentînd clasele de valori cu aceeași lungime. De la această regulă putem excepta, eventual, clasele extreme (prima și ultima). Lungimea acestor intervale nu este impusă de vreo regulă fixă, ea fiind la îndemîna statisticianului, care caută ca împărțirea în intervale să fie cît mai judicioasă. Acest mod de realizare a tabelelor (cu clase de valori de amplitudini egale sau nu) se impune totdeauna în cazul caracteristicilor continue. În tabela 7 este dată repartiția elevilor unei școli după înălțime.

Tabela 7

Talia în centimetri	Nr. elevi
150—154	38
154—158	65
158—162	175
162—166	189
166—170	111
peste 170	62
	640

Aici lungimea intervalelor alese (amplitudinea claselor) este de 4 cm.

3. Frecvența absolută. Frecvența relativă. Frecvențe cumulate

Numărul tuturor elementelor unei populații statistice se numește efectivul total al acelei populații.

Astfel, în tabela 7, populația este mulțimea elevilor unei școli și are un efectiv total de 640 de elevi.

Se numește frecvență absolută, a unei valori x a caracteristicii, numărul de unități ale populației corespunzătoare acelei valori. De exemplu, în tabela 2, valoarea 179 cm a caracteristicii are frecvența absolută egală cu 5, iar în tabela 3, nota 5 are frecvența absolută egală cu 4.

Este clar, că suma frecvențelor absolute ale tuturor valorilor caracteristice este egală cu efectivul total al populației.

Se numește frecvență relativă (sau pe scurt, frecvență) a unei valori x a caracteristicii, raportul dintre frecvența absolută a valorii x și efectivul total al populației. Vom scrie:

$$f(x) = \frac{n_x}{n},$$

unde $f(x)$ este frecvența relativă a valorii, x, n_x este frecvența absolută a acestei valori, iar n efectivul total al populației.

Deseori, frecvența relativă este dată în procente.

În tabela 2, frecvența relativă a valorii 179 este $\frac{5}{120} = \frac{1}{24}$, a valorii 175, $\frac{10}{120} = \frac{1}{12}$ etc., iar în tabela 3, frecvența valorii 5 este $\frac{4}{40} = 10\%$ etc.

În tabelele statistice cu două coloane putem înlocui în coloana a doua frecvențele absolute prin frecvențele relative. Astfel, tabela 3, care înfățișează rezultatele la teza de matematică, de către o clasă compusă din 40 de elevi, poate fi scrisă ca în tabela 8.

Tabela 8

Nota	Frecvența
2	0,025
3	0,025
4	0,050
5	0,100
6	0,175
7	0,375
8	0,150
9	0,075
10	0,025

Deci, în cazul caracteristicilor cantitative, aceste tabele scot în evidență o corespondență între două

Noțiunile introduse în raport cu valorile individuale ale variabilei pot fi extinse și în cazul tabelelor cu clase de valori. Astfel, frecvența absolută a unei clase este numărul de unități corespunzătoare valorilor variabilei care aparțin clasei respective, iar frecvența relativă (frecvența) unei clase este raportul dintre frecvența sa absolută și efectivul total al populației. Astfel, din tabela 6 rezultă că frecvența absolută a clasei 160—165 este 7, iar frecvența sa relativă este $\frac{7}{120}$.

Se numește frecvență absolută cumulată crescătoare a unei clase suma frecvențelor absolute ale tuturor claselor care apar pînă la clasa considerată inclusiv.

Se numește frecvență absolută cumulată descrescătoare a unei clase suma frecvențelor absolute a tuturor claselor care apar de la clasa considerată inclusiv.

Completînd cu aceste frecvențe tabela 6, obținem tabela 11.

Tabela 11

Clasa de valori	Frecvența absolută	Frecvența absolută cumulată crescătoare	Frecvența absolută cumulată descrescătoare
<160	2	2	120
160—165	7	9	118
165—170	16	25	111
170—175	37	62	95
175—180	40	102	58
180—185	11	113	18
185—190	7	120	7

Cum se face citirea unei astfel de tabele ?

Frecvența absolută cumulată crescătoare a unei clase ne dă numărul de unități corespunzătoare valorilor mai mici (strict) decît limita superioară a intervalului, iar frecvența absolută cumulată descrescătoare ne dă numărul unităților corespunzătoare valorilor mai mari (sau egale) cu limita inferioară a intervalului. Astfel din tabela 11 rezultă că dintre persoanele măsurate 62 au înălțimi sub 175 cm ; 95 persoane au cel puțin 170 cm înălțime etc.

Frecvența relativă cumulată crescătoare (frecvența cumulată) a unei clase este suma frecvențelor claselor care apar pînă la clasa considerată, *inclusiv* sau, ceea ce este totuna, raportul dintre frecvența absolută cumulată crescătoare și efectivul total al populației. În mod analog se definește frecvența relativă cumulată crescătoare a unei clase.

Din tabela 11 rezultă că frecvența cumulată a clasei 175—180 este $\frac{102}{120} = 0{,}85 = 85\%$. Vom spune că 85% din persoanele măsurate au înălțimea sub 180 cm. Frecvența cumulată descrescătoare a clasei 165—170 este 0,925, deci 92,5% din persoanele considerate au înălțimea de cel puțin 165 cm.

4. Serii cronologice

Tot în cadrul seriilor statistice sînt incluse și așa-numitele *serii cronologice* care prezintă evoluția în timp a unor mărimi.

În cazul unei tabele corespunzătoare unei serii cronologice, în prima coloană sînt trecute anumite momente sau intervale de timp, iar în coloana a doua valorile corespunzătoare ale mărimii considerate. Astfel să presupunem că se măsoară temperatura apei unui lac într-un anumit punct, în fiecare an, la 15 iulie ora 12, iar rezultatele obținute sînt trecute în tabela 12.

În coloana stîngă avem trecute momente precise de timp, spre deosebire de tabela 13, unde, în coloana stîngă, avem trecute intervale de timp.

Tabela 12

Data	Temperatura în °C
15 iulie 1960	20
15 iulie 1961	22
15 iulie 1962	19
15 iulie 1963	18
15 iulie 1964	20
15 iulie 1965	20
15 iulie 1966	21
15 iulie 1967	20

Tabela 13

NUMĂRUL ABSOLVENȚILOR UNEI ȘCOLI

1950—1954	1 000
1955—1959	997
1960—1964	1 002
1964—1968	1 018

(În tabelele 12 și 13 datele sînt imaginare).

§ 2. Reprezentarea grafică a seriilor statistice

În acest paragraf ne vom ocupa de reprezentarea grafică a seriilor statistice cu o singură caracteristică. Reprezentarea grafică a unei serii este uneori foarte sugestivă, ea contribuind la o primă interpretare intuitivă, pe cale vizuală, a datelor. Deseori reprezentarea grafică sugerează însăși legea pe care o urmează fenomenul studiat.

1. Reprezentarea grafică a seriilor cu caracteristică calitativă

Reprezentarea acestor distribuții constituie un capitol deosebit de important al reprezentării grafice, dat fiind că ilustrează, prin desene, anumite rapoarte numerice. Graficul corespunzător poartă numele de *diagramă.*

Să considerăm de exemplu distribuția investițiilor pe ramuri ale economiei naționale în anii 1975—1979 (tab. 14).

Tabela 14

Total investiții	100 %
Industrie	59,6%
Agricultură	11,2%
Transporturi și telecomunicații	11,3%
Gospodărie comunală și locuințe	10,7%
Învățămînt cultură și sănătate	3,0%
Celelalte ramuri	4,2%

Datele pot fi reprezentate prin dreptunghiuri de baze egale și cu înălțimi proporționale cu procentele sau prin sectoare de cerc, cu unghiurile proporționale cu aceleași numere (fig. 1).

