# CAPITOLUL 1

# PROBABILITATE, VARIABILE ALEATOARE ŞI PROCESE STOCHASTICE

## 1.1. SCURT ISTORIC AL TEORIEI PROBABILITĂȚII

Marele matematician american Claude Elwood Shannon (30.04.1916 – 24.02.2001), fondatorul teoriei moderne a informației, şi-a formulat ideile cu ajutorul teoriei probabilității. Faptul că el şi-a scris lucrările în limba engleză nu ne obligă să învățăm mai întâi această limbă pentru a le înțelege, căci o expunere a conținutului lor în româneşte este perfect posibilă şi a fost deja făcută foarte bine; nimic, însă, nu ne dispensează de asimilarea prealabilă a bazei teoriei probabilității pentru a avea acces la teoria matematică a informației. De altfel, numeroase alte discipline de studiu se bazează pe teoria probabilității, de la fizica statistică până la sociologie şi lingvistică, aşa încât efortul nostru nu e unul foarte special, numeroşi alți studenți făcându-l în cadrul educației lor.

Originile teoriei probabilității trebuie căutate în interesul purtat de unii nobili din Europa medievală jocurilor de noroc, o petrecere a timpului care le mobiliza de minune surplusul de inteligență, neconsumat în scopuri utilitariste, ca de exemplu născocirea unor maşinării şi mecanisme, îndeletnicire lăsată pe seama minților mai prozaice ale burghezilor. Germenii teoriei probabilității au apărut pe la mijlocul secolului XVII în lucrările lui Pierre de Fermat (1601–1665), Blaise Pascal (1623–1662) şi Christian Huygens (1629–1695). Deşi cercetările lor erau inspirate de jocurile de noroc, importanța noilor concepte introduse –, cel de probabilitate a unui eveniment stochastic şi cel de valoare medie sau aşteptată a unei variabile aleatoare – pare-se că le era clară, după cum dă de înțeles Huygens în primul text despre probabilitate tipărit (1657) *Cu privire la calculele din jocurile de noroc*: „Cititorul va binevoi să remarce că nu ne ocupăm numai cu jocurile de noroc, dar că se şi pun aici fundamentele unei foarte interesante şi profunde teorii." Menţionăm că excentricul savant şi mare amator de jocuri de noroc Girolamo Cardano (1501–1576) scrisese *Cartea jocurilor şi a norocului* pe la 1520, dar ea n-a fost publicată decât în 1663. Ulterior, Jacob Bernoulli (1654–1705), Abraham de Moivre (1667–1754), reverendul Thomas Bayes (1702–1761), marchizul Pierre Simon Laplace (1729–1827), Johann Friedrich Carl Gauss (1777–1855) şi Siméon Denis Poisson (1781–1840) au contribuit semnificativ la dezvoltarea teoriei probabilităţii. Şcoala rusă a dat mari matematicieni ca P. L. Cebîşev (1821–1894) şi studenţii săi A. Markov (1856–1922) şi A. M. Liapunov (1857–1918) cu contribuţii importante legate de legea numerelor mari. Germanul Richard von Mises, pe la începutul secolului XX, a introdus o teorie a probabilităţii bazată pe definiţia probabilităţii ca frecvenţă relativă. Dar teoria deductivă bazată pe definiţia axiomatică a probabilităţii, aşa cum o studiem în zilele noastre, îi este atribuită în principal lui Andrei Nicolaevici Kolmogorov, care, în anii 1930, împreună cu Paul Lévy, a fundamentat o conexiune strânsă între teoria probabilităţii şi teoria matematică a mulţimilor şi a funcţiilor de o variabilă reală. Se cuvine menţionat, totuşi, că matematicianul francez Émile Borel (1871–1956) ajunsese la aceste idei anterior.

## 1.2. CE ÎNŢELEGEM PRIN PROBABILITATE

O teorie se numeşte *deterministă* dacă stabileşte relaţii matematice precise între diversele mărimi cu care operează. De exemplu, de la bazele electrotehnicii ştim că, dacă legăm la bornele unei baterii de 9 V un rezistor având rezistenţa nominală de 1 kΩ, prin el va curge un curent de 9 mA, conform legii lui Ohm. Dacă vom măsura acest curent cu un miliampermetru analogic, acul va indica valoarea calculată, cea la care ne aşteptăm, dar în ce priveşte precizia, aceasta depinde printre altele şi de acuitatea noastră vizuală. Pentru o măsurătoare mai obiectivă, utilizăm un miliampermetru digital, care ne arată, să spunem, 9,08113 mA. Înlocuind rezistorul cu un altul de aceeaşi valoare nominală a rezistenţei şi repetând măsurătoarea, citim acum 8,97893 mA. Păstrând rezistorul, dar înlocuind bateria cu alta de aceeaşi tensiune nominală, măsurăm acum 9,11023 mA. Care este explicaţia? Să nu fie decât aproximativă legea lui Ohm? Desigur că nu. Dar rezistoarele produse industrial cu valoarea proiectată de 1 kΩ, ca şi bateriile cu tensiunea nominală de 9 V, rezultă din procesul tehnologic cu o oarecare abatere de la valoarea nominală, suficient de mică însă pentru a fi acceptabilă în practică. Adevărul este că trebuie să proiectăm circuitele noastre electronice astfel încât să funcționeze bine, deşi componentele pe care le folosim au parametri doar aproximativ egali cu cei nominali. Observăm însă un fapt interesant: dacă măsurăm un număr mare de rezistoare de 1 kΩ şi facem media aritmetică a măsurătorilor, aceasta este apropiată de 1 kΩ, iar cu cât numărul măsurătorilor este mai mare, cu atât este mai apropiată media aritmetică de cea nominală.

Vedem deci că, în anumite domenii, media tinde spre o valoare constantă pe măsură ce numărul observațiilor creşte şi că această valoare rămâne aceeaşi dacă media se efectuează pe orice subşir specificat mai înainte de realizarea experimentului. Teoria probabilității cu aceasta se ocupă, cu mediile unor fenomene de masă ce apar pe rând sau simultan: emisia electronilor, apelurile telefonice, defectările unor sisteme tehnice, rata natalității şi cea a mortalității, zgomotul şi multe altele.

Teoria are menirea de a prezice asemenea medii cu ajutorul probabilității evenimentelor. În vreme ce jocul de şah este absolut determinist –, făcând abstracție de caracterul conjunctural al condiției fizice şi mentale a celor doi adversari, jocurile de noroc (de la rişcă şi barbut până la loteria de stat, trecând prin jocurile de cazino, ca ruleta şi bacara) au un caracter probabilist prin excelență, încât, fără a le recomanda, le putem utiliza în scop ilustrativ.

Să catapultăm o monedă cu un bobârnac; după un zbor elegant prin aer, moneda va ateriza pe o suprafață plană cu una din cele două fețe în sus. În această carte, vom numi convențional cele două fețe „capul" şi „banul". Faptul că anumite monede au o stemă sau o pajură în loc de capul unei figuri istorice nu schimbă, desigur, datele experimentului nostru. Să presupunem că a căzut „capul". Putem deduce din aceasta că, la o nouă aruncare a monedei, va cădea din nou capul? (Să renunțăm la ghilimele). Nicidecum, după cum bine ştim din experiență. Să repetăm, totuşi, experimentul de un număr $n$ de ori, să spunem $n = 10$. De $n_c$ ori va cădea capul, de $n_b$ ori banul, astfel încât $n_b + n_c = n$. Pentru un număr $n$ mic de încercări, se poate ca $n_b \neq n_c$. Observăm, însă, că, pe măsură ce $n$ creşte, $n_b$ şi $n_c$ sunt tot mai apropiate. Pentru o monedă nemăsluită şi un număr mare $n$ de încercări, $n_b$ şi $n_c$ trebuie să fie egale.

Să începem acum să ne familiarizăm cu terminologia specifică teoriei probabilității. Aruncarea monedei o singură dată este un exemplu de *experiment*. Efectuarea experimentului este o *încercare*. Faptul că moneda a căzut cu o față în sus este *rezultatul* experimentului. Spațiul de probabilitate, sau spațiul eşantioanelor S al unui experiment constă din mulțimea tuturor rezultatelor posibile. În cazul aruncării monedei, S ={capul, banul}, iar în cazul aruncării zarului, $S = \{f_1, f_2, f_3, f_4, f_5, f_6\}$, unde $f_i$ este fațeta marcată cu $i$ puncte. Cele două, respectiv şase rezultate posibile sunt *punctele eşantion* ale experimentului. Un *eveniment* este o submulţime a lui S şi poate consta din orice număr de puncte eşantion.

Fie S spaţiul eşantioanelor pentru experimentul constând din aruncarea zarului. Un posibil eveniment este $A=\{f_2, f_4\}$ care constă din rezultatele $f_2$ şi $f_4$. Complementul evenimentului $A$, notat $\overline{A}$, constă din toate punctele eşantion din S care nu sunt în $A$: $\overline{A} = \{f_1, f_3, f_5, f_6\}$. Două evenimente se spune că sunt mutual exclusive dacă n-au puncte eşantion în comun, adică, dacă apariţia unui eveniment exclude apariţia celuilalt. Cu consecvenţă, de acum înainte, în locul cuvântului *întâmplător*, vom utiliza termenul ştiinţific de *aleator*. (În limba latină, *aleator* înseamnă jucător de zaruri). Pentru ca un experiment să fie aleator, trebuie să fie satisfăcute trei cerinţe:

1. Experimentul să fie repetabil în condiţii identice.
2. La orice încercare a experimentului, rezultatul să fie imprevizibil.
3. Pentru un număr mare de încercări ale experimentului, rezultatele să prezinte *regularitate statistică*. Aceasta înseamnă că, dacă se repetă experimentul de un număr mare de ori, se observă că rezultatele au *medii* definite.

Să luăm acum un zar. Presupunem că oricine a văzut un zar, dar să îl descriem totuşi ca pe un mic cub din os sau din plastic având marcate cele şase faţete cu un număr de puncte, de la 1 la 6. Dacă aruncăm un zar permiţându-i să se rostogolească de mai multe ori, el se va opri în cele din urmă, pe o suprafaţă plată orizontală, cu o faţetă în sus. Care va fi ea? Nu ştim cu anticipaţie decât că va fi una din cele şase. Apariţia unei faţete este un eveniment. Un eveniment apare cu o anumită *probabilitate*. Ştim cu toţii, fără prea multă teorie, că probabilitatea unei faţete este egală cu 1/6. Dar cum definim probabilitatea unui eveniment? Vom vedea aceasta în secţiunea următoare.

## 1.3. DEFINIŢIA PROBABILITĂŢII

Sunt trei definiţii ale probabilităţii, după cum urmează.

### Definiţia clasică

Timp de secole, s-a utilizat definiţia clasică, potrivit căreia probabilitatea $P(A)$ a unui eveniment $A$ se determină *a priori* fără a efectua vreun experiment. Ea este dată de raportul

$$P(A) = \frac{N_A}{N} \tag{1.1}$$

unde $N$ este numărul rezultatelor *posibile* iar $N_A$ este numărul rezultatelor care sunt *favorabile* evenimentului $A$.

În experimentul cu zarul, sunt şase rezultate posibile, astfel încât probabilitatea oricăruia dintre ele este egală cu 1/6. Rezultatele favorabile evenimentului compus *par* (adică, oricare din fațetele $f_2$, $f_4$, $f_6$) sunt în număr de trei astfel încât $P(par) = 3/6$.

În experimentul cu o monedă nemăsluită, $N = 2$ şi $N_c = N_b = 1$, astfel încât probabilitatea de a cădea *capul* este egală cu cea de a cădea *banul* şi deci $P(cap) = P(ban) = 1/2$.

Nu rareori, însă, semnificația numerelor $N$ şi $N_A$ nu este clară. Spre exemplu, care este probabilitatea $p$ ca, aruncând două zaruri, suma numerelor marcate pe fațetele vizibile după oprire să fie 7? Pentru a rezolva această problemă utilizând definiția clasică (1.1), trebuie să determinăm numerele $N$ şi $N_A$.

a) Am putea considera drept rezultate posibile cele 11 sume 2, 3, ... , 12. Dintre acestea, numai una dintre ele, şi anume 7, este favorabilă, astfel încât $p = 1/11$. Acest rezultat este greşit.

b) Am putea număra drept rezultate posibile toate perechile de numere fără a face deosebire între primul şi al doilea zar. Avem acum 21 de rezultate dintre care favorabile sunt perechile (3, 4), (5, 2) şi (6, 1). Deci $N_A = 3$ şi $N = 21$, de unde $p = 3/21$. Şi acest rezultat este greşit.

c) Soluțiile precedente sunt greşite fiindcă rezultatele nu sunt egal probabile. Trebuie să socotim toate perechile de numere făcând deosebire între primul şi al doilea zar. Numărul de rezultate posibile este acum 36, iar rezultatele favorabile sunt cele şase perechi (3, 4), (4, 3), (5, 2), (2, 5), (6, 1) şi (1, 6). Prin urmare, $p = 6/36$.

## Definiția probabilității ca frecvență relativă

Probabilitatea $P(A)$ a unui eveniment $A$ este limita

$$P(A) = \lim_{n\to\infty} \frac{n_A}{n} \tag{1.2}$$

unde $n_A$ este numărul de apariții ale lui $A$ iar $n$ este numărul de încercări.

În practică, vom aproxima infinitul printr-un număr mare de încercări. Spre exemplu, am putea calcula probabilitatea utilizând (1.2) pentru câteva valori ale lui $n$ din ce în ce mai mari, până când obţinem aproximativ aceeaşi valoare pentru $P(A)$.

## Definiţia axiomatică a probabilităţii

Ne vom baza pe teoria mulţimilor, cu care suntem cu toţii familiarizaţi. Să notăm cu $f_i$ faţetele unui zar. Ele sunt elementele mulţimii S $=\{f_1, f_2, f_3, f_4, f_5, f_6\}$. Numim această mulţime *spaţiu de probabilitate* şi în acelaşi timp *evenimentul sigur*, căci la fiecare încercare, este sigur că va apărea unul din elementele sale. Elementele sale se numesc *rezultatele experimentale*. Submulţimile sale se numesc *evenimente*. În cazul aruncării zarului, sunt $2^6 = 64$ de submulţimi: $\{\varnothing\}$, $\{f_1\}$,..., $\{f_1, f_2\}$,..., $\{f_1, f_2, f_3\}$,..., S. Mulţimea vidă $\{\varnothing\}$ este *evenimentul imposibil*, iar evenimentul $\{f_i\}$ constând dintr-un singur element $f_i$ este un *eveniment elementar*.

