TEORIA PROBABILITĂŢILOR ŞI STATISTICĂ MATEMATICĂ

Suport de curs

Eugen Păltănea

Universitatea Transilvania din Braşov
2018

Cuprins

0.1 Introducere

0.1.1 Date istorice

Teoria probabilităţilor[1] şi statistica matematică studiază fenomenele aleatoare, în scopul predicţiei şanselor de realizare a unor rezultate aşteptate. Studiul se bazează pe modelare matematică şi raţionamente deductive.

Bazele teoriei probabilităţilor şi statisticii (studiului matematic al "hazardului") au fost puse în secolele XVI-XVII. Astfel, Gerolamo Cardano scria în 1526 prima carte dedicată unor concepte probabiliste: *Liber de ludo aleae* ("Cartea jocurilor de noroc"), publicată în 1663. Contribuţii majore la punerea bazelor teoriei probabilităţilor au avut matematicienii francezi Blaise Pascal şi Pierre de Fermat (începutul sec. al XVII-lea). În 1657, Cristiaan Huygens publica lucrarea *De ratiociniis in ludo aleae* ("Asupra raţionamentelor în jocurile de noroc"). Doctrina probabilităţilor se conturează în sec. al XVIII-lea prin lucrările fundementale *Ars conjectandi*, publicată în 1713 de Jacob Bernoulli şi *The Doctrine of Chances*, publicată în 1718 de Abraham de Moivre. Un secol mai târziu, Pierre Simon Laplace publică lucrarea *Théorie analytique des probabilitées*, în care evidenţiază *legea erorilor* (convergenţa la *legea normală*). Teoria modernă a probabilităţilor (incluzând formalizarea axiomatică) este datorată în bună măsură matematicianului rus Andrei Nikolaevici Kolmogorov, autorul cărţii *Foundation of the Theory of Probability*, publicată în 1950. Printre lucrările de referinţă în domeniu ale secolului XX trebuiesc amintite *The Theory of Stochastic Processes* (D.R Cox, H.D. Miller, 1965), *An Introduction to Probability Theory* (W. Feller, 1968) şi *Probability and Measure* (P. Billingsley, 1979). In ultimii 60 de ani, studiul proceselor aleatoare cunoaşte o dezvoltare şi diversificare teoretică exponenţială, dar şi un câmp foarte larg de aplicabilitate.

Şcoala românească de probabilităţi se remarcă prin numeroase contribuţii originale. Bazele acestei şcoli se datorează în principal matematicienilor C. T. Ionescu-Tulcea, Gheorghe Mihoc, Octav Onicescu, Ion Cuculescu, Marius Iosifescu şi Constantin Tudor.

0.1.2 Descrierea cursului

Prezentul curs reprezintă o introducere în teoria probabilităţilor şi statisticii matematice. Sunt descrise şi caracterizate concepte de bază ale domeniului: câmpuri de probabilitate, variabile aleatoare şi convergenţa şirurilor de variabile aleatoare, eşantioane, estimatori statistici etc. Pentru studiul acestor noţiuni sunt necesare cunoştinţe elementare de analiză matematică şi teoria măsurii. Cunoştinţele teoretice sunt exemplificate şi aplicate în probleme şi lucrări de laborator, menite să familiarizeze cursantul cu raţionamentele specifice domeniului şi cu utilizarea unor softuri informatice.

0.1.3 Principiile de bază. Limbaj

Teoria probabilităţilor analizează şansele de realizare a unor *evenimente* în urma unei *experienţe* şi este fundamentată pe următoarele trei principii:

1. Fiecărui *eveniment* rezultat în urma unei *experienţe* i se atribuie un număr real din intervalul $[0, 1]$ numit *probabilitatea evenimentului.*

[1] Etimologie: *probabilitas* (lat.) = credibilitate

2. Oricare două *evenimente contrare* au suma probabilităţilor egală cu 1.

3. Probabilitatea de realizare simultană a două evenimente este egală cu produsul dinte probabilitatea realizării unuia dintre evenimente şi probabilitatea realizării celuilalt eveniment, condiţionată de realizarea primului eveniment.