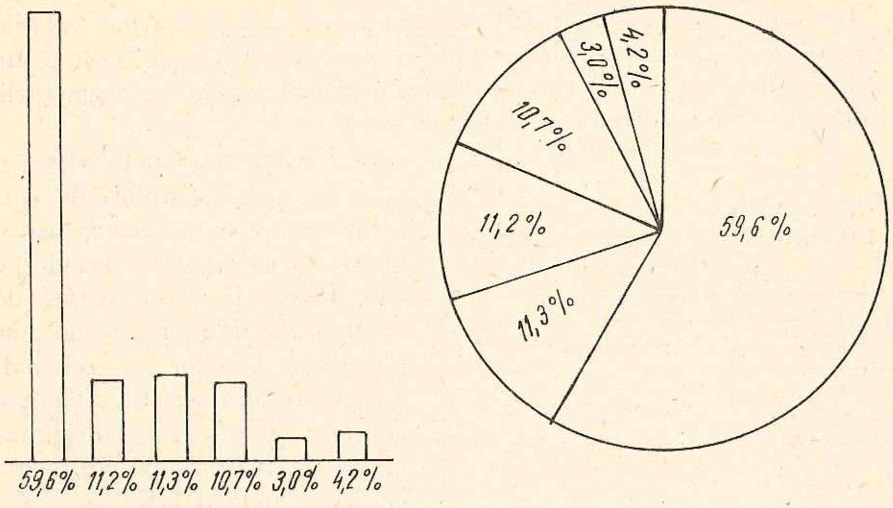

Fig. 1

2. Reprezentarea seriilor cu caracteristică cantitativă

Seriile cu caracteristică cantitativă se reprezintă grafic în raport cu un sistem de axe rectangulare. Alegerea unității pe fiecare dintre axe este la îndemîna statisticianului, care are grijă ca alegerea să ușureze obținerea concluziilor dorite, cît și ca desenul să rămînă în cadrul hîrtiei.

a) *Reprezentarea în batoane.* Această reprezentare se folosește mai ales pentru seriile în care variabila ia un număr mai mic de valori.

Să considerăm datele din tabela 15.

Obținem reprezentarea în batoane din figura 2.

Tabela 15

DISTRIBUȚIA FAMILIILOR DINTR-UN BLOC DUPĂ NUMĂRUL COPIILOR

Nr. de copii	Frecvența absolută
0	6
1	18
2	23
3	20
4	14
5	6
6	2
7	1

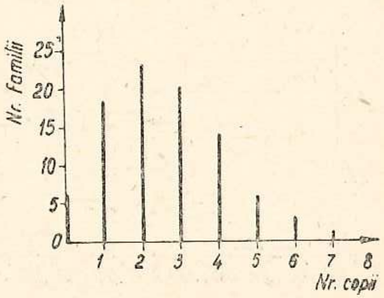

Fig. 2

Deci, pe axa orizontală sînt trecute punctele reprezentînd valorile variabilei şi din aceste puncte se ridică segmente verticale de lungime egală cu frecvenţa absolută sau relativă a valorii respective. Segmentele ridicate sînt măsurate cu unitatea de pe Oy.

b) *Histogramă.* Fiind dată o serie cu clase de valori de amplitudini egale obţinem histograma acestei serii, luînd pe axa orizontală o succesiune de segmente egale reprezentînd amplitudinea claselor, şi ridicînd, pe fiecare din aceste segmente considerate ca baze, dreptunghiuri de înălţimi proporţionale cu frecvenţele (relative sau absolute) ale claselor respective.

Tabela 16

DISTRIBUŢIA UNOR PIESE DUPĂ DIAMETRUL LOR

Mărimea diametrului mm	Frecvenţa absolută	Frecvenţa cumulată
10—20	10	10
20—30	15	25
30—40	12	37
40—50	15	52
50—60	8	60

Histograma corespunzătoare din tabela 16 este dată în figura 3. Histograma corespunzătoare tabelei 11 este dată în figura 4.

c) Dacă din mijlocul fiecărui segment de pe axa orizontală ridicăm segmente proporţionale cu frecvenţele claselor corespunzătoare fiecărui segment şi unim printr-o linie poligonală extremităţile superioare ale acestor segmente obţinem poligonul frecvenţelor. Astfel, poligonul frecvenţelor corespunzătoare tabelei 16 este dat în figura 5.

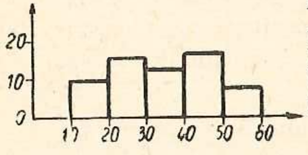

Fig. 3

d) Dacă aceleaşi puncte de la alineatul precedent le unim, nu printr-o linie poligonală, ci

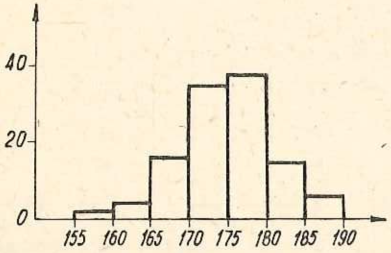

Fig. 4

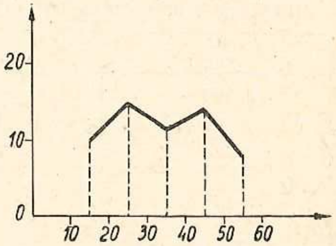

Fig. 5

printr-o curbă, obţinem curba de distribuţie a seriei respective.

Poligonul frecvenţelor cumulate (crescătoare) se obţine unind printr-o linie poligonală punctele (x, y), unde x este extremitatea dreaptă a intervalului unei clase, iar y frecvenţa cumulată a clasei respective, la care mai adăugăm punctul $(a, 0)$, unde a este limita inferioară a primei clase.

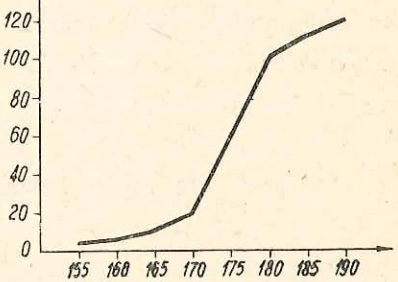

Fig. 6

Astfel, poligonul frecvenţelor cumulate corespunzător tabelei 11 este dat în figura 6.

În mod analog se defineşte curba frecvenţelor cumulate descrescătoare. Dacă punctele din figură le unim, nu printr-o linie poligonală, ci printr-o curbă, obţinem curba cumulativă a seriei considerate.

3. Reprezentarea seriilor cronologice

Să presupunem că într-un internat se dau zilnic note fiecărui dormitor pentru întreţinerea curăţeniei. Rezultatele obţinute de un anumit dormitor într-o săptămînă sînt date în tabela 17.

Diagrama corespunzătoare acestei tabele este dată în figura 7.

În acelaşi mod se fac graficele realizărilor zilnice (săptămînale, lunare) — individuale sau colective — într-o întreprindere. Abscisele punctelor care se unesc prin linia poligonală sînt mijloacele segmentelor

Tabela 17

L	10
M	9
M	9
J	10
V	8
S	10
D	9

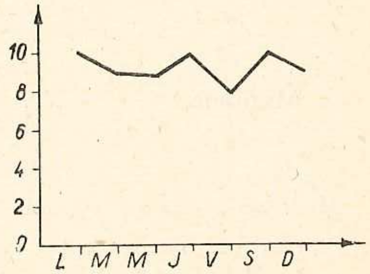

Fig. 7

reprezentînd intervale de timp (sau punctele reprezentînd momentele), iar ordonatele, valorile mărimii considerate, în intervalul de timp corespunzător.

§ 3. Elemente caracteristice ale unei serii statistice

În cele ce urmează *vom numi valoarea centrală a unei clase de variație, media aritmetică a extremităților acestei clase.* Astfel, valoarea centrală a clasei 165—170 din tabela 11 este 167,5.

1. Modul

Modul sau dominanta unei serii statistice se numește valoarea caracteristicii corespunzătoare celei mai mari frecvențe, în cazul cînd valorile caracteristicii sînt date individual, și valoarea centrală a clasei corespunzătoare celei mai mari frecvențe, în cazul variabilelor continue, cînd se dau clase de variație. Această noțiune prezintă interes mai ales în cazurile cînd avem o singură dominantă.

În cazul prezentat în tabela 3, dominanta este 7, iar în cazul tabelei 11 este 177,5.

2. Mediana

Mediana unei serii este un număr x astfel că există tot atîtea unități statistice corespunzătoare valorilor $<x$, ca și cele corespunzătoare valorilor $> x$.