Atribuim fiecărui eveniment $A$ un număr $P(A)$, pe care-l numim *probabilitatea evenimentului A*. Acest număr se alege astfel încât să fie satisfăcute următoarele trei condiţii, care sunt axiomele teoriei probabilităţii:

**A1.**
$$P(A) \geq 0 \tag{1.3}$$

**A2.**
$$P(S) = 1 \tag{1.4}$$

**A3.** Dacă $A \cap B = \{\varnothing\}$,
$$P(A \cup B) = P(A) + P(B) \tag{1.5}$$

Pe baza axiomelor, deducem următoarele proprietăţi.

**P1.** Probabilitatea evenimentului imposibil este zero:
$$P\{\varnothing\} = 0 \tag{1.6}$$

Într-adevăr, $A \cap \{\varnothing\} = \{\varnothing\}$ şi $A \cup \{\varnothing\} = A$. De aceea, conform axiomei **A3**, avem:
$$P(A) = P(A \cup \varnothing\} = P(A) + P\{\varnothing\}$$

**P2.** Pentru orice $A$,
$$P(A) = 1 - P(\overline{A}) \leq 1 \tag{1.7}$$

din cauză că $A \cup \overline{A} = S$ şi $A \cap \overline{A} = \{\varnothing\}$, de unde
$$1 = P(S) = P(A \cup \overline{A}) = P(A) + P(\overline{A})$$

**P3.** Pentru orice *A* şi *B*,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \le P(A) + P(B) \qquad (1.8)$$

Pentru a demonstra această proprietate, scriem evenimentele $A \cup B$ şi *B* ca reuniuni de două evenimente care se exclud reciproc:

$$A \cup B = A \cup (\overline{A} \cap B)$$
$$B = (A \cap B) \cup (\overline{A} \cap B)$$

De aceea, aplicând axioma **A3**, putem scrie:

$$P(A \cup B) = P(A) + P(\overline{A} \cap B)$$
$$P(B) = P(A \cap B) + P(\overline{A} \cap B)$$

Eliminând $P(\overline{A} \cap B)$, obţinem (1.8).

**P4.** Dacă $B \subset A$, avem

$$P(A) = P(B) + P(A \cap \overline{B}) \ge P(B) \qquad (1.9)$$

din cauză că $A = B \cup (A \cap \overline{B})$ şi $B \cap (A \cap \overline{B}) = \{\varnothing\}$.

## 1.4. CLASA *F* A EVENIMENTELOR

Evenimentele sunt submulţimi ale lui S cărora le-am atribuit probabilităţi. Pentru a elimina unele dificultăţi matematice legate de mulţimi cu un număr infinit de rezultate, nu vom considera drept evenimente toate submulţimile lui S, ci doar o clasă ***F*** de submulţimi.

Un câmp ***F*** este o clasă nevidă de mulţimi astfel încât:

**P1.** Dacă $A \in \boldsymbol{F}$, atunci $\overline{A} \in \boldsymbol{F}$ (1.10)

**P2.** Dacă $A \in \boldsymbol{F}$ şi $B \in \boldsymbol{F}$, atunci $(A \cup B) \in \boldsymbol{F}$ (1.11)

Aceste două proprietăţi sunt suficiente pentru ca ***F*** să fie un câmp. Din ele, rezultă proprietăţile următoare:

**P3.** Dacă $A \in \boldsymbol{F}$ şi $B \in \boldsymbol{F}$, atunci $(A \cap B) \in \boldsymbol{F}$ (1.12)

Într-adevăr, din (1.10) urmează că $\overline{A} \in \boldsymbol{F}$ şi $\overline{B} \in \boldsymbol{F}$. Aplicând **P1** şi **P2** la mulţimile $\overline{A}$ şi $\overline{B}$, rezultă că $\overline{A} \cup \overline{B} \in \boldsymbol{F}$ şi $\overline{\overline{A} \cup \overline{B}} \in \boldsymbol{F}$.

**P4.** Un câmp conţine evenimentul sigur şi evenimentul imposibil:

$$S \in F \text{ şi } \{\varnothing\} \in F \qquad (1.13)$$

Într-adevăr, fiindcă ***F*** nu este o mulţime vidă, conţine cel puţin un element $A$, iar conform **P1**, va conţine şi $\overline{A}$**.** Prin urmare, $A \cup \overline{A} = S \in$ ***F*** şi $A \cap \overline{A} = \{\varnothing\} \in$ ***F***.

Din cele de mai sus, rezultă că toate mulţimile ce se pot scrie ca reuniuni sau intersecţii ale unui număr finit de mulţimi din ***F*** sunt şi ele în ***F***. Aceasta, însă, nu este în mod necesar cazul şi pentru un număr infinit de mulţimi. Suntem astfel nevoiţi să introducem noţiunea de *câmp Borel*.

Dacă pentru orice şir infinit $A_1, \cdots, A_n, \cdots$ de mulţimi din ***F***, reuniunea şi intersecţia acestor mulţimi aparţin şi ele lui ***F***, atunci ***F*** se numeşte un câmp Borel.

Clasa tuturor submulţimilor unei mulţimi S este un câmp Borel. Să presupunem că o clasă C de submulţimi ale lui S nu este un câmp. Adăugându-i alte submulţimi ale lui S, toate submulţimile dacă este necesar, putem forma un câmp având C drept submulţime.

Pentru cazul unui număr infinit de mulţimi, trebuie să adăugăm la cele trei axiome ale probabilităţii o a patra, numită *axioma probabilităţii infinite*:

**A4.** Dacă evenimentele $A_1, A_2, \cdots$ sunt mutual exclusive, atunci

$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots \qquad (1.14)$$

## 1.5. DEFINIŢIA AXIOMATICĂ A UNUI EXPERIMENT

În teoria probabilităţii, specificăm un experiment cu ajutorul următoarelor concepte:

**1.** Mulţimea S a tuturor rezultatelor experimentale.
**2.** Câmpul Borel al tuturor evenimentelor din S.
**3.** Probabilităţile acestor evenimente.

Dacă spaţiul S constă din $N$ rezultate, unde $N$ este un număr finit, iar evenimentul elementar $\zeta_i$ are probabilitatea $p_i$, conform axiomelor probabilităţii, numerele $p_i$ trebuie să fie pozitive iar suma lor trebuie să fie egală cu 1:

$$p_i \geq 0 \text{ şi } p_1 + \cdots + p_N = 1 \qquad (1.15)$$

Probabilitatea oricărui eveniment $A$ constând din $r$ evenimente elementare se poate scrie ca suma probabilităţilor acestora.

## 1.6. DREAPTA REALĂ

Să presupunem că S este mulţimea numerelor reale. Submulţimile sale pot fi considerate drept mulţimi de puncte de pe dreapta numerelor reale. Se poate arăta că este imposibil să atribuim probabilităţi tuturor submulţimilor lui S astfel încât să satisfacă axiomele **A1**–**A4**. Pentru a construi un spaţiu de probabilitate pe dreapta numerelor reale, vom considera drept evenimente intervalele $x_1 \leq x \leq x_2$ precum şi reuniunile şi intersecţiile numărabile ale acestora. Aceste evenimente formează un câmp ***F*** definit drept cel mai mic câmp Borel ce include toate semidreptele $x \leq x_i$, unde $x_i$ este orice număr real. Acest cîmp conţine toate intervalele deschise şi închise, toate punctele şi, practic, toate mulţimile de puncte de pe dreapta numerelor reale ce prezintă interes în aplicaţii. Menţionăm că există mulţimi de puncte de pe dreapta numerelor reale care nu sunt reuniuni şi intersecţii numărabile de intervale, dar ele nu prezintă interes în majoritatea aplicaţiilor.

Presupunem că $p(x) \geq 0$ este o funcţie astfel încât

$$\int_{-\infty}^{\infty} p(x)\mathrm{d}x = 1 \qquad (1.16)$$

Definim probabilitatea evenimentului $\{x \leq x_i\}$ prin integrala

$$P\{x \leq x_i\} = \int_{-\infty}^{x_i} p(x)\mathrm{d}x \qquad (1.17)$$

Probabilitatea evenimentului $\{x_1 < x \leq x_2\}$ constând din toate punctele din intervalul $(x_1, x_2]$ este dată de

$$P\{x_1 < x \leq x_2\} = \int_{x_1}^{x_2} p(x)\mathrm{d}x \qquad (1.18)$$

Într-adevăr, evenimentele $\{x \leq x_1\}$ şi $\{x_1 < x \leq x_2\}$ sunt mutual exclusive iar reuniunea lor este egală cu $\{x \leq x_2\}$. Conform axiomei **A3**, avem

$$P\{x \le x_1\} + P\{x_1 < x \le x_2\} = P\{x \le x_2\}$$

şi (1.8) rezultă din (1.17).

Observăm că, dacă funcţia $p(x)$ este mărginită, integrala din (1.18) tinde spre 0 pentru $x_1 \to x_2$. Aceasta conduce la concluzia că probabilitatea evenimentului $\{x_1\}$ constând din rezultatul $x_1$ este 0 pentru orice $x_1$. Deci, probabilitatea tuturor evenimentelor elementare din S este egală cu 0, deşi probabilitatea reuniunii lor este egală cu 1. Aceasta nu vine în conflict cu axioma **A4**, căci mulţimea elementelor lui S nu este numărabilă.

## Mase de probabilitate

Putem interpreta probabilitatea $P(A)$ a unui eveniment $A$ drept masa figurii corespunzătoare din reprezentarea prin diagrama Venn. Fie, de exemplu, identitatea

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

Membrul stâng este egal cu masa evenimentului $A \cup B$. În suma $P(A) + P(B)$, masa lui $A \cap B$ este socotită de două ori. Pentru a obţine din această sumă $P(A \cup B)$, trebuie deci să scădem din ea $P(A \cap B)$.

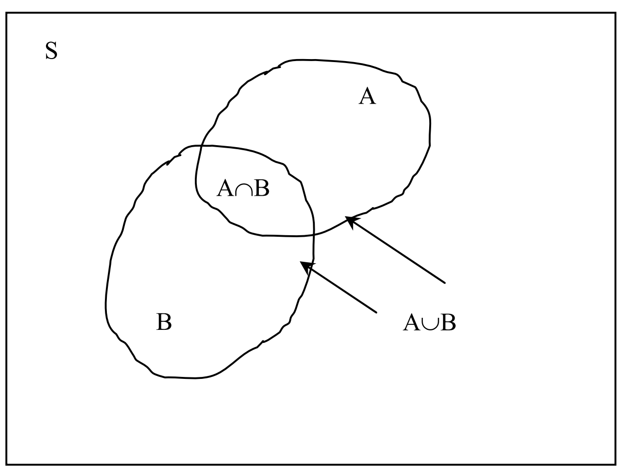


**Fig. 1.1.** Diagrama Venn a spaţiului de probabilitate S arătând două mulţimi de evenimente $A$ şi $B$ împreună cu reuniunea $A \cup B$ şi intersecţia $A \cap B$ lor.

## 1.7. EVENIMENTE COMUNE ŞI PROBABILITĂŢI COMUNE

Fie două experimente, primul constând din aruncarea unui zar nemăsluit

$$S_1 = \{f_1, \cdots, f_6\} \text{ cu } P_1(f_i) = \tfrac{1}{6}$$

şi al doilea constând din aruncarea unei monede nemăsluite

$$S_2 = \{c, b\} \text{ cu } P_2\{c\} = P_2\{b\} = ½.$$

Să considerăm experimentul combinat constând din efectuarea ambelor experimente $S_1$ şi $S_2$. Produsul cartezian $S_1 \times S_2$ este o mulţime S ale cărei elemente sunt toate perechile ordonate $\zeta_1\zeta_2$, unde $\zeta_1 \in S_1$ şi $\zeta_2 \in S_2$.

$$S = S_1 \times S_2$$

Dacă $A$ este o submulţime a lui $S_1$ iar $B$ este o submulţime a lui $S_2$, mulţimea $C = A \times B$, constând din toate perechile $\zeta_1\zeta_2$, unde $\zeta_1 \in A$ şi $\zeta_2 \in B$, este o submulţime a lui S.

Produsul cartezian al experimentelor $S_1$ şi $S_2$ este un nou experiment $S = S_1 \times S_2$ ale cărui evenimente sunt toate produsele carteziene de forma $A \times B$, unde $A$ este un eveniment din $S_1$ iar $B$ este un evenimente din $S_2$, precum şi reuniunile şi intersecţiile acestora.