In formalizarea matematică, evenimentele care pot rezulta în urma efectuării unei experienţe reprezintă submulţimi ale unei mulţimi Ω (interpretată ca ”evenimentul sigur”). Aceste evenimente au o probabilitate de apariţie (şansă de realizare). Matematic, probabilitatea este o *măsură finită*, definită pe mulţimea evenimentelor. Astfel, limbajului probabilistic îi corespunde limbajul mulţimilor.

Următoarea schemă evidenţiază dualitatea de limbaj (relativ la o anumită experienţă):

- *evenimentul sigur* = mulţimea totală (de referinţă): Ω;
- *evenimentul imposibil* = mulţimea vidă: $\emptyset$;
- *evemiment* = submulţime a mulţimii totale: $A \subset \Omega$;
- *realizarea unui eveniment asigură realizarea unui alt eveniment (implicaţia evenimentelor)* = incluziunea mulţimilor: $A \subset B$;
- *evenimentul contrar (unui eveniment)* = complementara unei submulţimi a mulţimii totale: $A^c = \Omega \setminus A$;
- *evenimentul realizării a cel puţin unuia dintre două evenimente (disjuncţia evenimentelor)* = reuniunea mulţimilor: $A \cup B$;
- *evenimentul realizării simultane a două evenimente (conjuncţia evenimentelor)* = intersecţia mulţimilor: $A \cap B$;

Capitolul 1

Spaţii măsurabile. Câmpuri de probabilitate. Scheme clasice de probabilitate

1.1 Introducere

Vom defini şi caracteriza în continuare noţiunea de spaţiu măsurabil (câmp de evenimente) şi funcţiile măsurabile, definite între spaţii măsurabile. Noţiunile respective modelează matematic rezultatele posibile ale unei experienţe (evenimentele posibile) şi corespondenţele specifice asociate unor experienţe. Astfel, definim noţiunile: corp, σ-corp, borelianul unui spaţiu topologic, spaţiu măsurabil, funcţie măsurabilă şi evidenţiem proprietăţile lor de bază. Se introduce noţiunea de *câmp de probabilitate* şi se pregăteşte definirea conceptului de *variabilă aleatoare*. Sunt evidenţiate proprietăţile probabilităţii, definită ca măsură finită pe spaţiul evenimentelor (σ-corpul evenimentelor). Este discutată independenţa evenimentelor şi noţiunea de probabilitate condiţionată. Exemplele elementare de câmpuri de probabilitate discrete sunt particularizate în modelele clasice (scheme de probabilitate), adaptate studiului unor fenomene aleatoare.

1.2 Corpuri

Noţiune de *corp* (de mulţimi) reprezintă cadrul minimal de modelare matematică a setului evenimentelor posibile rezultate în urma unei experienţe cu rezultate aleatoare.

Definiţia 1.2.1. *Fie Ω o mulţime nevidă. O familie nevidă $\mathcal{C}$ de submulţimi ale lui Ω ($\mathcal{C} \subset \mathcal{P}(\Omega)$) se numeşte corp a lui Ω dacă satisface proprietăţile:*

C_1. $A^c \in \mathcal{C},\ \forall\ A \in \mathcal{C}$;

C_2. $A \cup B \in \mathcal{C},\ \forall\ A, B \in \mathcal{C}$.

Definiţia de mai sus are o serie de consecinţe imediate.

Propoziţia 1.2.1. *Fie $\mathcal{C}$ un corp a lui Ω. Au loc proprietăţile următoare:*

C_3. $\cup_{i=1}^{n} A_i \in \mathcal{C},\ \forall\ A_i \in \mathcal{C},\ i = \overline{1,n},\ n \geq 2$;

C_4. $\cap_{i=1}^n A_i \in \mathcal{C},\ \forall\ A_i \in \mathcal{C},\ i = \overline{1,n},\ n \geq 2$;

C_5. $\emptyset, \Omega \in \mathcal{C}$;

C_6. $A \setminus B \in \mathcal{C},\ \forall\ A, B \in \mathcal{C}$.