Dacă o caracteristică ia valorile

$$1,\ 3,\ 3,\ 3,\ 4,\ 5,\ 6,\ 6,\ 7,\ 8,\ 9,$$

atunci 5 este mediana, deoarece există 5 valori <5 și 5 valori > 5. Dacă avem valorile

$$1,\ 3,\ 3,\ 3,\ 4,\ 5,\ 6,\ 7,\ 7,\ 9,$$

atunci vom lua ca mediană media aritmetică a numerelor situate la mijloc (dacă ele au fost în ordinea mărimii). În acest caz, mediana este 4,5. Uneori se consideră ca mediană oricare din cele două numere.

Cum se calculează mediana în cazul unei variabile continue, vom arăta pe un exemplu. Să considerăm pentru aceasta tabela 16.

Dacă piesele ar fi aranjate în ordinea diametrelor lor, noi vrem să calculăm diametrul celei de-a 30-a. Diametrul acestei piese este cuprins între 30 și 40 mm. Clasa 30—40 are frecvența absolută 12. Vom presupune că diametrul celor 12 piese corespunzătoare crește uniform de la 30 la 40. Deci creșterea diametrului de la o piesă la următoarea este $\frac{40-30}{12}$. Pe de altă parte, a 30-a piesă a populației este a $30 - 25 =$ $= $ a 5-a piesă a clasei (deoarece există 25 de piese cu diametrul <30). Deci, diametrul celei de-a 30-a piesă este de:

$$30 + (30 - 25) \times \frac{40 - 30}{12} = 34,17 \text{ mm}.$$

3. Media aritmetică

Dacă $x_1, x_2, \ldots, x_n$, sînt n valori, se știe că *media lor aritmetică este*

$$\frac{x_1 + x_2 + \ldots + x_n}{n}.$$

Valoarea	Frecvența
x_1	y_1
x_2	y_2
.	.
.	.
.	.
x_n	y_n

Fiind dată distribuția unei variabile x valoarea medie a variabilei respective este

$$x = \frac{x_1y_1 + x_2y_2 + \ldots + x_ny_n}{y_1 + y_2 + \ldots + y_n}. \qquad (1)$$

Dacă $N = y_1 + y_2 + \ldots + y_n$ este efectivul total al populației, atunci

$$\bar{x} = x_1 \frac{y_1}{N} + x_2 \frac{y_2}{N} + \ldots + x_n \frac{y_n}{N}$$

sau dacă notăm cu $f_i = \frac{y_i}{N}$ frecvența relativă a valorii x_i $(i = 1, 2, \ldots, n)$,

$$\bar{x} = x_1 f_1 + x_2 f_2 + \ldots + x_n f_n.$$

Expresia (1) are într-adevăr semnificația unei medii aritmetice. Variabila x ia, după cum reiese din tabelă, de y_1 ori valoarea x_1, de y_2 ori valoarea x_2 ş.a.m.d. Deci, pentru a calcula valoarea medie a variabilei, calculăm media aritmetică a numerelor

$$\underbrace{x_1, x_1, \ldots, x_1}_{y_1 \text{ ori}}, \quad \underbrace{x_2, x_2, \ldots, x_2}_{y_2 \text{ ori}}, \quad \ldots, \quad \underbrace{x_n, x_n, \ldots, x_n}_{y_n \text{ ori}}$$

și obținem chiar expresia din membrul drept al relației (1). Această expresie se mai numește media aritmetică ponderată a numerelor $x_1, x_2, \ldots, x_n$, numerele $y_1, y_2, \ldots, y_n$ fiind ponderile respective ale acestor valori.

Dacă vom considera datele tabelei 3, rezultă că media pe întreaga clasă la notele la matematică este

$$\frac{10 \times 1 + 9 \times 3 + 8 \times 6 + 7 \times 15 + 6 \times 7 + 5 \times 4 + 4 \times 2 + 3 \times 1 + 2 \times 1}{40} = 6{,}625.$$

Cazul seriilor cu variație continuă îl reducem la cazul precedent, substituind fiecare clasă cu valoarea sa centrală.

Astfel în exemplul prezentat în tabela 16 se obțin datele din tabela 18.

Tabela 18

Mărimea diametrului	Frecvența absolută y_i	Valoarea centrală x_i	$x_i y_i$
10–20	10	15	150
20–30	15	24	375
30–40	12	35	420
40–50	15	45	675
50–60	8	55	440
	60		2 060

$$\bar{x} = \frac{2\,060}{60} = 34{,}3.$$

Dacă vrem să lucrăm cu numere mai mici decît cele ce ne sînt date în tabele, facem următoarele observații. Avem, pentru orice i:

$$x_i = x_0 + (x_i - x_0),$$

$$x_i y_i = x_0 y_i + (x_i - x_0) y_i.$$

Dînd lui i valorile $1, 2, \ldots, n$, obținem n relații care adunate, termen cu termen, ne dau

$$x_1 y_1 + x_2 y_2 + \ldots + x_n y_n = x_0 (y_1 + y_2 + \ldots + y_n) +$$
$$+ (x_1 - x_0) y_1 + (x_2 - x_0) y_2 + \ldots + (x_n - x_0) y_n$$

sau, împărțind cu $N = y_1 + y_2 + \ldots + y_n$,

$$\bar{x} = x_0 + \frac{(x_1 - x_0) y_1 + (x_2 - x_0) y_2 + \ldots + (x_n - x_0) y_n}{N}.$$

Această relație mai poate fi scrisă:

$$\bar{x} = x_0 + \overline{x - x_0}.$$

În exemplul precedent, luînd $x_0 = 35$, avem calculele prezentate în tabela 19.

Tabela 19

Clase	Frecvența absolută y_i	Val. centrală x_i	$x_i - x_0$	$(x_i - x_0) y_i$
10–20	10	15	−20	−200
20–30	15	25	−10	−150
30–40	12	35	0	0
40–50	15	45	10	150
50–60	8	55	20	160
	60			−40

$$\bar{x} = 35 - \frac{40}{60} = 34{,}3.$$

4. Dispersia

Fiind date n valori $x_1, x_2, \ldots, x_n$ a căror medie este $\bar{x}$, se numește dispersia acestor valori, mărimea

$$\sigma^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{n}.$$

Fiind dată seria statistică:

Valori x_i	Frecvența absolută y_i
x_1	y_1
x_2	y_2
.	.
.	.
.	.
x_n	y_n

unde $N = y_1 + y_2 + \ldots + y_n$ și $\bar{x} = \frac{x_1y_1 + x_2y_2 + \ldots + x_ny_n}{N}$ sînt, respectiv, efectivul total al populației și valoarea medie. Dispersia corespunzătoare este

$$\sigma^2 = \frac{(x_1 - \bar{x})^2y_1 + (x_2 - \bar{x})^2y_2 + \ldots + (x_n - \bar{x}_n)^2y_n}{N}.$$

Mărimea

$$\sigma = \sqrt{\sigma^2}$$

se numește *abaterea* medie pătratică. Ea se exprimă în aceleași unități ca și caracteristica seriei.

În cazul caracteristicilor continue se substituie fiecare interval de variație prin valoarea sa centrală.

Să dăm și o altă formă dispersiei. Vom dezvolta expresia

$$\frac{1}{N}[(x_1 - \bar{x})^2y_1 = (x_2 - \bar{x})^2y_2 + \ldots + (x_n - \bar{x})^2y_n]$$

și obținem

$$\sigma^2 = \frac{1}{N}[x_1^2y_1 + x_2^2y_2 + \ldots + x_n^2y_n - 2\bar{x}(x_1y_1 + x_2y_2 + \ldots$$

$$\ldots + x_ny_n + \bar{x}^2(y_1 + y_2 + \ldots + y_n)] = \frac{x_1^2y_1 + x_2^2y_2 + \ldots + x_n^2y_n}{N} -$$

$$- 2\bar{x}^2 + \bar{x}^2 = \frac{x_1^2y_1 + x_2^2y_2 + \ldots + x_n^2y_n}{N} - \bar{x}^2. \qquad (2)$$

Deci:

$$\sigma^2 = \overline{x^2} - (\bar{x})^2.$$

Dacă am fi înlocuit mărimile x_i prin $x_i - x_0$ unde x_0 este o constantă, am fi obținut

$$\sigma^2 = \frac{(x_1 - x_0)^2y_1 + (x_2 - x_0)^2y_2 + \ldots + (x_n - x_0)^2y_n}{N} - (\bar{x} - x_0)^2, \qquad (2')$$

unde prin $\overline{x^2}$ am notat media aritmetică a mărimilor $x_1^2, x_2^2, \ldots, x_n^2$ cu ponderile $y_i, y_2, \ldots, y_n$. Am regăsit o proprietate a dispersiei unei variabile aleatoare.