În acest experiment, dacă $P_1(A)$ este probabilitatea evenimentului $A$ în experimentul $S_1$ iar $P_2(B)$ este probabilitatea evenimentului $B$ în experimentul $S_2$, probabilităţile evenimentelor $A \times S_2$ şi $S_1 \times B$ sunt

$$P(A \times S_2) = P_1(A) \tag{1.19}$$

$$P(S_1 \times B) = P_2(B) \tag{1.20}$$

### Experimente independente

Două evenimente $A$ şi $B$ se spune că sunt independente dacă

$$P(A \cap B) = P(A)P(B) \tag{1.21}$$

În multe aplicaţii, evenimentele $A \times S_2$ şi $S_1 \times B$ din experimentul combinat S sunt independente pentru orice $A$ şi $B$. Întrucât intersecţia acestor evenimente este egală cu $A \times B$, din (1.19), (1.20) şi (1.21) conchidem că

$$P(A \times B) = P(A \times S_2)P(S_1 \times B) = P_1(A)P_2(B) \tag{1.22}$$

Vom reformula conceptul de experiment combinat astfel: dacă un experiment are rezultatele posibile $A_i$, $i = 1,2,\cdots,n$, iar cel de al doilea experiment are rezultatele posibile $B_j$, $j = 1,2,\cdots,m$, experimentul combinat are rezultatele comune posibile $(A_i, B_j), i = 1,2,\cdots,n,\ j = 1,2,\cdots,m.$ Probabilitatea comună $P(A_i, B_j)$ asociată cu evenimentul comun $(A_i, B_j)$ satisface condiția

$$0 \le P(A_i, B_j) \le 1$$

Presupunând că rezultatele $B_j$, $j = 1,2,\cdots,m$ sunt mutual exclusive, urmează că

$$\sum_{j=1}^{m} P(A_i, B_j) = P(A_i) \tag{1.23}$$

Similar, dacă rezultatele $A_i, i = 1,2,\cdots,n$ sunt mutual exclusive, avem

$$\sum_{i=1}^{n} P(A_i, B_j) = P(B_j) \tag{1.24}$$

$P(A_i)$ din (1.23) şi $P(B_j)$ din (1.24) se spune că sunt probabilități *marginale*. Dacă toate rezultatele celor două experimente sunt mutual exclusive, avem

$$\sum_{i=1}^{n} \sum_{j=1}^{m} P(A_i, B_j) = 1 \tag{1.25}$$

## 1.8. PROBABILITĂȚI CONDIȚIONATE

Să considerăm un experiment combinat în care un eveniment comun apare cu probabilitate $P(A, B)$. Presupunem că evenimentul $B$ a apărut şi vrem să determinăm probabilitatea de apariție a evenimentului $A$. Aceasta se numeşte probabilitatea evenimentului $A$ condiționată de apariția evenimentului $B$ şi se defineşte prin relația

$$P(A|B) = \frac{P(A,B)}{P(B)} \tag{1.26}$$

pentru $P(B) > 0$. Similar, probabilitatea evenimentului $B$ condiționată de apariția evenimentului $A$ este

$$P(B|A) = \frac{P(A,B)}{P(A)} \tag{1.27}$$

pentru $P(A) > 0$. Combinând (1.26) şi (1.27), obținem:

$$P(A,B) = P(A|B)P(B) = P(B|A)P(A) \tag{1.28}$$

Aceste relații se aplică şi în cazul unui singur experiment în care $A$ şi $B$ sunt două evenimente definite pe spațiul eşantioanelor S iar $P(A, B)$ este interpretată drept probabilitatea lui $A \cap B$. Cu alte cuvinte, $P(A, B)$ este probabilitatea apariției simultane a evenimentelor $A$ şi $B$. Dacă două evenimente $A$ şi $B$ sunt mutual exclusive, $A \cap B = \{\varnothing\}$ şi deci $P(A|B) = 0$. Dacă $A$ este o submulțime a lui $B$, $A \cap B = A$ şi deci

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

Dacă însă $B$ este o submulțime a lui $A$, avem $A \cap B = B$ şi deci

$$P(A|B) = \frac{P(B)}{P(B)} = 1.$$

## Teorema probabilității totale

O partiție $U = [A_1, \cdots, A_n]$ a lui S este prin definiție o colecție de $n$ submulțimi $A_1, \cdots, A_n$ ale lui S disjuncte a căror reuniune este egală cu S. Fie $B$ un eveniment arbitrar. Avem atunci

$$P(B) = P(B|A_1)P(A_1) + \ldots + P(B|A_n)P(A_n) \tag{1.29}$$

**Demonstrație**

Este clar că

$$B = B \cap S = B \cap (A_1 \cup \cdots \cup A_n) = (B \cap A_1) \cup \cdots (B \cap A_n) .$$

Dar evenimentele $B \cap A_i$ şi $B \cap A_j$ sunt mutual exclusive din cauză că evenimentele $A_i$ şi $A_j$ sunt mutual exclusive. Prin urmare

$$P(B) = P(B \cap A_1) + \cdots + P(B \cap A_n) .$$

Conform (1.28), avem însă

$$P(B \cap A_i) = P(B|A_i)P(A_i)$$

Utilizând aceasta în expresia lui $P(B)$, rezultă (1.29).

**Teorema lui Bayes**

Aplicând (1.28), putem scrie că

$$P(A_i \mid B) = P(B \mid A_i)\frac{P(A_i)}{P(B)} \tag{1.30}$$

Introducând (1.29) în (1.30), obținem teorema lui Bayes:

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{P(B \mid A_1)P(A_1) + \cdots + P(B \mid A_n)P(A_n)} \tag{1.31}$$

**Independență statistică**

Să presupunem că apariția lui $A$ nu depinde de apariția lui $B$, ceea ce înseamnă că

$$P(A \mid B) = P(A) \tag{1.32}$$

Înlocuind (1.32) în (1.26), obținem că

$$P(A, B) = P(A)P(B) \tag{1.33}$$

Dacă probabilitatea comună a evenimentelor $A$ şi $B$ se descompune în produsul probabilităților marginale $P(A)$ şi $P(B)$, se spune că evenimentele sunt *independente statistic.*

Definiția independenței statistice se poate generaliza la un număr $n$ de evenimente. Spre exemplu, trei evenimente independente statistic $A_1$, $A_2$ şi $A_3$ trebuie să satisfacă următoarele condiții:

$$\begin{aligned} P(A_1, A_2) &= P(A_1)P(A_2) \\ P(A_1, A_3) &= P(A_1)P(A_3) \\ P(A_2, A_3) &= P(A_2)P(A_3) \\ P(A_1, A_2, A_3) &= P(A_1)P(A_2)P(A_3) \end{aligned} \tag{1.34}$$

## 1.9. CONCEPTUL DE VARIABILĂ ALEATOARE

Se dă un experiment având un spaţiu al eşantioanelor S şi elemente $s \in$ S. Fie o funcţie $X(s)$ al cărei domeniu de definiţie este S şi al cărei domeniu de existenţă este mulţimea numerelor reale, considerate drept coordonate ale punctelor de pe o dreaptă. Funcţia $X(s)$ se numeşte o *variabilă aleatoare*. Spre exemplu, în experimentul cu moneda, rezultatele posibile sunt capul ($c$) şi banul ($b$), astfel încât S conţine numai două puncte notate $c$ şi $b$. Putem defini o funcţie $X(s)$ astfel încât

$$X(s) = \begin{cases} 1 & (s = c) \\ -1 & (s = b) \end{cases} \tag{1.35}$$

În experimentul cu zarul, S $=\{f_1, f_2, f_3, f_4, f_5, f_6\}$. O variabilă aleatoare definită pe spaţiul eşantioanelor este $X(f_i) = i$, în care caz rezultatele experimentului sunt aplicate în întregii 1, ..., 6.

Deşi am dat exemple de experimente care au o mulţime finită de rezultate posibile, spaţiul eşantioanelor poate fi la fel de bine un domeniu continuu, în care caz spunem că variabila aleatoare este *continuă*. Pentru simplitate, variabila aleatoare $X(s)$ se scrie $X$. Adoptăm convenţia de a nota variabila aleatoare cu literă mare, iar o valoare pe care o poate ea lua cu litera mică respectivă. Vom scrie concis că variabila aleatoare $X$ are distribuţia de probabilitate $p(X)$ astfel: $X \sim p(x)$.

Se dă o variabilă aleatoare $X$. Să considerăm evenimentul $\{X \le x\}$ unde $x$ este orice număr real din intervalul $(-\infty, \infty)$. Scriem probabilitatea acestui eveniment $P(X \le x)$ şi o notăm cu $F_X(x)$

$$F_X(x) = P(X \le x), \quad -\infty < x < \infty \tag{1.36}$$

Indicele $X$ din $F_X(x)$ arată despre ce variabilă aleatoare este vorba şi poate fi omis dacă nu există ambiguitate. Funcţia $F(x)$ se numeşte *funcţia distribuţie de probabilitate* a variabilei aleatoare $X$. Ea se mai numeşte şi *funcţia de distribuţie cumulativă* (FDC). Fiindcă $F(x)$ este o probabilitate, domeniul ei de existenţă este limitat la intervalul $0 \le F(x) \le 1$. Conform definiţiei (1.36), avem:

$$F(-\infty) = P(X \le -\infty) = 0$$
$$F(+\infty) = P(X \le \infty) = 1.$$

Variabila aleatoare generată de aruncarea unei monede nemăsluite şi definită de ecuaţia (1.35) are FDC arătată în figura 1.2.

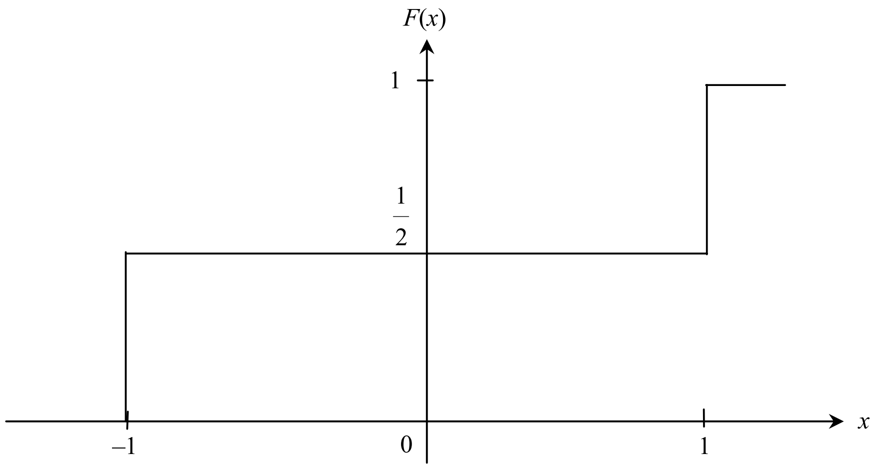


**Fig. 1.2.** Funcţia de distribuţie cumulativă a unei variabile aleatoare generate de aruncarea unei monede conform ecuaţiei (1.35).

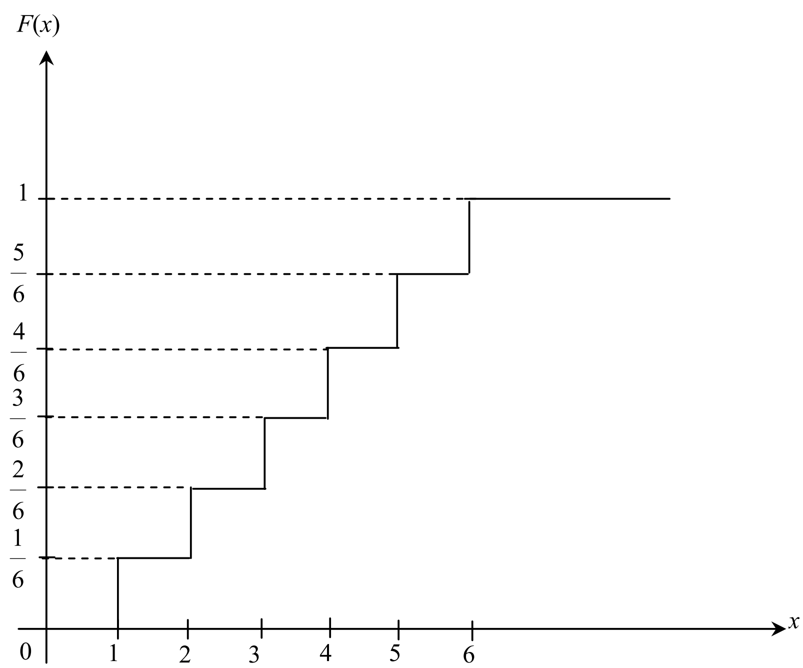


**Fig. 1.3.** Funcţia de distribuţie cumulativă a unei variabile aleatoare generate prin aruncarea zarului conform ecuaţiei $X(f_i) = i$ .

$F(x)$ are două discontinuităţi sau salturi, una la $x = -1$ şi alta la $x = +1$. Similar, variabila aleatoare $X(s) = s$ generată de aruncarea unui zar are FDC arătată în figura 1.3. În acest caz, $F(x)$ are şase salturi, câte unul la fiecare din punctele $x = 1, \ldots, 6$.

Funcţia de distribuţie cumulativă a unei variabile aleatoare continue, aşa cum se vede şi din figura 1.4, este o funcţie continuă, nedescrescătoare de $x$.

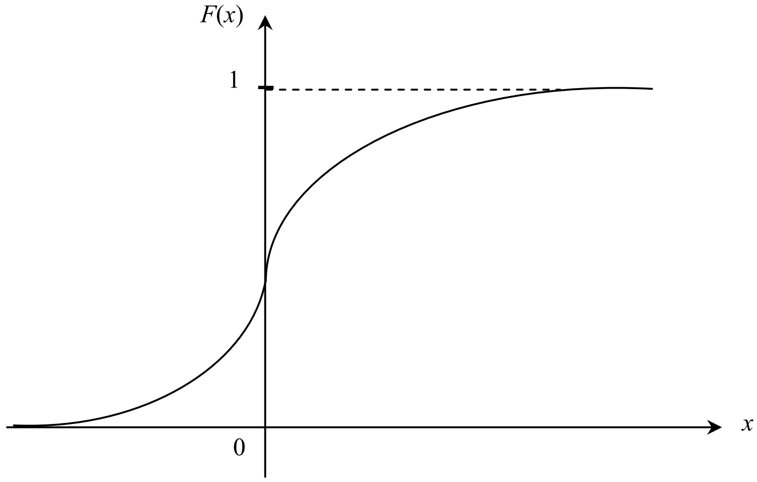


**Fig. 1.4.** Exemplu de funcţie de distribuţie cumulativă a unei variabile aleatoare continue.

În practică, putem întâlni şi variabile aleatoare de tip *mixt*, aşa cum se arată în figura 1.5. FDC a unei astfel de variabile aleatoare este o funcţie continuă pe porţiuni, nedescrescătoare şi conţine salturi într-un număr de valori discrete ale lui $x$.

Derivata lui $F(x)$, notată $p(x)$, se numeşte *funcţia densitate de probabilitate* (FDP) a variabilei aleatoare $X$. Deci, avem

$$p(x) = \frac{\mathrm{d}F(x)}{\mathrm{d}x} \qquad -\infty < x < \infty \tag{1.37}$$

sau, echivalent,

$$F(x) = \int_{-\infty}^{x} p(u)\mathrm{d}u, \quad -\infty < x < \infty \tag{1.38}$$

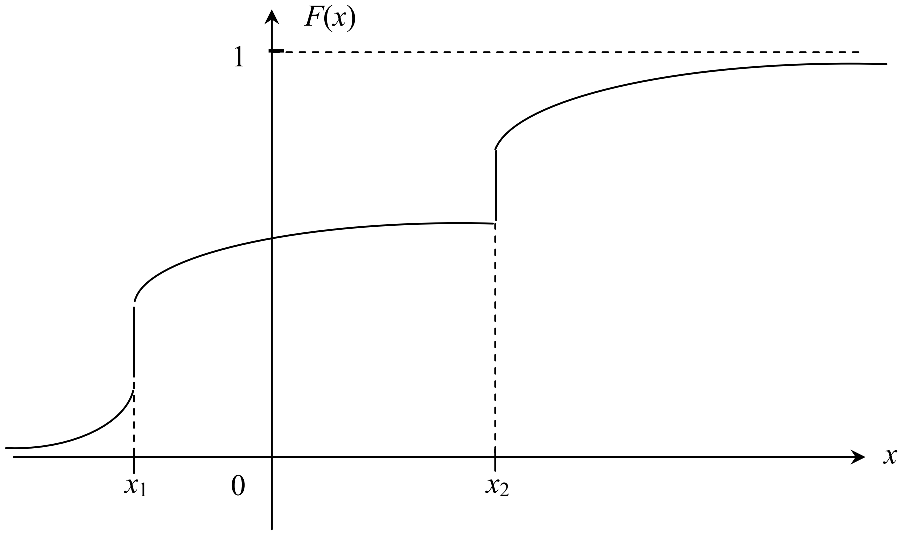


**Fig. 1.5.** Exemplu de funcție de distribuție cumulativă a unei variabile aleatoare de tip mixt.

Fiindcă $F(x)$ este o funcție nedescrescătoare, urmează că $p(x) \geq 0$. Dacă variabila aleatoare este discretă sau de tip mixt, FDP conține impulsuri în punctele de discontinuitate ale lui $F(x)$. În asemenea cazuri, partea discretă a lui $p(x)$ poate fi exprimată astfel:

$$p(x) = \sum_{i=1}^{n} P(X = x_i)\delta(x - x_i) \tag{1.39}$$

unde $x_i, i = 1,2,\ldots,n$ sunt valorile discrete posibile ale variabilei aleatoare, $P(X = x_i), i = 1,2,\ldots,n$ sunt probabilitățile iar cu $\delta(x)$ am notat un impuls Dirac în $x = 0$.