Demonstraţie. C_3 se verifică prin inducţie. C_4 rezultă din următoarea relaţie (De Morgan): $\cap_{i=1}^n A_i = (\cup_{i=1}^n A_i^c)^c$ şi proprietăţile C_1 şi C_3. Pentru C_5, considerăm $A \in \mathcal{C}$; avem $\Omega = A \cup A^c \in \mathcal{C}$ şi $\emptyset = A \cap A^c \in \mathcal{C}$. C_6 se obţine din relaţia: $A \setminus B = A \cap B^c$ şi C_4. $\square$

Vom considera în continuare intersecţii de corpuri ale unei mulţimi fixate.

Propoziţia 1.2.2. *O intersecţie de corpuri ale unei mulţimi date este un corp al acelei mulţimi.*

Demonstraţie. Fie $\{\mathcal{C}_i\ :\ i \in I\}$ o familie de corpuri ale mulţimii Ω. Notăm $\mathcal{C} = \cap_{i \in I} \mathcal{C}_i \subset \mathcal{P}(\Omega)$. Evident, $\Omega \in \mathcal{C}$ (cf. P_3), deci $\mathcal{C} \neq \emptyset$. Dacă $A \in \mathcal{C}$, atunci $A \in \mathcal{C}_i,\ \forall\ i \in I$, de unde $A^c \in \mathcal{C}_i,\ \forall\ i \in I$ (cf. C_1), deci $A^c \in \mathcal{C}$. Similar se verifică $A \cup B \in \mathcal{C},\ \forall\ A, B \in \mathcal{C}$. Rezultă că $\mathcal{C}$ este un corp a lui Ω. $\square$

Rezultatul de mai sus este util în studiul corpurilor generate de familii de părţi ale multimii totale considerate.

Definiţia 1.2.2. *Fie $\Omega \neq \emptyset$ şi $\mathcal{M} \subset \mathcal{P}(\Omega)$, $\mathcal{M} \neq \emptyset$. Fie $\mathcal{P} = \{\mathcal{C}\ -$ corp a lui $\Omega \,|\, \mathcal{M} \subset \mathcal{C}\}$. Atunci mulţimea*

$$\mathcal{C}(\mathcal{M}) = \bigcap_{\mathcal{C} \in \mathcal{P}} \mathcal{C}$$

se numeşte corpul generat de $\mathcal{M}$.

Vom observa la $\mathcal{C}(\mathcal{M})$ este cel mai mic corp a lui Ω care include familia de părţi $\mathcal{M}$ a lui Ω. Ne vom referi acum la cazul corpurilor finite.

Propoziţia 1.2.3. *Fie $\Omega \neq \emptyset$.*

1. *Dacă $\mathcal{M} \subset \mathcal{P}(\Omega)$, $\mathcal{M} \neq \emptyset$, este o mulţime finită, atunci corpul $\mathcal{C}(\mathcal{M})$ este finit.*

2. *Dacă $\mathcal{C}$ este un corp finit a lui Ω, atunci există o partiţie finită $\Delta = \{A_1, A_2, \cdots, A_n\}$ a lui Ω astfel ca*

$$\mathcal{C} = \{\cup_{i \in I} A_i\ :\ I \subset \{1, 2, \cdots, n\}\}.$$

Demonstraţie.
1. Fie mulţimile finite $\mathcal{M}^c = \{A^c\ :\ A \in \mathcal{M}\}$, $\mathcal{M}_1 = \mathcal{M} \cup \mathcal{M}^c$, $\mathcal{M}_2 = \{\cap_{A \in H} A\ :\ H \subset \mathcal{M}_1\}$ şi $\mathcal{M}_3 = \{\cup_{A \in G}\ :\ G \subset \mathcal{M}_2\}$. Atunci $\mathcal{C}(\mathcal{M}) = \mathcal{M}_3$, deci este finit.
2. Partiţia Δ este formată din mulţimile nevide $A \in \mathcal{C}$, având proprietatea următoare $B \subset A\ \Rightarrow\ B = A$ sau $B = \emptyset,\ \forall\ B \in \mathcal{C}$. $\square$

1.3 Spaţii măsurabile

Noţiunea de corp (de mulţimi), definită în paragraful enterior, va fi extinsă în cele ce urmează la cea de $\sigma - corp$. Aceasta va permite introducerea noţiunii de *spaţiu măsurabil* (câmp de evenimente).