Dispersia, sau mai bine zis, abaterea medie pătratică indică gradul de împrăștiere a valorilor în jurul valorii medii. O valoare mică a abaterii indică o pronunțată grupare a valorilor în jurul mediei aritmetice.

Să completăm tabela 19 cu datele necesare calculului dispersiei σ^2. Vom calcula σ^2 atît direct cît și folosind formula (2').

Luăm $x_0 = 35$ și știm că $\bar{x} = 34{,}3$. Se obțin datele din tabela 20.

Tabela 20

Clase	Frecv. y_i	Val. centr. x_i	$x_i - x_0$	$x_i - \bar{x}$	$(x_i - x_0)^2$	$(x_i - \bar{x})^2$	$y_i(x_i - x_0)^2$	$y_i(x_i - \bar{x})^2$
10—20	10	15	—20	—19,3	400	372,49	4 000	3 724,90
20—30	15	25	—10	— 9,3	100	86,49	1 500	1 297,35
30—40	12	35	0	0,7	0	0,49	0	5,88
40—50	15	45	10	10,7	100	114,49	1 500	1 717,35
50—60	8	55	20	20,7	400	428,49	3 200	3 427,92
	60						10 200	10 173,40

Deci:

$$\sigma^2 = \frac{10\ 173{,}4}{60} = 169{,}5$$

sau folosind (2')

$$\sigma^2 = \frac{1}{60} \times 10\ 200 - (34{,}3 - 35)^2 = 169{,}5.$$

Se observă că, alegînd convenabil valoarea lui x_0, calculele se simplifică.

§ 4. Sondaje

1. Generalități

Am văzut că în statistică întîlnim diverse populații, alcătuite dintr-un număr mare de unități, care pot fi persoane, obiecte, informații etc. Studiul direct al populațiilor statistice este de multe ori greu de realizat din cauza numărului mare de unități. Un asemenea studiu poate fi prea costisitor și să pretindă prea mult timp execuția lui. Alteori, dacă numărul unităților nu este determinat, ca de exemplu numărul pieselor pe care le poate face o mașină, populația totală nu poate fi evaluată. În toate aceste cazuri, pentru a culege informații privitoare la populația considerată, efectuăm o statistică numai pentru o fracțiune din populația totală și rezultatul obținut îl extindem pentru toată populația. Spunem că am executat un *sondaj*, iar fracțiunea din populația totală pentru care am făcut statistica poartă numele de *eșantion*.

De exemplu, la o policlinică s-au prezentat într-o lună 23 000 de persoane. Vă interesează distribuția pe vîrste a acestor bolnavi. Pentru aceasta, din fișele întocmite pentru fiecare bolnav în parte, se aleg la întîmplare 1 000 de fișe. Efectuăm statistica pe vîrste pentru bolnavii corespunzători acestor 1 000 de fișe. Am făcut astfel de sondaj în baza unui eșantion de 1 000 de unități, dintr-o populație totală de 23 000 de unități. Deoarece fișele au fost scoase la întîmplare, putem presupune că modelul matematic pentru operația pe care am făcut-o este dat de o urnă în care se găsesc 23 000 de bile și extragem la întîmplare 1 000 de bile, adică un eșantion de 1 000 de bile. Cunoscînd statistica referitoare la eșantion, ne propunem să căpătăm informații privind populația totală. Evident, informațiile le putem căpăta cu o anumită probabilitate.

Sondajele sînt mult folosite în practica statisticii. Unele dintre ele, numite sondaje de opinie, sînt făcute cu scopul de a afla părerea unor oameni în vederea alcătuirii unui program, unei lucrări etc. De exemplu, în vederea îmbunătățirii programelor de radio, se poate cere părerea abonaților pe baza unui sondaj de opinie, trimițîndu-le un chestionar pentru completare.

Avem două cazuri, după cum aplicăm schema lui Bernoulli sau schema hipergeometrică.

2. Schema lui Bernoulli

Să considerăm mai întîi urna lui Bernoulli, pentru care probabilitatea de a scoate o bilă albă este p. Efectuăm n extracții succesive din urnă, punînd de fiecare dată bila extrasă înapoi în urnă. Fie

$$f_n = \frac{\alpha}{n},$$

frecvența numărului de bile albe obținute în n extracții. Ne interesează probabilitatea dublei inegalități

$$np - k\sqrt{npq} \leqslant \alpha \leqslant np + k\sqrt{npq}, \quad q = 1 - p$$

sau

$$f_n \in \left[p - k\sqrt{\frac{pq}{n}},\ p + k\sqrt{\frac{pq}{n}}\right]$$

unde k este un număr real.
Această probabilitate este

$$P\left[f_n \in \left(p - k\sqrt{\frac{pq}{n}},\ p + k\sqrt{\frac{pq}{n}}\right)\right] = \sum_{\alpha=[np - k\sqrt{npq}]}^{\alpha=[np + k\sqrt{npq}]} C_n^{\alpha} p^{\alpha} q^{n-\alpha},$$

unde $[a]$ este cel mai mare număr întreg cuprins în a.

Probabilitatea de mai sus a fost calculată pentru diferitele valori date lui p, k, n, s-au întocmit tabele numerice în acest scop și s-a constatat că dacă este îndeplinită condiția

$$npq > 9$$

se obțin următoarele rezultate numerice:

$$P\left[f_n \in \left[p - 1{,}96\sqrt{\frac{pq}{n}},\ p + 1{,}96\sqrt{\frac{pq}{n}}\right]\right] = 0{,}950;$$

3) $$P\left[f_n \in \left[p - 2{,}58\sqrt{\frac{pq}{n}},\ p + 2{,}58\sqrt{\frac{pq}{n}}\right]\right] = 0{,}990;$$

$$P\left[f_n \in \left[p - 3\sqrt{\frac{pq}{n}},\ p + 3\sqrt{\frac{pq}{n}}\right]\right] = 0{,}997.$$

Se observă că probabilitățile din membrul al doilea sînt foarte mari. Deci, în afară de rare excepții, frecvența f_n este cuprinsă în intervalul

$$\left[p - k\sqrt{\frac{pq}{n}},\ p + k\sqrt{\frac{pq}{n}}\right],\ k \geqslant 1{,}96.$$

Vom da cîteva aplicații ale formulelor (3).

A p l i c a ț i a 1. *La o mașină, din 1 000 de piese 25 sînt rebut. Să se determine cu probabilitatea egală cu 0,95 un interval în care se găsește numărul de piese bune din 5 000 de piese fabricate.*

Sîntem în cazul schemei lui Bernoulli.

Aplicînd prima formulă din (3) avem

$$P\left[f_{5\,000} \in \left(0{,}975 - 1{,}96\sqrt{\frac{0{,}975 \times 0{,}025}{5\,000}},\right.\right.$$

$$\left.\left. 0{,}975 + 1{,}96\sqrt{\frac{0{,}975 \times 0{,}025}{5\,000}}\right)\right] = 0{,}95,$$

deoarece $p = 0{,}975$, $q = 0{,}025$, $n = 5\,000$.

Efectuînd calculele obținem:

$$P[f_{5\,000} \in (0{,}9707\ ;\ 0{,}9793)] = 0{,}95.$$

Numărul pieselor bune se obține înmulțind frecvența cu 5 000 și obținem:

$$P[5\,000 f_{5\,000} \in (4\,853\ ;\ 4\,897)] = 0{,}95.$$

Cu o probabilitate egală cu 0,95 vom găsi un număr de piese bune cuprins între 4 853 și 4 897.

A p l i c a ț i a 2. 51% *dintre copiii nou-născuți sînt băieți și* 49% *fete. Să se determine, cu o probabilitate egală cu* 0,99, *între ce limite variază numărul băieților la* 10 000 *copii născuți.*

Sîntem în cazul urnei lui Bernoulli cu $n = 10\,000$, $p = 0{,}51$, $q = 0{,}49$.

Condiția $npq > 9$ este îndeplinită.