Să determinăm probabilitatea ca o variabilă aleatoare $X$ să cadă într-un interval $(x_1, x_2)$, unde $x_2 > x_1$. Este evident că evenimentul $\{X \leq x_2\}$ este reuniunea a două evenimente mutual exclusive $\{X \leq x_1\}$ şi $\{x_1 < X \leq x_2\}$. Conform axiomei **A3**, putem exprima probabilitatea evenimentului $\{X \leq x_2\}$ drept suma probabilităților celor două evenimente mutual exclusive. Avem deci

$$P(X \le x_2) = P(X \le x_1) + P(x_1 < X \le x_2)$$
$$F(x_2) = F(x_1) + P(x_1 < X \le x_2)$$

sau, echivalent,

$$P(x_1 < X \le x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} p(x)\mathrm{d}x \tag{1.40}$$

Cu alte cuvinte, probabilitatea evenimentului $\{x_1 < X \le x_2\}$ este aria suprafeței de sub curba FDP în intervalul $x_1 < X \le x_2$.

Fie acum două variabile aleatoare $X_1$ şi $X_2$, fiecare din ele putând fi continuă, discretă sau mixtă. Prin definiție, FDC comună a celor două variabile aleatoare este

$$\begin{aligned} F(x_1, x_2) &= P(X_1 \le x_1, X_2 \le x_2) \\ &= \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} p(u_1, u_2)\mathrm{d}u_1\mathrm{d}u_2 \end{aligned} \tag{1.41}$$

unde $p(x_1, x_2)$ este FDP comună. Putem exprima această FDP comună în forma

$$p(x_1, x_2) = \frac{\partial^2}{\partial x_1 \partial x_2} F(x_1, x_2) \tag{1.42}$$

Dacă integrăm $p(x_1, x_2)$ pe una din variabile, obținem FDP a celeilalte variabile:

$$\int_{-\infty}^{\infty} p(x_1, x_2)\mathrm{d}x_1 = p(x_2) \tag{1.43}$$

$$\int_{-\infty}^{\infty} p(x_1, x_2)\mathrm{d}x_2 = p(x_1) \tag{1.44}$$

Aceste funcții densitate de probabilitate $p(x_1)$ și $p(x_2)$ obținute integrând pe una din variabile se numesc funcții densitate de probabilitate *marginale*. Integrând $p(x_1, x_2)$ pe ambele variabile, obținem:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 = F(\infty, \infty) = 1 \tag{1.45}$$

Se observă că $F(-\infty, -\infty) = F(-\infty, x_2) = F(x_1, -\infty) = 0$.

Vrem acum să determinăm probabilitatea ca variabila aleatoare $X_1 \le x_1$ condiționată de faptul că $x_2 - \Delta x_2 < X_2 \le x_2$ unde $\Delta x_2$ este un increment pozitiv. Cu alte cuvinte, dorim să determinăm probabilitatea evenimentului $(X_1 \le x_1 \mid x_2 - \Delta x_2 < X_2 \le x_2)$. Conform definiției probabilității condiționate a unui eveniment, putem exprima probabilitatea evenimentului $(X_1 \le x_1 \mid x_2 - \Delta x_2 < X_2 \le x_2)$ drept probabilitatea evenimentului comun $(X_1 \le x_1, x_2 - \Delta x_2 < X_2 \le x_2)$ împărțită la probabilitatea evenimentului $(x_2 - \Delta x_2 < X_2 \le x_2)$. Deci

$$\begin{aligned} P(X_1 \le x_1 \mid x_2 - \Delta x_2 < X_2 \le x_2) &= \frac{\int_{-\infty}^{x_1} \int_{x_2-\Delta x_2}^{x_2} p(u_1, u_2) \mathrm{d}u_1 \mathrm{d}u_2}{\int_{x_2-\Delta x_2}^{x_2} p(u_2) \mathrm{d}u_2} \\ &= \frac{F(x_1, x_2) - F(x_1, x_2 - \Delta x_2)}{F(x_2) - F(x_2 - \Delta x_2)} \end{aligned} \tag{1.46}$$

În ipoteza că $p(x_1, x_2)$ şi $p(x_2)$ sunt funcții continue pe intervalul $(x_2 - \Delta x_2, x_2)$, putem împărți atât numărătorul cât şi numitorul din (1.46) cu $\Delta x_2$ şi trece la limită pentru $\Delta x_2 \to 0$. Obținem astfel

$$\begin{aligned} P(X_1 \le x_1 \mid X_2 = x_2) \equiv F(x_1 \mid x_2) &= \frac{\partial F(x_1, x_2) / \partial x_2}{\partial F(x_2) / \partial x_2} \\ &= \frac{\partial \left[ \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} p(u_1, u_2) \mathrm{d}u_1 \mathrm{d}u_2 \right] / \partial x_2}{\partial \left[ \int_{-\infty}^{x_2} p(u_2) \mathrm{d}u_2 \right] / \partial x_2} \\ &= \frac{\int_{-\infty}^{x_1} p(u_1, x_2) \mathrm{d}u_1}{p(x_2)} \end{aligned} \tag{1.47}$$

Aceasta este FDC a variabilei aleatoare $X_1$ condiționată de variabila aleatoare $X_2$. Observăm că $F(-\infty \mid x_2) = 0$ şi $F(\infty \mid x_2) = 1$. Derivând ecuația (1.47) în raport cu $x_1$, obținem funcția densitate de probabilitate $p(x_1 \mid x_2)$ în forma

$$p(x_1 \mid x_2) = \frac{p(x_1, x_2)}{p(x_2)} \tag{1.48}$$

Putem exprima FDP comună $p(x_1, x_2)$ cu ajutorul funcțiilor densitate de probabilitate condiționate astfel:

$$p(x_1, x_2) = p(x_1 \mid x_2) p(x_2) = p(x_2 \mid x_1) p(x_1) \tag{1.49}$$

## 1.10. FUNCȚII DE O VARIABILĂ ALEATOARE

Fie $X$ o variabilă aleatoare şi $g(x)$ o funcție de variabila reală $x$. Expresia

$$Y = g(X) \tag{1.50}$$

este o nouă variabilă aleatoare definită după cum urmează: pentru un eveniment $\zeta$ dat, $X(\zeta)$ este un număr iar $g[X(\zeta)]$ este un alt număr specificat în funcție de $X(\zeta)$ şi de $g(x)$. Acest număr este valoarea $Y(\zeta) = g[X(\zeta)]$ atribuită variabilei aleatoare $Y$. Deci, o funcție de o variabilă aleatoare $X$ este o funcție compusă $Y = g(X) = g[X(\zeta)\}$ al cărei domeniu de definiție este mulțimea S a rezultatelor experimentale.

Funcția de distribuție $F_Y(y)$ a variabilei aleatoare astfel formate este probabilitatea evenimentului $\{Y \le y\}$ constând din toate rezultatele $\zeta$ astfel încât $Y(\zeta) = g[X(\zeta)] \le y$. Deci

$$F_Y(y) = P\{Y \le y\} = P\{g(X) \le y\} \tag{1.51}$$

Pentru un $y$ particular, valorile lui $x$ astfel încât $g(x) \le y$ formează o mulțime de pe axa $x$ notată cu $R_y$. În mod clar, $g[X(\zeta)] \le y$ dacă $X(\zeta)$ este un număr din mulțimea $R_y$. Prin urmare

$$F_Y(y) = P\{X \in R_y\} \tag{1.52}$$

Conchidem că, pentru ca $g(X)$ să fie o variabilă aleatoare, funcția $g(x)$ trebuie să aibă următoarele proprietăți:

**1.** Domeniul ei de definiție trebuie să includă domeniul de existență al variabilei aleatoare $X$.

**2.** Trebuie să fie o funcție *Borel*, adică, pentru orice $y$, mulțimea $R_y$ astfel încât $g(x) \le y$ trebuie să constea din reuniunea şi intersecția unei mulțimi numărabile de intervale. Numai atunci $\{Y \le y\}$ este un eveniment.

**3.** Evenimentele $\{g(X) = \pm\infty\}$ trebuie să aibe probabilitate zero.

**EXEMPLUL 1.1:** Fie variabila aleatoare $Y$ definită prin

$$Y = aX + b \tag{1.53}$$

unde $a$ şi $b$ sunt constante.

(a) Dacă $a > 0$, avem $ax + b \le y$ pentru $x \le \dfrac{y-b}{a}$. Prin urmare,

$$\begin{aligned} F_Y(y) &= P(Y \le y) = P(aX + b \le y) = P\left(X \le \frac{y-b}{a}\right) \\ &= F_X\left(\frac{y-b}{a}\right) = \int_{-\infty}^{(y-b)/a} p_X(x)\mathrm{d}x \end{aligned} \tag{1.54}$$

Derivând ecuaţia (1.54) în raport cu $y$, obţinem relaţia dintre funcţiile densitate de probabilitate respective:

$$p_Y(y) = \frac{1}{a} p_X\left(\frac{y-b}{a}\right) \tag{1.55}$$

(b) Dacă $a < 0$, avem $ax + b \le y$ pentru $x > (y-b)/a$. Prin urmare,

$$\begin{aligned} F_Y(y) &= P(Y \le y) = P\left(X > \frac{y-b}{a}\right) = 1 - P\left(X \le \frac{y-b}{a}\right) \\ &= 1 - F_X\left(\frac{y-b}{a}\right) \end{aligned} \tag{1.56}$$

**EXEMPLUL 1.2:** Fie variabila aleatoare $Y$ definită prin

$$Y = aX^3 + b, a > 0 \tag{1.57}$$

Funcţia $y = ax^3 + b$ este bijectivă şi, deci, inversabilă. Prin urmare,

$$\begin{aligned} F_Y(y) &= P(Y \le y) = P(aX^3 + b \le y) \\ &= P\left[X \le \left(\frac{y-b}{a}\right)^{1/3}\right] = F_X\left[\left(\frac{y-b}{a}\right)^{1/3}\right] \end{aligned} \tag{1.58}$$

Derivând ecuaţia (1.58) în raport cu $y$, obţinem relaţia dintre cele două FDP:

$$p_Y(y) = \frac{1}{3a[((y-b)/a)^{2/3}]} p_X\left[\left(\frac{y-b}{a}\right)^{1/3}\right] \tag{1.59}$$

**EXEMPLUL 1.3:** Definim variabila aleatoare $Y$ prin

$$Y = aX^2 + b, a > 0 \tag{1.60}$$

Spre deosebire de exemplele precedente, funcția $y = g(x)$ nu este bijectivă. Pentru a determina FDC a lui $Y$, observăm că

$$F_Y(y) = P(Y \le y) = P(aX^2 + b \le y) = P\left(|X| \le \sqrt{\frac{y-b}{a}}\right)$$

Prin urmare,

$$F_Y(y) = F_X\left(\sqrt{\frac{y-b}{a}}\right) - F_X\left(-\sqrt{\frac{y-b}{a}}\right) \tag{1.61}$$

Derivând ecuația (1.61) în raport cu $y$, obținem FDP a lui $Y$ în funcție de FDP a lui $X$ astfel:

$$p_Y(y) = \frac{p_X[\sqrt{(y-b)/a}]}{2a\sqrt{[(y-b)/a]}} + \frac{p_X[-\sqrt{(y-b)/a}]}{2a\sqrt{[(y-b)/a]}} \tag{1.62}$$

În Exemplul 1.3, observăm că ecuația

$$g(x) = ax^2 + b = y$$

are două rădăcini reale,

$$x_1 = \sqrt{\frac{y-b}{a}} \quad \text{şi} \quad x_2 = -\sqrt{\frac{y-b}{a}}$$

şi că $p_Y(y)$ constă din doi termeni corespunzători acestor două soluții.

## 1.11. MEDII STATISTICE ALE VARIABILELOR ALEATOARE

Fie o variabilă aleatoare $X$ de tip discret caracterizată prin distribuția de probabilitate $P(x_i)$, $i = 1, \cdots, n$. Prin definiție, *valoarea medie* sau *aşteptată* a lui $X$ este

$$E(X) \equiv m_x = \sum_{i=1}^{n} x_i P(x_i) \tag{1.63}$$

Fie o variabilă aleatoare *X* de tip continuu caracterizată prin FDP a ei $p(x)$. Prin definiţie, *valoarea medie* sau *aşteptată* a lui *X* este

$$E(X) \equiv m_x = \int_{-\infty}^{\infty} xp(x)\mathrm{d}x \tag{1.64}$$

Operatorul $E(\bullet)$ are o deosebită importanţă în teoria probabilităţii. Notaţia sa prin litera *E* se explică prin aceea că în engleză (*expectation*), în franceză (*espérance*) şi în germană (*Erwartung*), această noţiune fundamentală se exprimă prin cuvinte cu iniţiala *e*. În limba română există cuvântul „expectaţie“ ca termen livresc pentru „aşteptare, speranţă“[1]. Ar fi indicat să-l utilizăm şi noi în teoria probabilităţii şi în teoria informaţiei, fără a pierde din vedere că sensul său este de „lucru, valoare la care trebuie să ne aşteptăm în medie“. Valoarea medie sau expectaţia, definită în (1.64), este şi momentul de ordinul unu al variabilei aleatoare *X*. În general, momentul de ordinul *n* se defineşte prin

$$E(X^n) = \int_{-\infty}^{\infty} x^n p(x)\mathrm{d}x \tag{1.65}$$

Fie acum o variabilă aleatoare $Y = g(X)$, unde $g(X)$ este o funcţie arbitrară de variabila aleatoare *X*. Media (expectaţia) lui *Y* este

$$E(Y) = E[g(X)] = \int_{-\infty}^{\infty} g(x)p(x)\mathrm{d}x \tag{1.66}$$

În particular, dacă $Y = (X - m_x)^n$ unde $m_x$ este valoarea medie a lui *X*, definim *momentul central de ordinul n* al variabilei aleatoare *X* astfel:

$$E(Y) = E[(X - m_x)^n] = \int_{-\infty}^{\infty} (x - m_x)^n p(x)\mathrm{d}x \tag{1.67}$$

Pentru $n = 2$, momentul central se numeşte *varianţa* variabilei aleatoare şi se notează $\sigma_x^2$:

$$\sigma_x^2 = \int_{-\infty}^{\infty} (x - m_x)^2 p(x)\mathrm{d}x \tag{1.68}$$

Rădăcina pătrată a varianţei, $\sigma_x$, se numeşte *deviaţia standard* a lui *X*.