Definiţia 1.3.1. *Fie Ω o mulţime nevidă. O mulţime nevidă $\mathcal{F} \subset \Omega$ se numeşte $\sigma - corp$ a lui Ω dacă satisface următoarele condiţii:*

σ_1. $A^c \in \mathcal{F},\ \forall\ A \in \mathcal{F}$;

σ_2. *dacă* $A_n \in \mathcal{F},\ \forall n \in \mathbb{N}^*$ *atunci* $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Perechea $(\Omega, \mathcal{F})$ se numeşte spaţiu măsurabil sau câmp de evenimente.

Vom observa că orice $\sigma - corp$ $\mathcal{F}$ a lui Ω este un corp a lui Ω. Astfel, pentru $A, B \in \mathcal{F}$ considerăm şirul $(A_n)_{n\geq 1}$ definit prin $A_1 = A$ şi $A_n = B,\ \forall\ n \geq 1$. Conform σ_1, avem $A \cup B = \cup_{n=1}^{\infty} A_n \in \mathcal{F}$. Deducem că σ-corpul $\mathcal{F}$ este corp. Ca urmare, proprietăţile $C_3 - C_6$ rămân valabile într-un σ-corp. Alte proprietăţi sunt descrise în propoziţia următoare.

Propoziţia 1.3.1. *Fie $(\Omega, \mathcal{F})$ un spaţiu măsurabil. Au loc următoarele proprietăţi:*

σ_3. $A_n \in \mathcal{F},\ \forall n \in \mathbb{N}^* \quad \Rightarrow \quad \cap_{n=1}^{\infty} A_n \in \mathcal{F}$;

σ_4. $A_n \in \mathcal{F},\ \forall n \in \mathbb{N}^* \quad \Rightarrow \quad \limsup_{n\to\infty} A_n \in \mathcal{F},\quad \liminf_{n\to\infty} A_n \in \mathcal{F}$,
unde $\limsup_{n\to\infty} A_n = \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k$ *şi* $\liminf_{n\to\infty} A_n = \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k$.

Demonstraţie. σ_3. Conform relaţiilor lui De Morgan şi Definiţiei 1.3.1, avem: $\cap_{n=1}^{\infty} A_n = (\cup_{n=1}^{\infty} A_n^c)^c \in \mathcal{F}$.
σ_4. Se aplică σ_2 şi P_5. $\square$

Remarcăm că familiile $\mathcal{F}_1 = \mathcal{P}(\Omega)$ şi $\mathcal{F}_2 = \{\emptyset, \Omega\}$ reprezintă σ-corpuri (deci şi corpuri) elementare ale lui Ω.

Intersecţia unor σ-corpuri ale unei mulţimi este de asemenea un σ-corp al mulţimii respective, demonstraţia fiind similară celei din Propoziţia 1.2.2.

Propoziţia 1.3.2. *Dacă $\mathcal{F}_i,\ i \in I$, reprezintă o familie oarecare de $\sigma-$corpuri ale mulţimii Ω, atunci $\cap_{i\in I}\mathcal{F}_i$ reprezintă de asemenea un corp al mulţimii Ω.*

Proprietatea de mai sus permite definirea noţiunii de σ-corp generat de o mulţime de părţi (submulţimi) ale unei mulţimi Ω.

Definiţia 1.3.2. *Fie Ω o mulţime nevidă şi $\mathcal{M} \subset \mathcal{P}(\Omega)$. Intersecţia tuturor $\sigma-$corpurilor lui Ω care includ mulţimea $\mathcal{M}$ se numeşte $\sigma-$corpul generat de $\mathcal{M}$ şi se notează $\mathcal{B}(\mathcal{M})$.*

Mulţimea $\mathcal{B}(\mathcal{M})$ este cel mai mic σ-corp a lui Ω care include pe $\mathcal{M}$. Un caz particular foarte important este cel în care mulţimea $\mathcal{M}$ este o topologie pe spaţiul Ω.