Deci:

$$P\left[f_{10\,000} \in \left(0{,}51 - 2{,}58\sqrt{\frac{0{,}51 \times 0{,}49}{10\,000}},\ 0{,}51 + 2{,}58\sqrt{\frac{0{,}51 \times 0{,}49}{10\,000}}\right)\right] = 0{,}99;$$

$$P[f_{10\,000} \in (0{,}4971\ ;\ 0{,}5229]) = 0{,}99.$$

Deci cu o probabilitate egală cu 0,99, la 10 000 copii nou-născuți numărul băieților va fi cuprins între 4 971 și 5 229. Dacă o statistică efectuată într-o anumită regiune desminte acest fapt, înseamnă că în acea regiune nu putem admite că probabilitatea ca un nou-născut să fie băiat este egală cu 0,51.

A p l i c a ț i a 3. *S-a aruncat un zar de* 600 *de ori și s-a obținut fața 1 de* 70 *de ori. Se poate admite că zarul a fost just, adică probabilitatea de a obține o față dată este* 1/6 *?*

Sîntem în cazul urnei lui Bernoulli cu $p = \frac{1}{6}$, $q = \frac{5}{6}$, $n = 600$.

Construim un interval cu probabilitatea 0,95.

$$P\left[f_{600} \in \left[\frac{1}{6} - 1{,}96\sqrt{\frac{1}{600}\cdot\frac{1}{6}\cdot\frac{5}{6}},\ \frac{1}{6} + 1{,}96\sqrt{\frac{1}{600}\cdot\frac{1}{6}\cdot\frac{5}{6}}\right]\right] = 0{,}95.$$

$$P[f_{600} \in [0{,}1368\ ;\ 0{,}1965]] = 0{,}95.$$

Numărul $\frac{70}{600}$ se găsește în afara intervalului (0,137 ; 0,197). Probabilitatea ca să obținem un număr în afara intervalului este 0,05. Această probabilitate fiind mică, conchidem că nu putem admite ipoteza $p = \frac{1}{6}$, adică zarul nu e just. În aplicațiile teoriei probabilităților se admite că trebuie să considerăm ca anormale evenimentele care se produc rar, cu o probabilitate mică.

3. *Schema hipergeometrică*. Se pot stabili intervale care cuprind cu mare probabilitate și frecvența corespunzătoare schemei hipergeometrice. Să considerăm astfel o urnă, în care avem a bile albe și b bile negre

$\left(a + b = N,\ p = \frac{a}{N},\ q = \frac{b}{N}\right)$. Extragem din urnă $n \leqslant N$, bile fără ca să punem bila extrasă înapoi în urnă. Fie

$$f_n = \frac{\alpha}{n},$$

frecvența numărului de bile albe obținute în n extracții.

Avem :

Calculîndu-se membrul al doilea al acestei egalități pentru diverse valori ale lui p, n, k, s-a constatat că pentru

$$npq > 9$$

avem formulele :

$$P\left[f_n \in \left[p - 1{,}96\sqrt{\frac{pq}{n}\,\frac{N-n}{N-1}};\ p + 1{,}96\sqrt{\frac{pq}{n}\,\frac{N-n}{N-1}}\right]\right] = 0{,}950;$$

$$P\left[f_n \in \left[p - 2{,}58\sqrt{\frac{pq}{n}\,\frac{N-n}{N-1}};\ p + 2{,}58\sqrt{\frac{pq}{n}\,\frac{N-n}{N-1}}\right]\right] = 0{,}990; \tag{4}$$

$$P\left[f_n \in \left[p - 3\sqrt{\frac{pq}{n}\,\frac{N-n}{N-1}};\ p + 3\sqrt{\frac{pq}{n}\,\frac{N-n}{N-1}}\right]\right] = 0{,}997.$$

În intervalele de forma

$$\left[p - k\sqrt{\frac{pq}{n}\,\frac{N-n}{N-1}},\ p + k\sqrt{\frac{pq}{n}\,\frac{N-n}{N-1}}\right],\ k \geqslant 1{,}96,\ npq > 9,$$

frecvența f_n este cuprinsă cu probabilități foarte mari, conform formulelor (4). Pentru alte valori decît $k = 1{,}96$, $k = 2{,}58$, $k = 3$, au fost întocmite tabele numerice, pe care noi însă nu le folosim, pentru exemplele din această carte, fiind suficiente aceste trei valori date lui k.

A p l i c a ț i e. *Într-un oraș cu 1 756 000 de locuitori sînt 157 320 mai vîrstnici de 65 ani. Cîți locuitori depășesc vîrsta de 65 de ani într-un sector al orașului cu 420 000 de locuitori ?*

Locuitorii din sector reprezintă un eșantion din numărul total al locuitorilor din oraș. Sîntem în cazul schemei hipergeometrice cu

$$N = 1\,756\,000,\ n = 420\,000,\ p = \frac{157\,320}{1\,756\,000} = 0{,}0\,896,\ q = 0{,}9104.$$

Vom aplica formulele (4) cu o probabilitate egală cu 0,99 :

$$P\left[f_{420\,000} \in \left[0{,}0896 - 2{,}58\sqrt{\frac{0{,}0896 \times 0{,}9104}{420\,000} \cdot \frac{1\,756\,000 - 420\,000}{1\,756\,000 - 1}};\right.\right.$$

$$\left.\left.0{,}896 + 2{,}58\sqrt{\frac{0{,}0896 \times 0{,}9104}{420\,000} \cdot \frac{1\,756\,000 - 420\,000}{1\,756\,000 - 1}}\right]\right] = 0{,}99.$$

$$P[f_{420\,000} \in [0{,}088598;\ 0{,}090582]] = 0{,}99.$$

Prin urmare, cu probabilitatea 0,99 frecvența $f_{420\,000}$ se găsește cuprinsă în intervalul [0,088598 ; 0,090582). Pentru a găsi limitele în care variază numărul locuitorilor mai vîrstnici de 65 de ani în sectorul considerat trebuie să înmulțim rezultatele obținute cu 420 000. Notînd cu M acest număr găsim

$$37\,211 < M < 38\,044.$$

Cu o probabilitate egală cu 0,99, M se găsește cuprins între 37 211 și 38 044.

4. *Intervale de încredere pentru determinarea probabilității în cazul sondajelor de volum mare.*

Inegalitățile de care ne-am ocupat pînă acum sînt de forma :

$$np - k\sqrt{npq} \leqslant \alpha \leqslant np + k\sqrt{npq}.$$

Notînd pentru simplificare

$$f_n = f,$$

inegalitățile, de mai sus pot fi scrise succesiv :

$$|p - f| \leqslant k\sqrt{\frac{pq}{n}};$$

$$|p - f|^2 \leqslant k^2\frac{pq}{n};$$

$$\left(1 + \frac{k^2}{n}\right)p^2 - 2\left(f + \frac{k^2}{2n}\right)p + f^2 \leqslant 0,$$

$$p_1 \leqslant p \leqslant p_2$$

unde p_1 şi p_2 sînt rădăcinile ecuaţiei :

$$\left(1+\frac{k^2}{n}\right)p^2-2\left(f+\frac{k^2}{2n}\right)p+f^2=0,$$

$$p_1=\frac{f+\frac{k^2}{n}-k\sqrt{\frac{f(1-f)}{n}+\frac{k^2}{4n^2}}}{1+\frac{k^2}{n}},$$

$$p_2=\frac{f+\frac{k^2}{n}+k\sqrt{\frac{f(1-f)}{n}+\frac{k^2}{4n^2}}}{1+\frac{k^2}{n}}.$$

Cum $k \leqslant 3$, pentru eşantioanele mari, adică pentru valori ale lui n suficient de mari, putem lua

$$p_1 \approx f-k\sqrt{\frac{f(1-f)}{n}},$$

$$p_2 \approx f+k\sqrt{\frac{f(1-f)}{n}}.$$

Avem deci pentru eşantioanele mari

$$p-k\sqrt{\frac{pq}{n}} \leqslant f \leqslant p+k\sqrt{\frac{pq}{n}} \Leftrightarrow f-k\sqrt{\frac{f(1-f)}{n}} \leqslant p \leqslant f+k\sqrt{\frac{f(1-f)}{n}}.$$

Formulele pot fi scrise sub forma :

$$P\left[p\in\left[f-1{,}96\sqrt{\frac{f(1-f)}{n}},\ f+1{,}96\sqrt{\frac{f(1-f)}{n}}\right]\right]=0{,}950\,;$$

$$(3')\quad P\left[p\in\left[f-2{,}58\sqrt{\frac{f(1-f)}{n}},\ f+2{,}58\sqrt{\frac{f(1-f)}{n}}\right]\right]=0{,}990\,;$$

$$P\left[p\in\left[f-3\sqrt{\frac{f(1-f)}{n}},\ f+3\sqrt{\frac{f(1-f)}{N}}\right]\right]+0{,}997,$$

pentru $k=1{,}96$; $2{,}58$; 3.