Varianţa este un parametru care ne dă o măsură a dispersiei variabilei aleatoare *X*. Să dezvoltăm termenul $(x - m_x)^2$ din integrala ecuaţiei (1.68):

---

[1] Vezi „Dicţionarul explicativ al limbii române“ (DEX) redactat de Institutul de lingvistică „Iorgu Iordan“şi editat sub egida Academiei Române, p. 358.

$$(x-m_x)^2 = x^2 - 2m_x x + m_x^2$$

Întrucât valoarea medie a unei constante este egală cu constanta, obținem expresia care leagă varianța de momentele de ordinul unu şi doi:

$$\sigma_x^2 = E(X^2) - [E(X)]^2 = E(X^2) - m_x^2 \qquad (1.69)$$

Fie acum două variabile aleatoare $X_1$ şi $X_2$ cu FDP comună $p(x_1, x_2)$. Definim *momentul comun* astfel:

$$E(X_1^k X_2^n) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1^k x_2^n p(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 \qquad (1.70)$$

Să notăm mediile $m_i = E(X_i)$ pentru $i$ = 1, 2. Definim *momentul central comun* astfel:

$$E[(X_1-m_1)^k (X_2-m_2)^n] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_1-m_1)^k (x_2-m_2)^n p(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 \qquad (1.71)$$

De o importanță deosebită pentru noi sunt momentul comun şi momentul central comun corespunzătoare lui $k = n = 1$.

Corelația dintre $X_1$ şi $X_2$ este dată de momentul comun

$$E(X_1 X_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 p(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 \qquad (1.72)$$

Covarianța lui $X_1$ şi $X_2$ este

$$\begin{aligned}\mu_{12} &\equiv E[(X_1-m_1)(X_2-m_2)] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_1-m_1)(x_2-m_2) p(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 \qquad (1.73) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 p(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 - m_1 m_2 = E(X_1, X_2) - m_1 m_2\end{aligned}$$

## 1.12. FUNCȚII CARACTERISTICE

Prin definiție, *funcția caracteristică* a variabilei aleatoare $X$ este media statistică

$$E(e^{jvX}) \equiv \psi(jv) = \int_{-\infty}^{\infty} e^{jvx} p(x)\mathrm{d}x \qquad (1.74)$$

unde variabila $v$ este reală iar $j = \sqrt{-1}$.

Observăm că $\psi(jv)$ seamănă cu transformata Fourier a lui $p(x)$, deosebirea constând în semnul exponentului. Prin analogie cu transformata Fourier inversă, putem scrie:

$$p(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \psi(jv)e^{-jvx}\mathrm{d}v \tag{1.75}$$

Funcţia caracteristică are o relaţie foarte utilă cu momentele variabilei aleatoare. Să derivăm ecuaţia (1.74) în raport cu $v$:

$$\frac{d\psi(jv)}{dv} = j\int_{-\infty}^{\infty} xe^{jvx}p(x)\mathrm{d}x \tag{1.76}$$

Evaluând derivata în $v = 0$, obţinem momentul de ordinul unu, sau media statistică

$$E(X) = m_x = -j\left.\frac{\mathrm{d}\psi(jv)}{\mathrm{d}v}\right|_{v=0} \tag{1.77}$$

Repetând procesul de derivare, găsim că derivata de ordinul $n$ a lui $\psi(jv)$ evaluată în $v = 0$ ne dă momentul de ordinul $n$:

$$E(X^n) = (-j)^n\left.\frac{\mathrm{d}^n\psi(jv)}{\mathrm{d}v^n}\right|_{v=0} \tag{1.78}$$

## 1.13. DISTRIBUŢII DE PROBABILITATE

### Distribuţie binomială

Fie $X$ o variabilă aleatoare discretă care are două valori posibile, să spunem $X = 1$ şi $X = 0$ cu probabilităţi $p$ şi $1 - p$, respectiv. FDP a lui $X$ este arătată în figura 1.6.

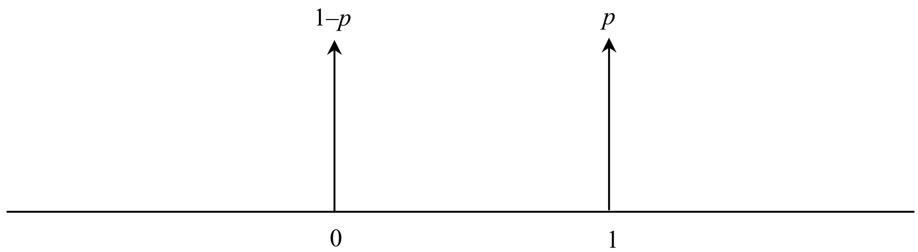


**Fig.1.6.** Funcția distribuție de probabilitate a lui *X*.

Formăm variabila aleatoare

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

unde $X_i, i = 1,2,\cdots,n$ sunt variabile aleatoare distribuite identic și independente statistic cu FDP arătată în figura 1.6. Vrem să determinăm funcția distribuție de probabilitate a lui *Y*.

Observăm că domeniul de existență al lui Y este mulțimea întregilor de la 0 la *n*. Probabilitatea ca $Y = 0$ nu este nimic altceva decât probabilitatea ca toate variabilele aleatoare $X_i$ să fie egale cu zero. Fiindcă aceste $X_i$ sunt independente statistic,

$$P(Y = 0) = (1 - p)^n \tag{1.79}$$

Probabilitatea ca $Y = 1$ este probabilitatea ca una din variabilele aleatoare $X_i = 1$, toate celelalte fiind $X_i = 0$. Întrucât acest eveniment poate surveni în *n* moduri diferite,

$$P(Y = 1) = np(1 - p)^{n-1} \tag{1.80}$$

Probabilitatea ca $Y = k$ este probabilitatea ca *k* din variabilele aleatoare $X_i$ să fie egale cu unu, iar celelalte $n - k$ să fie egale cu zero. Vom nota numărul combinațiilor de *n* obiecte luate câte *k* utilizând simbolul

$$\binom{n}{k} \equiv \frac{n!}{k!(n-k)!} \tag{1.81}$$

Reamintim de la cursul de algebră că $\binom{n}{k}$ este și coeficientul binomial din dezvoltarea binomului lui Newton $(x + y)^n$. Deci

$$P(Y=k)=\binom{n}{k}p^k(1-p)^{n-k} \tag{1.82}$$

Funcția densitate de probabilitate a lui $Y$ se poate exprima astfel:

$$p(y)=\sum_{k=0}^{n}P(Y=k)\delta(y-k)=\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}\delta(y-k) \tag{1.83}$$

Funcția de distribuție cumulativă a lui $Y$ este:

$$F(y)=P(Y\le y)=\sum_{k=0}^{[y]}\binom{n}{k}p^k(1-p)^{n-k} \tag{1.84}$$

unde am notat cu $[y]$ cel mai mare întreg $m$ astfel încât $m\le y$. FDC din ecuația (1.84) caracterizează o funcție aleatoare distribuită binomial.

Să calculăm media statistică (expectația) lui $Y$. Conform definiției (1.63),

$$\begin{aligned}E(Y)=m_y&=\int_{-\infty}^{\infty}yp(y)\mathrm{d}y\\&=\int_{-\infty}^{\infty}y\left[\sum_{k=0}^{n}p^k(1-p)^{n-k}\delta(y-k)\right]dy\\&=\int_{-\infty}^{\infty}\left[\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}y\delta(y-k)\right]\mathrm{d}y\end{aligned} \tag{1.85}$$

Având în vedere că integrarea şi însumarea sunt operații liniare, ordinea lor poate fi inversată:

$$m_y=\sum_{k=0}^{\infty}\int_{-\infty}^{\infty}\binom{n}{k}p^k(1-p)^{n-k}y\delta(y-k)\mathrm{d}y \tag{1.86}$$

Aplicăm acum proprietatea de „filtrare“ a impulsului Dirac:

$$\int_{-\infty}^{\infty}f(x)\delta(x-x_0)\mathrm{d}x=f(x_0) \tag{1.87}$$

Ecuația (1.86), aplicând (1.87), devine:

$$m_y=\sum_{k=0}^{n}\binom{n}{k}kp^k(1-p)^{n-k}=np\sum_{k=1}^{n}\binom{n-1}{k-1}p^{k-1}(1-p)^{n-k} \tag{1. 88}$$

Fie $m=k-1$. Suma din (1.88) se scrie:

$$\sum_{m=0}^{n-1}\binom{n-1}{m}p^m(1-p)^{n-1-m}.$$

Dar aceasta este dezvoltarea binomului lui Newton $(p+1-p)^{n-1}=1$. Prin urmare,

$$m_y = np \tag{1.89}$$

Procedând similar, să calculăm şi momentul de ordinul doi al lui *Y*. Conform definiției (1.65), avem

$$\begin{aligned}
E(Y^2) &= \int_{-\infty}^{\infty} y^2 p(y)\mathrm{d}y = \int_{-\infty}^{\infty} y^2\left[\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}\delta(y-k)\right]\mathrm{d}y\\
&= \sum_{k=0}^{n}\int_{-\infty}^{\infty}\binom{n}{k}p^k(1-p)^{n-k}y^2\delta(y-k)\mathrm{d}y = \sum_{k=1}^{n}\binom{n}{k}k^2p^k(1-p)^{n-k}\\
&= \sum_{k=1}^{n}\frac{n!}{k!(n-k)!}k^2p^k(1-p)^{n-k} = \sum_{k=1}^{n}\frac{n!}{(k-1)!(n-k)!}kp^k(1-p)^{n-k}\\
&= np\left[\sum_{m=0}^{n-1}\frac{(n-1)!}{m!(n-m-1)!}(m+1)p^m(1-p)^{n-m-1}\right]\\
&= np\left[\sum_{m=1}^{n-1}\frac{(n-1)!}{(m-1)!(n-m-1)!}p^m(1-p)^{n-m-1}+\right.\\
&\quad\left.+\sum_{m=0}^{n-1}\frac{(n-1)!}{m!(n-m-1)!}p^m(1-p)^{n-m-1}\right]
\end{aligned} \tag{1.90}$$

Fie $q = m - 1$. Avem atunci

$$\begin{aligned}
E(Y^2) &= np\left[\sum_{q=0}^{n-2}\frac{(n-1)!}{q!(n-q-2)!}p^{q+1}(1-p)^{n-q-2}+\right.\\
&\quad\left.+\sum_{m=0}^{n-1}\frac{(n-1)!}{m!(n-m+1)!}p^m(1-p)^{n-m-1}\right]\\
&= np\left[(n-1)p\sum_{q=0}^{n-2}\frac{(n-2)!}{q!(n-q-2)!}p^q(1-p)^{n-q-2}+\right.\\
&\quad\left.+\sum_{m=0}^{n-1}\frac{(n-1)!}{m!(n-m+1)!}p^m(1-p)^{n-m-1}\right]\\
&= np[(n-1)p+1] = n(n-1)p^2+np = np(1-p)+n^2p^2
\end{aligned}$$

(1-91)

Varianța $\sigma^2$ rezultă imediat din (1.89) şi (1.91) conform ecuației (1.69):

$$\sigma^2 = E(Y^2) - m_y^2 = np(1-p) + n^2 p^2 - n^2 p^2 = np(1-p) \qquad (1.92)$$

Vom găsi acum funcția caracteristică a variabilei aleatoare $Y$. Conform ecuației de definiție (1-74), avem:

$$\psi(jv) = \int_{-\infty}^{\infty} e^{jvy} p(y)\mathrm{d}y = \int_{-\infty}^{\infty} e^{jvy} \left[ \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \delta(y-k) \right] \mathrm{d}y$$

$$= \sum_{k=0}^{n} \int_{-\infty}^{\infty} \binom{n}{k} p^k (1-p)^{n-k} e^{jvy} \delta(y-k)\mathrm{d}y = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} e^{jkv} \qquad (1.93)$$

$$= \sum_{k=0}^{n} \binom{n}{k} (pe^{jv})^k (1-p)^{n-k} = (1-p+pe^{jv})^n$$

### Distribuție uniformă

FDP a unei variabile aleatoare distribuite uniform în intervalul $[a, b]$ este:

$$p(x) = \begin{cases} 0, & x < a \\ \dfrac{1}{b-a}, & a \le x \le b \\ 0, & x > b \end{cases} \qquad (1.94)$$

FDC corespunzătoare este:

$$F(x) = \begin{cases} 0, & x < 0 \\ \dfrac{x-a}{b-a}, & a \le x \le b \\ 0, & x > 0 \end{cases} \qquad (1.95)$$

Media statistică (expectația) a lui $X$ este:

$$E(X) = \int_a^b \frac{x}{b-a}\mathrm{d}x = \frac{1}{b-a} \left. \frac{x^2}{2} \right|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{1}{2}(a+b) \qquad (1.96)$$

Momentul de ordinul doi este:

$$E(X^2) = \int_a^b \frac{x^2}{b-a}\mathrm{d}x = \frac{1}{b-a} \left. \frac{x^3}{3} \right|_a^b = \frac{b^3 - a^3}{3(b-a)} = \frac{1}{3}(a^2 + ab + b^2) \qquad (1.97)$$

Din (1.96) şi (1.97), rezultă imediat:

$$\sigma^2 = \frac{1}{12}(a-b)^2 \tag{1.98}$$

Funcția caracteristică:

$$\psi(jv) = \int_a^b e^{jvx}\frac{1}{b-a}\mathrm{d}x = \frac{1}{b-a}\left.\frac{e^{jvx}}{jv}\right|_a^b = \frac{e^{jvb}-e^{jva}}{jv(b-a)} \tag{1.99}$$