Intervalul

$$\left[f-k\sqrt{\frac{f(1-f)}{n}},\ f+k\sqrt{\frac{f(1-f)}{n}}\right],$$

poartă numele de interval de încredere. Ele variază după hazard, dar acoperă cu o probabilitate mare, cunoscută, probabilitatea teoretică p. La fel, pentru eşantioane mari, formulele (4), referitoare la schema hipergeometrică sînt echivalente cu

$$P\left[p\in\left[f-1{,}96\sqrt{\frac{f(1-f)}{n}\,\frac{N-n}{N-1}}\,;\ f+1{,}96\sqrt{\frac{f(1-f)}{n}\,\frac{N-n}{N-1}}\right]\right]=0{,}950\,;$$

$$(4')\quad P\left[p\in\left[f-2{,}58\sqrt{\frac{f(1-f)}{n}\,\frac{N-n}{N-1}}\,;\ f+2{,}58\sqrt{\frac{f(1-f)}{n}\,\frac{N-n}{N-1}}\right]\right]=0{,}990\,;$$

$$P\left[p\in\left[f-3\sqrt{\frac{f(1-f)}{n}\,\frac{N-n}{N-1}}\,;\ f+3\sqrt{\frac{f(1-f)}{n}\,\frac{N-n}{N-1}}\right]\right]=0{,}997.$$

A p l i c a ţ i a I. *Într-un oraş au fost sancţionaţi în timpul unui an* 15 300 *de conducători de automobile pentru abateri de la legile circulaţiei.*

Efectuîndu-se un sondaj, s-a constatat că din 2 000 *de conducători sancţionaţi,* 590 *au fost femei.*

Să se evalueze numărul total de femei sancţionate.

Modelul matematic pentru această problemă este dat de o urnă în care avem bile de două culori. Numărul total al bilelor este de 15 300. Efectuîndu-se din această urnă, 2 000 extracţii, fără ca să se pună bila extrasă înapoi în urnă, s-a găsit frecvenţa

$$f_{2\,000}=\frac{590}{2\,000}.$$

Sondajul s-a făcut în modul următor : s-au ales la întîmplare 2 000 de conducători sancţionaţi din totalul de 15 300, avînd grijă ca nici unul dintre ei să nu fie ales de două ori. Este acelaşi lucru, ca şi cum într-o urnă am aşezat 15 300 bileţele, în care sînt trecute numele conducătorilor sancţionaţi, şi am scoate 2 000 dintre ele. Bileţelele sînt de două feluri : unele conţin numele conducătorilor femei şi altele conducătorilor bărbaţi. Deci sîntem tocmai în cazul schemei hipergeometrice.

Aplicînd formulele (4') obținem, ținînd seama că

$$N = 15\,300,\ n = 2\,000,\ f_{2\,000} = \frac{590}{2\,000} = 0{,}295;$$

$$P\left[p \in \left(0{,}295 - 1{,}96\sqrt{\frac{0{,}295 \times 0{,}705}{2\,000} \cdot \frac{15\,300 - 2\,000}{15\,300 - 1}};\right.\right.$$

$$\left.\left.0{,}295 + 1{,}96\sqrt{\frac{0{,}295 \times 0{,}705}{2\,000} \cdot \frac{15\,300 - 2\,000}{15\,300 - 1}}\right)\right] = 0{,}95;$$

$$P\left[p \in \left((0{,}295 - 2{,}58\sqrt{\frac{0{,}295 \times 0{,}705}{2\,000} \cdot \frac{13\,300}{15\,299}};\right.\right.$$

$$0{,}295 + 2{,}58\sqrt{\frac{0{,}295 \times 0{,}705}{2\,000} \cdot \frac{13\,300}{15\,299}} = 0{,}99;$$

$$P\left[p \in \left(0{,}295 - 3\sqrt{\frac{0{,}295 \times 0{,}705}{2\,000} \cdot \frac{13\,300}{15\,299}};\right.\right.$$

$$\left.\left.0{,}295 + 3\sqrt{\frac{0{,}95 \times 0{,}705}{2\,000} \cdot \frac{13\,300}{15\,299}}\right)\right] = 0{,}997.$$

Efectuînd calculele găsim:

$$P[0{,}276 < p < 0{,}314] = 0{,}95;$$

$$P[0{,}270 < p < 0{,}320] = 0{,}99;$$

$$P[0{,}266 < p < 0{,}324] = 0{,}997.$$

Prima egalitate ne arată că probabilitatea, pe care n-o cunoaștem ca un contravenient luat la întîmplare să fie femeie, este cuprinsă în intervalul (0,276 ; 0,314). Acest fapt se produce cu probabilitatea 0,95. Intervalul (0,276 ; 0,314) este un interval de încredere cu probabilitatea 0,95.

Dacă intervalul de încredere crește este evident că și probabilitatea corespunzătoare crește. Astfel, probabilitatea evenimentului $p \in (0{,}266;\ 0{,}324)$ este 0,997. Sîntem aproape siguri în afară de 3‰ cazuri de excepție, că probabilitatea necunoscută p ca un contravenient să fie femeie este cuprinsă în intervalul (0,266 ; 0,324).

Am arătat că, notînd prin p, probabilitatea de a scoate o bilă albă la schema hipergeometrică, valoarea medie a numărului de bile albe obținute în N probe este N_p. Prin urmare, pentru a determina inter-

valul în care se găsește numărul de femei contraveniente trebuie să înmulțim rezultatele găsite pentru p, cu $N = 15\,300$.

Găsim astfel, dacă notăm prin A numărul total de femei contraveniente:

$$P[4\,228 < A < 4\,799] = 0{,}95,$$

$$P[4\,138 < A < 4\,889] = 0{,}99,$$

$$P[4\,077 < A < 4\,950] = 0{,}997.$$

Cu o siguranță desmințită în 3‰ din cazuri, putem afirma că numărul de femei dintre cei 15 300 conducători sancționați este cuprins în intervalul (4 077, 4 950).

A p l i c a ț i a II. *O mașină automată a fabricat* 5 000 *de piese, dintre care* 243 *nu sînt acceptabile. Care este probabilitatea ca mașina să fabrice o piesă rebut?*

Admitem că probabilitatea ca mașina să fabrice o piesă care să nu poată fi acceptată rămîne aceeași în timpul fabricației. Fie p această probabilitate. Fiecărei piese i se atașează o variabilă aleatoare, care cu probabilitatea p poate să ia valoarea 1, adică piesa este inacceptabilă, și cu probabilitatea $q = 1 - p$, poate lua valoarea 0, adică piesa este acceptabilă. Obținem astfel, fabricînd un număr n de piese, un șir de variabile aleatoare, care presupunem că sînt independente între ele. Se vede că, schematic, ne situăm în cazul unei urne, în care avem bile de două culori, și efectuăm extrageri succesive, punînd de fiecare dată bila extrasă înapoi în urnă. Modelul matematic al aplicației de mai sus este dat deci de schema lui Bernoulli, în care probabilitatea de realizare a evenimentului în fiecare probă este p. Putem deci aplica formulele (3').

$$\text{Avem: } n = 5\,000 \qquad f_{5\,000} = \frac{243}{5\,000};$$

$$P\left[p \in \left(\frac{243}{5\,000} - 2{,}58\sqrt{\frac{243}{5\,000}\left(1 - \frac{243}{5\,000}\right)\frac{1}{5\,000}};\right.\right.$$

$$\left.\left.\frac{243}{5\,000} + 2{,}58\sqrt{\frac{243}{5\,000}\left(1 - \frac{243}{5\,000}\right)\frac{1}{5\,000}}\right)\right] = 0{,}99;$$

$$P\left[p \in \left(\frac{204}{5\,000},\ \frac{282}{5\,000}\right)\right] = 0{,}99.$$

Cu o probabilitate egală cu 0,99, probabilitatea teoretică de a avea un rebut se găsește cuprinsă în intervalul (0,041 ; 0,056).