## Distribuție normală (Gaussiană)

FDP a unei variabile aleatoare distribuite normal (Gaussian) este

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-(x-m_x)^2/2\sigma^2} \tag{1.100}$$

unde $m_x$ este media iar $\sigma^2$ este varianța variabilei aleatoare. Din păcate, funcția $p(x)$ din (1.100) nu are o primitivă. Pentru a exprima FDC, definim mai înainte o funcție deosebit de importantă în teoria informației şi a sistemelor de comunicație, funcția *eroare*:

$$fer(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\mathrm{d}t \tag{1.101}$$

Cu aceasta, FDC este

$$\begin{aligned} F(x) &= \int_{-\infty}^x p(u)\mathrm{d}u = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^x e^{-(u-m_x)^2/2\sigma^2}\mathrm{d}u \\ &= \frac{1}{2}\frac{2}{\sqrt{\pi}}\int_{-\infty}^{(x-m_x)/\sqrt{2}\sigma} e^{-t^2}\mathrm{d}t = \frac{1}{2}+\frac{1}{2}fer\left(\frac{x-m_x}{\sqrt{2}\sigma}\right) \end{aligned} \tag{1.102}$$

Definim şi funcția *eroare complementară*:

$$ferc(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-t^2}\mathrm{d}t = 1 - fer(x) \tag{1.103}$$

Cu aceasta, $F(x)$ se mai poate scrie şi astfel:

$$F(x) = 1 - \frac{1}{2}ferc\left(\frac{x-m_x}{\sqrt{2}\sigma}\right) \tag{1.104}$$

Vom arăta acum că $p(x)$ din (1.100) este normată corect în sensul că aria suprafeței de sub curba $y = p(x)$ este egală cu 1. Pentru aceasta, să notăm cu $I$ integrala lui $p(x)$:

$$I \stackrel{\Delta}{=} \int_{-\infty}^\infty p(x)\mathrm{d}x = \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}\sigma}e^{-(x-m_x)^2/2\sigma^2}\mathrm{d}x \tag{1.105}$$

Facem schimbarea de variabilă $u = (x - m_x)/\sigma$ şi avem:

$$I = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-u^2/2}\sigma \mathrm{d}u = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-u^2/2}\mathrm{d}u \tag{1.106}$$

Vom arăta că $I^2$ este egală cu 1:

$$\begin{aligned} I^2 &= \left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-x^2/2}\mathrm{d}x\right]\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-y^2/2}\mathrm{d}y\right] \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)/2}\mathrm{d}x\mathrm{d}y \end{aligned} \tag{1.107}$$

Trecem acum de la coordonate carteziene la coordonate polare. Fie $r^2 = x^2 + y^2$ şi $\vartheta = \mathrm{arctg}(y/x)$. Avem atunci

$$I^2 = \frac{1}{2\pi}\int_0^{2\pi}\left[\int_0^{\infty} e^{-r^2/2} r\mathrm{d}r\right]\mathrm{d}\vartheta = \frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}\vartheta = 1. \tag{1.108}$$

Deci $I^2 = 1$ şi, prin urmare, $I = 1$.

Până acum, am presupus că parametrii $m_x$ şi $\sigma^2$ din (1.100) sunt media şi varianţa distribuţiei. Pentru a arăta că este într-adevăr aşa, să rescriem funcţia lui Gauss utilizând parametrii arbitrari $\alpha$ şi $\beta$:

$$p(x) = \frac{1}{\sqrt{2\pi}\beta} e^{-(x-\alpha)^2/2\beta^2} \tag{1.109}$$

Această funcţie este normată corect, după cum am demonstrat în (1.108). Vom arăta mai întâi că parametrul $\alpha$ este media:

$$m_x = \int_{-\infty}^{\infty} xp(x)\mathrm{d}x = \frac{1}{\sqrt{2\pi}\beta}\int_{-\infty}^{\infty} xe^{-(x-\alpha)^2/2\beta^2}\mathrm{d}x \tag{1.110}$$

Facem schimbarea de variabilă $y = (x - \alpha)/\beta$, obţinând:

$$\begin{aligned} m_x &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(\beta y + \alpha)e^{-y^2/2}\mathrm{d}y \\ &= \frac{\beta}{\sqrt{2\pi}}\int_{-\infty}^{\infty} ye^{-y^2/2}\mathrm{d}y + \alpha\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-y^2/2}\mathrm{d}y \end{aligned} \tag{1.111}$$

Prima integrală din membrul drept al lui (1.111) este zero din cauză că integrandul este o funcţie impară iar integrala este evaluată între limite simetrice. Cea de a doua integrală din membrul drept este integrala unei funcții distribuție de probabilitate Gaussiene normate corect, astfel încât integrala are o valoare egală cu 1. Prin urmare, (1.111) devine:

$$m_x = \alpha \tag{1.112}$$

şi am demonstrat că $\alpha$ este valoarea medie.

Varianța este

$$\begin{aligned}\sigma^2 &= \int_{-\infty}^{\infty} (x - m_x)^2 p(x)\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}\beta}\int_{-\infty}^{\infty} (x - m_x)^2 e^{-(x-m_x)^2/2\beta^2}\mathrm{d}x\end{aligned} \tag{1.113}$$

Trebuie să arătăm că $\sigma^2 = \beta^2$. Pentru aceasta, facem schimbarea de variabilă $y = (x - m_x)/\beta$, obținând:

$$\sigma^2 = \frac{\beta^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} y^2 e^{-y^2/2}\mathrm{d}y \tag{1.114}$$

Fie acum $J = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} y^2 e^{-y^2/2}\mathrm{d}y$. Vom calcula $J^2$:

$$\begin{aligned}J^2 &= \left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x^2 e^{-x^2/2}\mathrm{d}x\right]\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} y^2 e^{-y^2/2}\mathrm{d}y\right] \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2 y^2 e^{-(x^2+y^2)/2}\mathrm{d}x\mathrm{d}y\end{aligned} \tag{1.115}$$

Trecând de la coordonate carteziene la coordonate polare, avem:

$$J^2 = \frac{1}{2\pi}\int_0^{2\pi} \cos^2\theta \sin^2\theta \mathrm{d}\theta \int_0^{\infty} r^5 e^{-r^2/2}\mathrm{d}r \tag{1.116}$$

Prima integrală se scrie succesiv:

$$\frac{1}{4}\int_0^{2\pi} \sin^2 2\theta \mathrm{d}\theta = \frac{1}{8}\int_0^{2\pi}(1 - \cos 4\theta)\mathrm{d}\theta = \frac{\pi}{4} \tag{1.117}$$

Notăm integrala a doua cu $I_r$. Cu schimbarea de variabilă $r^2/2 = t$, avem:

$$I_r = \int_0^{\infty} r^5 e^{-r^2/2}\mathrm{d}r = 4\int_0^{\infty} t^2 e^{-t}\mathrm{d}t \tag{1.118}$$

Vom integra prin părți. Notând $u = t^2$; $v = -e^{-t}$, avem:

$$\int_0^\infty t^2 e^{-t} \mathrm{d}t = t^2\left(-e^{-t}\right)\Big|_0^\infty - \int_0^\infty -e^{-t}(2t)\mathrm{d}t = 2\int_0^\infty te^{-t}\mathrm{d}t \qquad (1.119)$$

Integrăm din nou prin părți:

$$\int_0^\infty te^{-t}\mathrm{d}t = -te^{-t}\Big|_0^\infty - \int_0^\infty -e^{-t}\mathrm{d}t = \int_0^\infty e^{-t}\mathrm{d}t = 1. \qquad (1.120)$$

Prin urmare, $I_r = 4 \cdot 2 = 8$. Rezultă că

$$J^2 = \frac{1}{2\pi} \cdot \frac{\pi}{4} \cdot 8 = 1. \qquad (1121)$$

Din (1.121), rezultă că $J = 1$. Cu această valoare, (1-114) devine $\sigma^2 = \beta^2$.

Funcția caracteristică a unei variabile aleatoare Gaussiene cu medie $m_x$ şi varianță $\sigma_x^2$ este

$$\begin{aligned}
\psi(jv) &= \int_{-\infty}^{\infty} e^{jvx}\left[\frac{1}{\sqrt{2\pi}\sigma_x} e^{-(x-m_x)^2/2\sigma_x^2}\right]\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}\sigma_x}\int_{-\infty}^{\infty} e^{jvx-(x-m_x)^2/2\sigma_x^2}\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}\sigma_x}\int_{-\infty}^{\infty} e^{-[x-(m_x+jv\sigma_x^2)]^2/2\sigma_x^2+jvm_x-(v\sigma_x)^2/2}\mathrm{d}x \qquad (1.122)\\
&= e^{jvm_x-(v^2\sigma_x^2)/2}\left\{\frac{1}{\sqrt{2\pi}\sigma_x}\int_{-\infty}^{\infty} e^{-[x-(m_x+jv\sigma_x^2)^2/2\sigma_x^2}\mathrm{d}x\right\} \\
&= e^{jvm_x-(v^2\sigma_x^2)/2}
\end{aligned}$$

deoarece termenul dintre acolade este egal cu 1, după cum am demonstrat mai sus (vezi (1.108)).

În cazul particular al unei variabile aleatoare normale cu medie zero şi varianță egală cu 1, funcția caracteristică ia forma simplificată

$$\psi(jv) = e^{-v^2/2} \qquad (1.123)$$

**Alte distribuții**

Să reținem că în teoria sistemelor de comunicație se utilizează multe alte distribuții, între care:

- Chi pătrat
- Rayleigh
- Rice
- Nakagami-*m*
- Student-*t*
- Lognormală
- Gama
- Beta.

In economia acestui curs, ele nu îşi au locul, dar cunoştințele acumulate până acum sunt suficiente pentru a ne însuşi aceste distribuții şi altele prin studiu individual, dacă este cazul, recurgând la un manual de teoria probabilității.

## 1.14. INEGALITATEA LUI CEBÎŞEV

Dacă reprezentăm grafic funcția densitate de probabilitate $p(x)$ a unei variabile aleatoare *X*, obținem o curbă ce sugerează un clopot sub care se concentrează „grosul“ probabilității, evenimentele mai puțin probabile formând două „cozi“ laterale (spre -∞ şi spre ∞). Prin *probabilitatea cozii* înțelegem aria suprafeței de sub curba $p(x)$ corespunzătoare cozii. Este adeseori necesar să determinăm această mărime atunci când evaluăm performanța unui sistem de comunicație. Probabilitatea cozii de la dreapta lui *x* este dată de

$$1 - F_X(x) = \int_x^{\infty} p(u)\mathrm{d}u \tag{1.124}$$

Întrucât un calcul exact al integralei din (1.122) nu este întotdeauna posibil, ne interesează o margine superioară a probabilității cozii, care să aproximeze cât mai bine valoarea reală. Începem cu *inegalitatea lui Cebîşev.*

Presupunem că *X* este o variabilă aleatoare arbitrară cu medie finită $m_x$ şi varianță finită $\sigma_x^2$. Pentru orice număr pozitiv δ, avem:

$$P(|X - m_x| \geq \delta) \leq \frac{\sigma_x^2}{\delta^2} \tag{1.125}$$

**Demonstrație:**

$$\sigma_x^2 = \int_{-\infty}^{\infty} (x - m_x)^2 p(x)\mathrm{d}x \geq \int\limits_{|x-m_x|\geq\delta} (x - m_x)^2 p(x)\mathrm{d}x \qquad (1.126)$$
$$\geq \delta^2 \int\limits_{|x-m_x|\geq\delta} p(x)\mathrm{d}x = \delta^2 P(|X - m_x| \geq \delta).$$

Fie variabila aleatoare $Y = X - m_x$ de medie zero şi FDP $p(y)$. Inegalitatea lui Cebîşev este o margine superioară a ariei suprafeței de sub cozile lui $p(y)$, adică, aria lui $p(y)$ în intervalele $(-\infty,-\delta)$ şi $(\delta,\infty)$. Putem deci scrie

$$\begin{aligned} P(|X - m_x| \geq \delta) &= P(|Y| \geq \delta) \\ &= P(Y \leq -\delta) + P(Y \geq \delta) \\ &= F_Y(-\delta) + 1 - F_Y(\delta) \end{aligned} \qquad (1.127)$$

Putem deci exprima inegalitatea lui Cebîşev astfel:

$$1 - [F_Y(\delta) - F_Y(-\delta)] \leq \frac{\sigma_x^2}{\delta^2} \qquad (1.128)$$

sau, echivalent, astfel:

$$1 - [F_X(m_x + \delta) - F_X(m_x - \delta)] \leq \frac{\sigma_x^2}{\delta^2} \qquad (1.129)$$

Inegalitatea lui Cebîşev nu ne oferă o aproximație prea bună. Pentru a vedea de ce, să definim o funcție $g(Y)$ astfel:

$$g(Y) = \begin{cases} 1 & \text{pentru } |Y| \geq \delta \\ 0 & \text{pentru } |Y| < \delta \end{cases} \qquad (1.130)$$

Întrucât variabila aleatoare $g(Y)$ este 0 cu probabilitate $P(|Y| < \delta)$ şi 1 cu probabilitate $P(|Y| \geq \delta)$, valoarea ei medie este

$$E[g(Y)] = P(|Y| \geq \delta) \qquad (1.131)$$

Să încercăm să mărginim $g(Y)$ prin forma pătratică $(Y/\delta)^2$:

$$g(Y) \leq \left(\frac{Y}{\delta}\right)^2 \qquad (1.132)$$

Graficul lui $g(Y)$ şi marginea superioară se arată în figura 1.7.

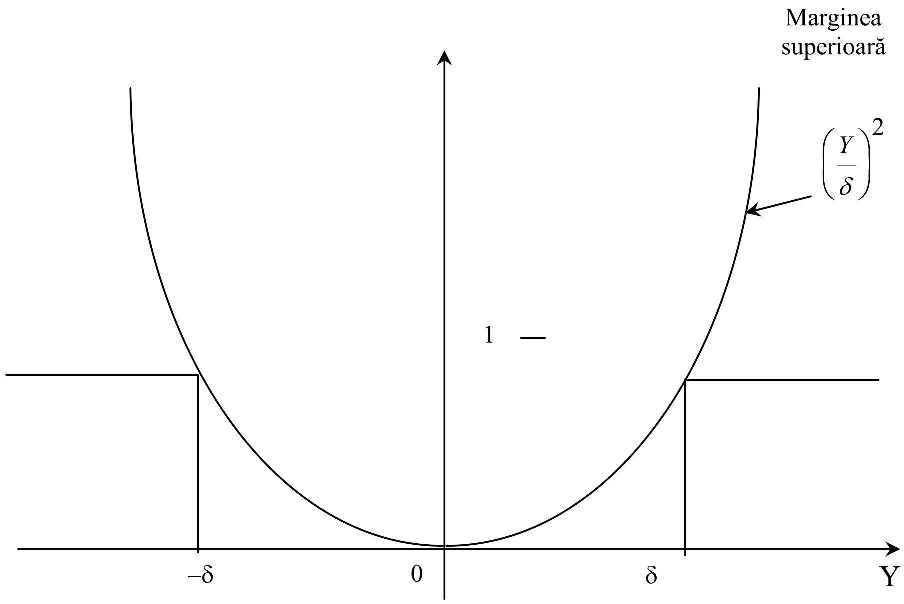


**Fig.1.7.** Margine superioară pătratică a lui $g(Y)$ utilizată pentru a obține probabilitatea cozii (marginea Cebîşev).

Din monotonia mediei, urmează că

$$E[g(Y)] \leq E\left(\frac{Y^2}{\delta^2}\right) = \frac{E(Y^2)}{\delta^2} = \frac{\sigma_y^2}{\delta^2} = \frac{\sigma_x^2}{\delta^2} \tag{1.133}$$

Întrucât $E[g(Y)]$ este probabilitatea cozilor, am regăsit astfel inegalitatea lui Cebîşev.

## 1.15. MARGINEA CHERNOFF

Să presupunem că, pentru o aplicație dată, nu ne interesează decât aria suprafeței de sub una dintre cozi, fie în intervalul $(-\infty,-\delta)$, fie în intervalul $(\delta,\infty)$. Putem obține o aproximație foarte bună dacă mărginim funcția $g(Y)$ printr-o exponențială al cărei parametru poate fi optimizat. Fie probabilitatea cozii în intervalul $(\delta, \infty)$. Definim deci funcția $g(Y)$ astfel:

$$g(Y) = \begin{cases} 1 & \text{pentru } Y \geq \delta \\ 0 & \text{pentru } Y < \delta \end{cases} \tag{1.134}$$

Fie $\nu \geq 0$ parametrul ce urmează a fi optimizat pentru a obține o aproximație cât mai bună. Mărginim superior funcția $g(Y)$ astfel:

$$g(Y) \leq e^{\nu(Y-\delta)} \tag{1.135}$$

Graficul lui $g(Y)$ şi exponențiala sunt arătate în figura 1.8.

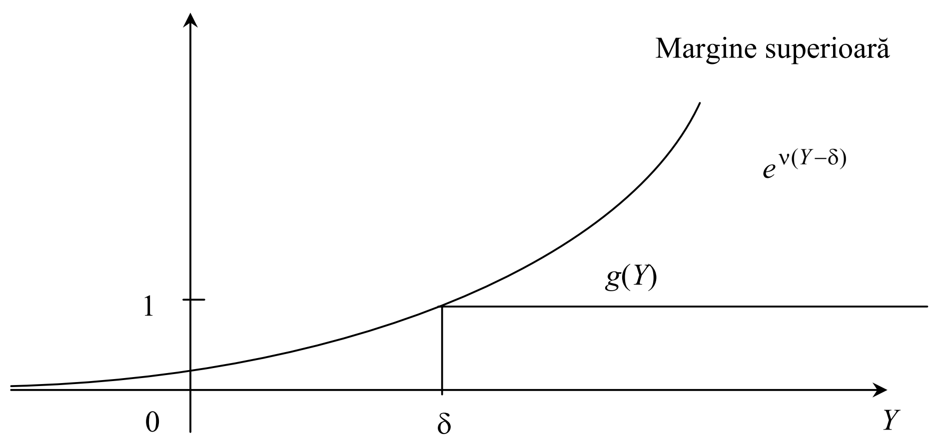


**Fig.1.8.** Margine superioară exponențială a lui $g(Y)$ utilizată pentru a obține probabilitatea cozii (marginea Chernoff).

Valoarea medie (expectația) lui $g(Y)$ este

$$E[g(Y)] = P(Y \geq \delta) \leq E\left(e^{\nu(Y-\delta)}\right) \tag{1.136}$$

Această margine este valabilă pentru orice $\nu \geq 0$. Cea mai bună aproximație se obține minimizând $E\left(e^{\nu(Y-\delta)}\right)$. Pentru a avea un minim, o condiție necesară este ca

$$\frac{\mathrm{d}}{\mathrm{d}\nu} E\left(e^{\nu(Y-\delta)}\right) = 0 \tag{1.137}$$

Schimbând ordinea de efectuare a derivării şi a integrării, avem succesiv

$$\frac{\mathrm{d}}{\mathrm{d}\nu}E\left(e^{\nu(Y-\delta)}\right)=E\left(\frac{\mathrm{d}}{\mathrm{d}\nu}e^{\nu(Y-\delta)}\right)$$
$$=E\left[(Y-\delta)e^{\nu(Y-\delta)}\right]=e^{-\nu\delta}\left[E\left(Ye^{\nu Y}\right)-\delta E\left(e^{\nu Y}\right)\right]=0$$

Valoarea lui ν care ne dă cea mai bună aproximaţie este, deci, soluţia ecuaţiei

$$E\left(Ye^{\nu Y}\right)-\delta E\left(e^{\nu Y}\right)=0 \tag{1.138}$$

Fie $\hat{\nu}$ soluţia ecuaţiei (1.138). Din (1.136), rezultă că marginea superioară a probabilităţii uneia dintre cozi este

$$P(Y\geq\delta)\leq e^{-\hat{\nu}\delta}E\left(e^{\hat{\nu}Y}\right) \tag{1.139}$$

Aceasta este marginea Chernoff pentru probabilitatea cozii din dreapta în cazul unei variabile aleatoare de tip discret sau continuu având medie zero.

Pentru ν real, $E\left(e^{\nu Y}\right)$ se numeşte *funcţia generatoare de momente* a lui *Y*.

## 1.16. SUME DE VARIABILE ALEATOARE

Fie *n* variabile aleatoare distribuite identic şi independente statistic $X_i,\ i=1,2,\cdots,n,$ fiecare având o medie $m_x$ finită şi o varianţă $\sigma_x^2$ finită. Definim o nouă variabilă aleatoare *Y* drept sumă normată, denumită *medie a eşantioanelor*:

$$Y=\frac{1}{n}\sum_{i=1}^{n}X_i \tag{1.140}$$

Variabila aleatoare *Y* definită în (1-140) se întâlneşte frecvent la estimarea mediei unei variabile aleatoare *X* dintr-un număr de observaţii $X_i,\ i=1,2,\cdots,n$.

Media lui *Y* este

$$E(Y) = m_y = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = m_x \tag{1.141}$$

Varianța lui $Y$ este

$$\begin{aligned}
\sigma_y^2 &= E(Y^2) - m_y^2 = E(Y^2) - m_x^2 \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} E(X_i X_j) - m_x^2 \\
&= \frac{1}{n^2}\sum_{i=1}^{n} E(X_i^2) + \frac{1}{n^2}\sum_{\substack{i=1 \\ i \neq j}}^{n}\sum_{j=1}^{n} E(X_i)E(X_j) - m_x^2 \\
&= \frac{1}{n}(\sigma_x^2 + m_x^2) + \frac{1}{n^2} n(n-1)m_x^2 - m_x^2 = \frac{\sigma_x^2}{n}.
\end{aligned} \tag{1.142}$$

Dacă se consideră $Y$ drept o estimație pentru media $m_x$, observăm că expectația sa este egală cu $m_x$ iar varianța sa descreşte cu numărul de eşantioane $n$. Când $n$ tinde spre infinit, varianța $\sigma_y^2$ tinde spre zero.

O estimație a unui parametru (în acest caz media $m_x$) ce satisface condițiile ca expectația sa să conveargă spre valoarea adevărată a parametrului iar varianța sa să conveargă spre zero când $n \to \infty$ se spune că este o *estimație compatibilă*.

Să aplicăm inegalitatea lui Cebîşev la $Y$. Avem

$$\begin{gathered}
P(|Y - m_y| \geq \delta) \leq \frac{\sigma_y^2}{\delta^2} \\
P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - m_x\right| \geq \delta\right) \leq \frac{\sigma_x^2}{n\delta^2}
\end{gathered} \tag{1.143}$$

La limită când $n \to \infty$ (1.143) devine:

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - m_x\right| \geq \delta\right) = 0 \tag{1.144}$$

Rezultă că probabilitatea ca estimația mediei să difere de adevărata medie $m_x$ cu mai mult decât $\delta\,(\delta > 0)$ tinde către zero când $n$ tinde la infinit. Aceasta este o formă a legii numerelor mari. Având în vedere că marginea superioară converge la zero relativ lent (invers cu $n$), expresia din (1.143) se numeşte *legea slabă a numerelor mari*.

Vom aplica acum marginea Chernoff care ne oferă o aproximație mai bună a probabilității unei cozi.

$$
\begin{aligned}
P(Y-m_y \geq \delta) &= P\left(\frac{1}{n}\sum_{i=1}^{n} X_i - m_x \geq \delta\right) \\
&= P\left(\sum_{i=1}^{n} X_i \geq n\delta_m\right) \leq E\left\{\exp\left[\nu\left(\sum_{i=1}^{n} X_i - n\delta_m\right)\right]\right\}
\end{aligned} \tag{1.145}
$$

unde $\delta_m = m_x + \delta$ iar $\delta > 0$. Dar variabilele aleatoare $X_i, i = 1, 2, \cdots, n$ sunt statistic independente şi distribuite identic, aşa că

$$
\begin{aligned}
E\left\{\exp\left[\nu\left(\sum_{i=1}^{n} X_i - n\delta_m\right)\right]\right\} &= e^{-\nu n\delta_m} E\left[\exp\left(\nu\sum_{i=1}^{n} X_i\right)\right] \\
&= e^{-\nu n\delta_m}\prod_{i=1}^{n} E(e^{\nu X_i}) = \left[e^{-\nu\delta_m} E(e^{\nu X})\right]^n
\end{aligned} \tag{1.146}
$$

unde $X$ este oricare din $X_i$. Parametrul $\nu$ care ne dă marginea superioară cea mai strânsă se obține derivând (1.146) şi făcând derivata egală cu zero. Aceasta ne dă ecuația

$$E(Xe^{\nu X}) - \delta_m E(e^{\nu X}) = 0 \tag{1.147}$$

Notăm cu $\hat{\nu}$ soluția ecuației (1.147). Marginea probabilității cozii din dreapta este atunci

$$P\left(\frac{1}{n}\sum_{i=1}^{n} X_i \geq \delta_m\right) \leq \left[e^{-\hat{\nu}\delta_m} E(e^{\hat{\nu}X})\right]^n, \quad \delta_m > m_x \tag{1.148}$$

**EXEMPLUL 1.4:** Fie $X_i$, $i = 1, 2, \cdots, n$ o mulțime de variabile aleatoare independente statistic definite astfel

$$
X_i = \begin{cases} 1 & \text{cu probabilitate} \quad p < \dfrac{1}{2} \\ -1 & \text{cu probabilitate} \quad 1-p \end{cases}
$$

Vrem să determinăm o margine superioară strânsă a probabilității ca suma acestor $X_i$ să fie mai mare decât zero. Întrucât $p < \frac{1}{2}$, observăm că suma va avea o valoare negativă drept medie; de aceea, ne interesează probabilitatea cozii din dreapta. Făcând $\delta_m = 0$ în (1.148), avem

$$P\left(\sum_{i=1}^{n} X_i \geq 0\right) \leq \left[E(e^{\hat{\nu}X})\right]^n \tag{1.149}$$

unde $\hat{\nu}$ este soluția ecuației

$$E(Xe^{\nu X}) = 0 \tag{1.150}$$

Dar

$$E(Xe^{\nu X}) = -(1-p)e^{-\nu} + \nu e^{\nu} = 0 \tag{1.151}$$

De unde

$$\hat{\nu} = \ln\left(\sqrt{\frac{1-p}{p}}\right) \tag{1.152}$$

De asemenea,

$$E(e^{\hat{\nu}X}) = pe^{\hat{\nu}} + (1-p)e^{-\hat{\nu}} \tag{1.153}$$

Marginea Chernoff devine

$$\begin{aligned} P\left(\sum_{i=1}^{n} X_i \geq 0\right) &\leq \left[pe^{\hat{\nu}} + (1-p)e^{-\hat{\nu}}\right]^n \\ &\leq \left[p\sqrt{\frac{1-p}{p}} + (1-p)\sqrt{\frac{p}{1-p}}\right]^n \leq [4p(1-p)]^{n/2} \end{aligned} \tag{1.154}$$

Se observă că marginea superioară descreşte exponențial cu $n$, aşa cum şi era de aşteptat.