Probleme

1. Să se dea exemple de populații statistice și caracteristici.

2. Stabiliți repartiția elevilor clasei din care faceți parte, în raport cu numărul fraților.

3. Stabiliți repartiția elevilor clasei din care faceți parte după media la matematică în clasa a XI-a.

4. Aruncați un zar de 120 de ori. Întocmiți tabela rezultatelor obținute. Comparați frecvențele absolute ale fiecărei fețe cu numărul 20.

4. Faceți tabela distribuției elevilor clasei după media la matematică și media la fizică obținută în clasa a XI-a (tabela cu dublă intrare).

6. Completați tabelele corespunzătoare problemelor 2, 3, 4 cu frecvențele relative și cu frecvențele cumulate. În cazul problemei 4 să se spună : la cîte aruncări s-a obținut o față cu un număr $\leqslant 3$ puncte ? La cîte aruncări s-a obținut o față cu un număr $\geqslant 5$ puncte ?

7. În tabela 25 sînt trecute rezultatele obținute într-o cursă de 100 m de către 40 de participanți.

Tabela 25

Timp (s)	Nr. de alergători	Timp (s)	Nr. de alergători
10,5—10,7	1	11,5—11,7	7
10,7—10,9	1	11,7—11,9	6
10,9—11,1	2	11,9—12,1	5
11,1—11,3	5	12,1—12,3	4
11,3—11,5	7	12,3—12,5	2

Să se completeze tabela cu valorile centrale ale claselor, cu frecvențele relative și cu frecvențele cumulate. Să se citească pe tabela obținută cîți alergători au realizat timpuri mai bune de 11,7 s. Care este procentul de alergători care au realizat timpuri mai bune de 11,5 s ?

8. Să se reprezinte grafic, în batoane, seriile de la punctele 2, 3, 4.

9. Să se construiască histograma, poligonul frecvențelor și poligonul frecvențelor cumulate crescătoare pentru seria din problema 7.

10. Să se calculeze dominanta, mediana și valoarea medie (media aritmetică) pentru seriile de la problemele 2, 3, 4.

11. Să se stabilească timpul mediu realizat pe 100 m de către 40 alergători, ale căror rezultate sînt trecute în tabela 25.

12. În tabela 26 este prezentată distribuția elevilor dintr-o școală generală după talie.

Să se completeze tabela. Să se construiască histograma corespunzătoare. Să se calculeze și să se interpreteze dominanta și mediana. Să se calculeze înălțimea medie a elevilor școlii.

Tabela 26

Talia (cm)	Număr elevi
150—154	38
154—158	65
158—162	175
162—166	189
166—170	111
170—174	62
	640

13. Să se calculeze dispersiile seriilor de la problemele 2, 3, 4, 7, 12.

14. S-a aruncat o monedă de 1 000 de ori și s-a obținut o față de 576 de ori. Se poate admite că moneda a fost justă, în sensul că probabilitatea obținerii uneia din fețe este 1/2 ?

15. O mașină produce 6% rebuturi. Într-un lot de 1 000 de piese fabricate sînt 85 piese rebutate. Funcționează mașina normal ?

16. La o bibliotecă sînt, după numărul formularelor completate la înscriere, 9 700 de cititori, dintre care 2 020 sînt mai tineri de 20 de ani. Luîndu-se la întîmplare 500 formulare, între ce limite variază, cu probabilitatea 0,990, numărul cititorilor mai tineri de 20 de ani cuprinși în formulare ?

17. Într-o localitate sînt 125 372 locuințe. Dintr-un sondaj de 1 200 locuințe s-a constatat că 571 dintre ele posedă spațiu excedentar. Să se evalueze numărul total al locuințelor avînd spațiu excedentar.

18. În biblioteca unei întreprinderi au fost eliberate într-un an fișe de împrumut de cărți. Luîndu-se la întîmplare 5 000 de fișe s-a întocmit următoarea situație :

	Bărbați	Femei	Total
Lucrători	1 960	1 211	3 171
Funcționari	1 010	819	1 829
Total	2 970	2 030	5 000

Se cere să se evalueze pentru întreaga colectivitate :

1. Numărul cititoarelor.
2. Numărul lucrătorilor, care au împrumutat cărți.

Indicații și răspunsuri

Capitolul I

7. A, B incompatibile: $P(A) = \frac{[\sqrt{n}]}{n}$;

$P(B) = \frac{\left[\frac{n+1}{3}\right]}{n}$ ($[x]$ = partea întreagă a lui x).

8. $a)\ \frac{3}{8}$; $b)\ \frac{1}{8}$; $c)\ \frac{1}{4}$; $d)\ \frac{1}{4}$; $e)\ \frac{1}{2}$; $f)\ \frac{1}{2}$.

9. Calculăm mai întîi probabilitatea evenimentului contrar: ca prima fișă extrasă să nu conțină cifra 5. Numărul cazurilor favorabile este egal cu numărul sistemelor de 4 cifre ce se pot forma cu:

$$0,\ 1,\ 2,\ 3,\ 4,\ 6,\ 7,\ 8,\ 9,$$

adică $9^4 = 6\,561$ și probabilitatea căutată este $\frac{6\,561}{10\,000} = 0{,}6561$. Probabilitatea din enunț este $1 - 0{,}6561 = 0{,}3439$.

Observație. Am considerat că numărul fiecărei fișe este format din 4 cifre. Astfel fișa cu numărul 1 poartă scris numărul 0001, iar cea cu numărul 35 este notată cu 0035. În loc de numărul 0000 considerăm numărul 10 0000.

10. *Indicație.* Se calculează mai întîi probabilitatea evenimentului contrar.

Răspuns: $1 - \frac{C_{90}^5}{C_{100}^5}$.

11. Cele 7 cifre pot fi scrise în 7 ! moduri. Dacă cifrele 1 și 2 au locurile stabilite, atunci celelalte cifre se pot scrie în 5 ! moduri. Dar cifrele 1 și 2 pot ocupa locuri consecutive în ordinea 1, 2 în 6 moduri.

Răspuns: $\frac{5!6}{7!} = \frac{1}{7}$.

12. $\frac{6!}{6^6} = 0{,}015$; $1 - \frac{5^6}{6^6} = 0{,}665$.

13. **Răspuns**: $\frac{7}{8}$.

14. Fie A evenimentul ca prima bilă extrasă să fie albă și a doua bilă extrasă să fie tot albă, iar B evenimentul ca prima să fie neagră, iar a doua albă. Avem de calculat $P(A \cup B)$.

Răspuns: $\frac{\alpha}{\alpha + \beta}$.

15. $a)\ \frac{6}{35}$; $b)\ \frac{12}{35}$; $c)\ \frac{23}{35}$; $d)\ \frac{18}{35}$.

16. *a)* Fie k numărul multiplilor de a, care se găsesc printre numerele $1, 2, 3, \ldots, n$. Acești multipli sînt: $a, 2a, 3a, \ldots, ka$. Rezultă că $n = ka + r$, unde $0 < r < a$. Putem scrie $\frac{n}{a} = k + \frac{r}{a}$ și deci $\left[\frac{n}{a}\right] = k$. Pentru extragerea unui multiplu de a, avem k cazuri favorabile din n posibile. Probabilitatea căutată este $\frac{k}{n} = \frac{\left[\frac{n}{a}\right]}{n}$.

b) Calculăm mai întîi probabilitatea p a obținerii unui multiplu de a. Deoarece $n = ka$, pentru $p = \frac{k}{ka} = \frac{1}{a}$.

17. $a)\ \frac{1}{10}$; $b)\ \frac{1}{90}$.

18. Se scrie $P(A_1 \cup A_2 \cup A_3) = P((A_1 \cup A_2) \cup A_3) = P(A_1 \cup A_2) + P(A_3) - P((A_1 \cup A_2) \cap A_3)$ etc.

La fel pentru $P(A_1 \cup A_2 \cup A_3 \cup A_4)$.