## 1.17. TEOREMA LIMITĂ CENTRALĂ

Fie $n$ variabile aleatoare distribuite identic şi independente statistic $X_i$, $i = 1, 2, \cdots, n$, fiecare având o medie $m_x$ finită şi o varianță $\sigma_x^2$ finită. Definim variabilele aleatoare normate:

$$U_i = \frac{X_i - m_x}{\sigma_x}, \; i = 1, 2, \cdots, n \tag{1.155}$$

Se vede că $U_i$ are medie zero şi varianță egală cu unu. Fie

$$Y = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} U_i \qquad (1.156)$$

Funcția caracteristică a lui $Y$ este

$$\begin{aligned} \psi_Y(jv) &= E(e^{jvY}) = E\left[\exp\left(\frac{jv\sum_{i=1}^{n} U_i}{\sqrt{n}}\right)\right] \\ &= \prod_{i=1}^{n} \psi_{U_i}\left(\frac{jv}{\sqrt{n}}\right) = \left[\psi_U\left(\frac{jv}{\sqrt{n}}\right)\right]^n \end{aligned} \qquad (1.157)$$

unde $U$ este oricare din variabilele aleatoare $U_i$, care sunt distribuite identic. Dezvoltăm funcția caracteristică în serie Taylor:

$$\Psi_U\left(j\frac{v}{\sqrt{n}}\right) = 1 + j\frac{v}{\sqrt{n}}E(U) - \frac{v^2}{n\,2!}E(U^2) + \frac{(jv)^3}{\left(\sqrt{n}\right)^3 3!}E(U^3) - \cdots \qquad (1.158)$$

Întrucât $E(U) = 0$ şi $E(U^2) = 1,$ (1.158) se simplifică devenind

$$\Psi_U\left(\frac{jv}{\sqrt{n}}\right) = 1 - \frac{v^2}{2n} + \frac{1}{n}R(v,n) \qquad (1.159)$$

unde $R(v,n)/n$ este restul. Observăm că $R(v,n)$ tinde la zero când $n \to \infty$. Înlocuind (1.159) în (1.157), obținem funcția caracteristică a lui $Y$ în forma

$$\Psi_Y(jv) = \left[1 - \frac{v^2}{2n} + \frac{R(v,n)}{n}\right]^n \qquad (1.160)$$

Luând logaritmul natural în (1.160), obținem

$$\ln \Psi_Y(jv) = n \ln\left[1 - \frac{v^2}{2n} + \frac{R(v,n)}{n}\right] \qquad (1.161)$$

Pentru valori mici ale lui $x$, $\ln(1+x)$ se poate dezvolta în seria de puteri

$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \cdots \qquad (1.162)$$

Aplicând această dezvoltare ecuaţiei (1-161), avem

$$\ln \psi_Y(jv) = n\left[ -\frac{v^2}{2n} + \frac{R(v,n)}{n} - \frac{1}{2}\left( -\frac{v^2}{2n} + \frac{R(v,n)}{n} \right)^2 + \cdots \right] \qquad (1.163)$$

Trecem acum la limită când $n \to \infty$.

$$\lim_{n\to\infty} \ln \psi_Y(jv) = -\frac{1}{2}v^2 \qquad (1.164)$$

Aceasta echivalează cu

$$\lim_{n\to\infty} \psi_Y(jv) = e^{-v^2/2} \qquad (1.165)$$

Dar aceasta este chiar funcţia caracteristică a unei variabile aleatoare normale (Gaussiene) cu medie zero şi varianţă egală cu unu (vezi(1.122)). Am obţinut astfel un rezultat cunoscut drept *teorema limită centrală*: suma unor variabile aleatoare independente statistic şi distribuite identic cu medie finită şi varianţă finită tinde către o funcţie de distribuţie cumulativă Gaussiană când $n \to \infty$.

## 1.18. PROCESE STOCHASTICE

Până acum, vorbind despre experimente repetate ca aruncarea zarului sau a monedei de *n* ori, nu am luat niciodată factorul timp expres în consideraţie. În realitate, numeroase fenomene aleatoare, atât naturale cât şi artificiale, sunt funcţii de timp. În particular, diversele mărimi aleatoare ce intervin în teoria transmisiunii informaţiei sunt funcţii de timp.

Stochastic este un adjectiv care îşi are etimologia într-un verb din greaca veche având sensul de „a ghici, a conjectura” şi înseamnă aleator, adică întâmplător. În orice moment de timp dat, un proces stochastic este caracterizat printr-o variabilă aleatoare indexată cu parametrul *t*. Vom nota un proces stochastic cu $X(t)$. În general, parametrul *t* este continuu, dar $X$ poate fi de tip continuu sau discret, în funcţie de caracteristicile sursei care generează procesul stochastic.

Tensiunea de zgomot generată de un singur rezistor sau o singură sursă de informație reprezintă o singură realizare a procesului stochastic. De aceea, se numeşte o *funcție eşantion* a procesului stochastic. Mulțimea tuturor funcțiilor eşantion posibile, de exemplu mulțimea tuturor formelor de undă ale tensiunii de zgomot generate de rezistoarele dintr-un lot, constituie un *ansamblu* de funcții eşantion sau, echivalent, procesul stochastic $X(t)$. Numărul funcțiilor eşantion din ansamblu poate fi foarte mare, chiar şi infinit.

Un exemplu de proces stochastic este zgomotul generat de rezistoarele, nominal identice, dintr-un lot de fabricație, ilustrat în figura 1.9.

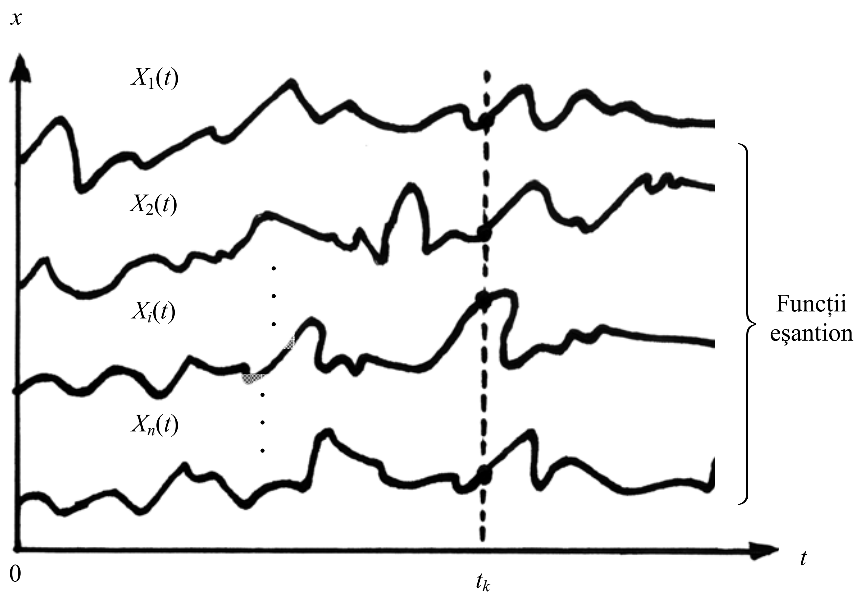


**Fig. 1.9.** Proces stochastic reprezentat de tensiunile de zgomot generate de un lot de rezistoare.

## Procese stochastice staționare

Atunci când ne ocupăm cu procese aleatoare întâlnite în realitate, constatăm adeseori că proprietățile statistice ale unui proces sunt independente de momentul în care se începe observarea procesului. Cu alte cuvinte, dacă se împarte un astfel de proces într-un număr de intervale de timp, diversele secțiuni ale procesului prezintă aceleaşi proprietăți statistice.

Un astfel de proces se spune că este *staționar*. În caz contrar, se spune că este *nestaționar*.

Să exprimăm aceasta mai riguros. Variabilele aleatoare $X_{t_i}$, $i = 1, 2, \cdots, n$, obținute de la un proces stochastic $X(t)$ pentru orice mulțime de momente de timp $t_1 \succ t_2 \succ t_3 \succ \cdots \succ t_n$, unde $n$ este orice număr natural, sunt caracterizate statistic de FDP comună $p(x_{t_1}, x_{t_2}, \cdots, x_{t_n})$. Să considerăm o altă mulțime de $n$ variabile aleatoare $X_{t_i+t} \equiv X(t_i + t)$, $i = 1, 2, \cdots, n$, unde $t$ este o deplasare de timp arbitrară. Aceste variabile aleatoare sunt caracterizate de FDP comună $p(x_{t_1+t}, x_{t_2+t}, \cdots, x_{t_n+t})$. Dacă

$$p(x_{t_1}, x_{t_2}, \cdots, x_{t_n}) = p(x_{t_1+t}, x_{t_2+t}, \cdots, x_{t_n+t}) \tag{1.166}$$

pentru orice $t$ şi orice $n$, procesul stochastic se spune că este *staționar în sens strict*.

### Media

Prin definiție, *media* procesului $X(t)$ este expectația variabilei aleatoare obținute observând procesul la un timp $t$:

$$\mu_X(t) = E[X(t)] = \int_{-\infty}^{\infty} x p_{X(t)}(x) \mathrm{d}x \tag{1.167}$$

unde $p_{X(t)}(x)$ este funcția densitate de probabilitate de ordinul unu a procesului. Pentru un proces aleator strict staționar, $p_{X(t)}(x)$ este independentă de timpul $t$. Drept urmare, media unui proces strict staționar este o constantă:

$$\mu_X(t) = \mu_X \text{ pentru orice } t \tag{1.168}$$

### Funcția de autocorelație

Funcția de autocorelație a procesului $X(t)$ este expectația produsului dintre două variabile aleatoare $X(t_1)$ şi $X(t_2)$ obținute observând procesul $X(t)$ la timpii $t_1$ şi $t_2$, respectiv:

$$\begin{aligned} R_X(t_1, t_2) &= E[X(t_1)X(t_2)] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 p_{X(t_1),X(t_2)}(x_1, x_2) \mathrm{d}x_1 \mathrm{d}x_2 \end{aligned} \tag{1.169}$$

unde $p_{X(t_1),X(t_2)}(x_1,x_2)$ este funcția densitate de probabilitate de ordinul doi a procesului. Pentru un proces aleator strict staționar, $p_{X(t_1),X(t_2)}(x_1,x_2)$ depinde numai de diferența dintre timpii de observație $t_1$ şi $t_2$. Aceasta implică faptul că funcția de autocorelație a unui proces strict staționar depinde numai de diferența de timp $t_2 - t_1$:

$$R_X(t_1,t_2) = R_X(t_2 - t_1) \text{ pentru orice } t_1 \text{ şi } t_2 \tag{1.170}$$

Să redefinim funcția de autocorelație a unui proces staționar $X(t)$ astfel:

$$R_X(\tau) = E[X(t+\tau)X(t)] \text{ pentru orice } t \tag{1.171}$$

Această funcție de autocorelație are următoarele proprietăți importante:

**1.** Valoarea medie pătratică a procesului se obține din $R_X(\tau)$ pentru $\tau = 0$:

$$R_X(0) = E[X^2(t)] \tag{1.172}$$

**2.** Funcția de autocorelație $R_X(\tau)$ este o funcție pară de $\tau$:

$$R_X(\tau) = R_X(-\tau) \tag{1.173}$$

Această proprietate rezultă direct din ecuația de definiție (1.171). Aceasta înseamnă că am fi putut defini $R_X(\tau)$ şi astfel:

$$R_X(\tau) = E[X(t)X(t-\tau)]$$

**3.** Funcția de autocorelație $R_X(\tau)$ este maximă pentru $\tau = 0$:

$$|R_X(\tau)| \le R_X(0) \tag{1.174}$$

Pentru a demonstra această proprietate, să considerăm mărimea nenegativă

$$E[(X(t+\tau) \pm X(t))^2] \ge 0$$

Ridicând la pătrat şi luând expectațiile termenilor, obținem

$$E[X^2(t+\tau)] \pm 2E[X(t+\tau)X(t)] + E[X^2(t)] \ge 0$$

Având în vedere (1.171) şi (1.172), aceasta se reduce la relația

$$2R_X(0) \pm 2R_X(\tau) \ge 0$$

Echivalent, aceasta se poate scrie

$$-R_X(0) \le R_X(\tau) \le R_X(0)$$

din care rezultă (1.173).

**Funcţia de autocovarianţă**

Funcţia de autocovarianţă a unui proces strict staţionar $X(t)$ se scrie:

$$\begin{aligned} C_X(t_1,t_2) &= E[(X(t_1)-\mu_X)(X(t_2)-\mu_X)] \\ &= R_X(t_2-t_1)-\mu_X^2 \end{aligned} \tag{1.175}$$

# 1.19. PROCESE ERGODICE

Există două feluri de a efectua medii ale unui proces stochastic $X(t)$. Primul este de a considera o „secţiune prin proces“. Spre exemplu, media unui proces stochastic $X(t)$ la un timp $t_k$ fixat este expectaţia variabilei aleatoare $X(t_k)$ ce descrie *toate valorile posibile* ale funcţiilor eşantion ale procesului observate la timpul $t=t_k$. Al doilea este de a defini medii în timp "de-a lungul procesului". Întrucât medierea în timp reprezintă un mijloc practic care ne stă la dispoziţie pentru estimarea medierilor pe ansamblu, ne interesează să facem o legătură între medierea pe ansamblu şi medierea în timp. Când, însă, avem voie să înlocuim medierea pe ansamblu cu medierea în timp? Pentru a răspunde la această întrebare, să considerăm funcţia eşantion $x(t)$ a unui proces staţionar $X(t)$. Definim intervalul de observaţie astfel: $-T \le t \le T$. Valoarea de curent continuu (c.c.) a lui $x(t)$ este prin definiţie media în timp

$$\mu_x(T)=\frac{1}{2T}\int_{-T}^{T} x(t)\mathrm{d}t \tag{1.176}$$

Media în timp $\mu_x(T)$ este o variabilă aleatoare, căci valoarea ei depinde de intervalul de observaţie şi de funcţia eşantion particulară a procesului stochastic $X(t)$ ce se utilizează în ecuaţia (1.176). În ipoteza că procesul $X(t)$ este staţionar, media mediei în timp $\mu_x(T)$, după inversarea ordinii de efectuare a operaţiilor de expectaţie şi de integrare, este dată de

$$E[\mu_x(T)]=\frac{1}{2T}\int_{-T}^{T} E[x(t)]\mathrm{d}t=\frac{1}{2T}\int_{-T}^{T}\mu_X\,\mathrm{d}t=\mu_X \tag{1.177}$$

unde $\mu_X$ este media procesului $X(t)$. În mod corespunzător, $\mu_x(T)$ reprezintă o estimație nedeplasată a mediei mediate pe ansamblu $\mu_X$. Spunem că procesul $X(t)$ este *ergodic în medie* dacă sunt satisfăcute două condiții:

**1**. Media în timp $\mu_x(T)$ tinde către media pe ansamblu $\mu_X$ la limită pentru intervalul de observație tinzând la infinit:

$$\lim_{T\to\infty} \mu_x(T) = \mu_X$$

**2.** Varianța lui $\mu_x(T)$, considerată drept variabilă aleatoare, tinde către zero la limită pentru intervalul de observație tinzând la infinit

$$\lim_{T\to\infty} \mathrm{var}[\mu_x(T)] = 0$$

Definim funcția de *autocorelație mediată în timp* a unei funcții eşantion $x(t)$ astfel:

$$R_x(\tau, T) = \frac{1}{2T}\int_{-T}^{T} x(t+\tau)x(t)\mathrm{d}t \qquad (1./178)$$

Considerăm $R_x(\tau, T)$ drept o variabilă aleatoare, având medie şi varianță. Spunem că procesul $X(t)$ este ergodic în funcția de autocorelație dacă sunt îndeplinite următoarele două condiții:

$$\lim_{T\to\infty} R_x(\tau, T) = R_X(\tau)$$

$$\lim_{T\to\infty} \mathrm{var}[R_x(\tau, T)] = 0$$

Cu aceste cunoştințe minime de teoria probabilității, vom aborda în capitolul următor obiectul nostru de studiu propriu-zis, care este teoria matematică a informației.