19. $1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \ldots + (-1)^{n+1}\frac{1}{n!}$.

20. Se folosește formula: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ și generalizarea sa din problema 18.

Capitolul II

1. $a)\ \frac{5}{7}$; $\frac{3}{7}$; $b)\ \frac{27}{49}$; $c)\ \frac{31}{49}$.

2. *b)* $\frac{31}{70}$; *c)* $\frac{3}{14}$; *d)* $\frac{39}{70}$.

3. *a)* $\frac{1}{6}$; *b)* $\frac{10}{21}$.

6. $\frac{55}{126}$.

8. *a)* Un rezultat al experienței $\mathcal{E}$ este o pereche de numere (m, n) : m — numărul obținut la extragerea din U_1 și n — numărul obținut din U_2. Deci $\mathcal{E}$ are $7 \cdot 8 = 56$ rezultate (egal) posibile :

(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)	(1, 7)
(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)	(2, 7)
(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)	(3, 7)
(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)	(4, 7)
(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)	(5, 7)
(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)	(6, 7)
(7, 1)	(7, 2)	(7, 3)	(7, 4)	(7, 5)	(7, 6)	(7, 7)
(8, 1)	(8, 2)	(8, 3)	(8, 4)	(8, 5)	(8, 6)	(8, 7)

b) Ca eveniment legat de $\mathcal{E}_1$, A are 4 cazuri favorabile : 2, 4, 6, 8. Ca eveniment legat de $\mathcal{E}$, A are $4 \cdot 7 = 28$ cazuri favorabile (cele subliniate în tabloul de mai sus).

9. $\frac{95}{99}$.

10. $\frac{119}{120}$.

11. $\frac{5}{36} \cdot \frac{6}{36} \cdot \frac{5}{36} = \frac{25}{7\,776}$; $6 \cdot \frac{25}{7\,776} = \frac{25}{1\,296}$.

12. *a)* $\frac{3}{7} \cdot \frac{4}{9} \cdot \frac{5}{11} + \frac{4}{7} \cdot \frac{5}{9} \cdot \frac{6}{11} = \frac{20}{77}$; *b)* $\frac{3}{7} \cdot \frac{5}{9} \cdot \frac{6}{11} + \frac{4}{7} \cdot \frac{4}{9} \cdot \frac{6}{11} +$

$$+ \frac{4}{7} \cdot \frac{5}{9} \cdot \frac{5}{11} = \frac{376}{693}.$$

13. $\frac{3}{4}$.

Capitolul III

1. *Indicație. a)* Sîntem în cazul schemei binomiale generalizate. Se calculează coeficientul lui x^2 din polinomul

$$\left(\frac{19}{20}x + \frac{1}{20}\right)\left(\frac{18}{20}x + \frac{2}{20}\right)\left(\frac{17}{29}x + \frac{3}{20}\right)\left(\frac{16}{20}x + \frac{4}{20}\right).$$

b) Este suma coeficienților puterilor a 3-a și a 4-a ale lui x din polinomul de mai sus.

2. *Indicație.* Probabilitatea ca la o aruncare a zarurilor să obținem 7 este $\frac{1}{6}$.

Sîntem în cazul schemei lui Bernoulli, cu $n = 10$; $p = \frac{1}{6}$, $k = 3$.

Răspuns : $C_{10}^3 \frac{5^7}{6^{10}}$.

3. $C_8^4 \frac{1}{2^8} = \frac{35}{128}$; $C_8^4 \frac{1}{2^8} + C_8^5 \frac{1}{2^8} + \ldots + C_8^8 \frac{1}{2^8} = \frac{163}{256}$.

4. *Indicație.* Probabilitatea obținerii unei fețe cu mai mult de 4 puncte la o aruncare a zarurilor este $\frac{1}{3}$. Sîntem în cazul schemei lui Bernoulli, cu $n = 10$, $k = 4$, $p = \frac{1}{3}$.

Răspuns : $C_{10}^4 \frac{1}{3^4} \frac{2^6}{3^6}$.

5. *Indicație.* Dacă A este evenimentul obținerii a 2 bile albe și una neagră din prima urnă, iar B evenimentul obținerii a 2 bile albe și una neagră din următoarele 3 urne, aplicăm formula $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, ținînd cont că A și B sînt evenimente independente. $P(A)$ se calculează pe baza schemei lui Bernoulli, iar $P(B)$ pe baza schemei binomiale generalizate.

Răspuns : $\frac{108}{343} + \frac{146}{343} - \frac{108 \cdot 146}{343^2}$.

6. *Indicație.* Fie p_i $(i = 1, 2, 3)$ probabilitatea obținerii a 2 bile albe și 3 negre în cele 5 extrageri din urna U_i. Probabilitatea cerută se calculează pe baza schemei lui Poisson cu $n = 3$, $k = 2$ și cu probabilitățile p_1, p_2, p_3. Probabilitățile p_i se calculează fiecare pe baza schemei lui Bernoulli.

7. *Indicație.* Pe baza schemei lui Poisson se calculează probabilitatea p_i ca făcînd cîte o extragere din fiecare urnă să obținem o bilă albă și 2 negre. Probabilitatea cerută se calculează din schema lui Bernoulli, cu parametrii p, $n = 5$, $k = 3$.

8. *a)* Să numim cele două cutii *a* și *b*. Pentru a ajunge în situația dată fumătorul a scos de $2n - k + 1$ ori o cutie din buzunar (de n ori o cutie pentru a se goli, de $n - k$ ori cealaltă — pentru a-i mai rămîne k bețe — și din nou prima cutie — pentru a constata că este goală). Deci, „în $2n - k$ cazuri apare de n ori cutia *a* și de $n - k$ ori cutia *b*, în al $2n - k + 1$-lea caz apare cutia *a* „*sau*" în $2n - k$ cazuri apare de n ori cutia *b* și

$$M(X) = \sum_{k=0}^{n} kp^k = \sum_{k=0}^{n} kC_n^k p^k q^{n-k} = np,$$

$$M(X^2) = \sum_{k=0}^{n} k^2 C_n^k p^k q^{n-k} = n^2p^2 + npq.$$

$$D^2(X) = M(X^2) - (M(X))^2 = npq.$$

Metoda 2. Dacă X_k este numărul de bile albe obținut la extragerea k, atunci X_k are distribuția $\begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}$ și deci $M(X_k) = p$, $D^2(X_k) = pq$; În acest caz numărul total de bile albe obținut este

$$X = X_1 + X_2 + \ldots + X_n$$

și deci

$$M(X) = M(X_1) + M(X_2) + \ldots + M(X_n) = np$$

$$D^2(X) = D^2(X_1) + D^2(X_2) + \ldots + D^2(X_n) = npq;$$

$$(X_1, X_2, \ldots, X_n \text{ independente}).$$

24. Se raționează ca în problema precedentă.

Media $= 10 \cdot \frac{1}{2} = 5$; Dispersia $= 10 \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{5}{2}$.

Capitolul V

11. 11,62 s; **12.** Dominanta = 164 cm. Mediana = 162,88. Mediana = 162,85.

13. 0,1796; 26,76.

14. Se aplică formula (3) cu probabilitatea 0,99. Se obține

$$f_{1\,000} \in (0,459;\ 0,541).$$

Frecvența 0,576 iese din acest interval. Conchidem cu probabilitatea 0,99, că moneda nu e justă.

15. Aplicînd formulele (3), rezultă cu probabilitatea 0,99:

$$f_{1\,000} \in [0,041;\ 0,079].$$

Frecvența $f_{1\,000} = 0,085$ iese din interval. Deci mașina nu funcționează normal.

16. Se aplică formulele (4) pentru

$$p = \frac{2\,020}{9\,760},\ N = 9\,760,\ n = 500.$$

17. Se aplică formulele (4'), avînd $f = \frac{571}{1\,200}$, $N = 125\,372$, $n = 1\,200$.

18. Se aplică formulele (4') pentru $N = 96\,350$, $n = 5\,000$. Frecvențele corespunzătoare se iau din tabloul statistic.

Cuprins

Coli de tipar 6. B.T. 19.01.1982.
Format 16/61×86. Apărut 1982.

I. P. „Oltenia" Craiova
Str. M. Viteazul, nr. 4
Republica Socialistă România
Plan 28264/299/1981