Definiţia 1.3.3. *Fie* $(\Omega, \mathcal{O})$ *un spaţiu topologic, unde* $\mathcal{O}$ *reprezintă topologia lui* Ω *(familia mulţimilor deschise). Atunci* $\sigma -$ *corpul generat de mulţimea* $\mathcal{O}$ *se numeşte borelianul spaţiului topologic* $(\Omega, \mathcal{O})$ *şi se notează (atunci când nu este pericol de confuzie) prin* $\mathcal{B}_\Omega$ *.*

Topologia $\mathcal{O}_\mathbb{R}$ a mulţimii numerelor reale $\mathbb{R}$ este familia reuniunilor cel mult numărabile de intervale deschise disjuncte. Se poate demonstra că borelianul $\mathcal{B}_\mathbb{R}$ al spaţiului topologic $(\mathbb{R}, \mathcal{O}_\mathbb{R})$ conţine toate tipurile de intervale şi reuniunile cel mult numărabile ale acestora. Reciproc, $\mathcal{B}_\mathbb{R}$ coincide cu σ - corpul generat de oricare tip de intervale reale.

1.4 Funcţii măsurabile

Descriem în continuare funcţiile ”specifice” spaţiilor măsurabile.

Definiţia 1.4.1. *Fie* $(\Omega, \mathcal{F})$ *şi* $(\Omega', \mathcal{F}')$ *o pereche de spaţii măsurabile (câmpuri de evenimente). O funcţie* $f : \Omega \to \Omega'$ *se numeşte funcţie măsurabilă dacă*

$$f^{-1}(A) \in \mathcal{F}, \quad \forall\, A \in \mathcal{F}',$$

unde $f^{-1}(A) = \{\omega \in \Omega \,|\, f(\omega \in A\}$.

În particular, dacă $\Omega' = \mathbb{R}$ şi $\mathcal{F}' = \mathcal{B}_\mathbb{R}$, funcţia f se va numi *funcţie reală măsurabilă.* De remarcat faptul că, pentru o funcţie reală măsurabilă $f : \Omega \to \mathbb{R}$, avem:

- $\{f \le a\} := f^{-1}((-\infty, a]) \in \mathcal{F}, \quad \forall\, a \in \mathbb{R}$
- $\{f > a\} := f^{-1}((a, \infty)) \in \mathcal{F}, \quad \forall\, a \in \mathbb{R}$
- $\{a \le f \le b\} := f^{-1}([a, b]) \in \mathcal{F}, \quad \forall\, a, b \in \mathbb{R},\ a < b.$

Conservarea măsurabilităţii prin operaţii algebrice este stabilită de propoziţia următoare.

Propoziţia 1.4.1. *Fie* $(\Omega, \mathcal{F})$ *un spaţiu măsurabil. Dacă* $f, g : \Omega \to \mathbb{R}$ *sunt funcţii reale măsurabile şi* $a \in \mathbb{R}$ *atunci funcţiile reale* $f + a,\ af,\ |f|,\ f + g,\ fg,\ \min\{f, g\}$ *şi* $\max\{f, g\}$ *sunt măsurabile.*

Funcţiie continue sunt măsurabile. Astfel, are loc:

Propoziţia 1.4.2. *Dacă* $(\Omega, \mathcal{B}_\Omega)$ *este un spaţiu măsurabil, cu* $\mathcal{B}_\Omega$ *reprezentând borelianul definit de o topologie pe* Ω*, iar* $f : \Omega \to \mathbb{R}$ *este o funcţie continuă, atunci* f *este o funcţie reală măsurabilă.*

Noţiunea de *funcţie reală măsurabilă* serveşte definirii noţiunii de *variabilă aleatoare.*

1.5 Câmpuri de probabilitate

Noţiunea de câmp de probabilitate este fundamentală în teoria probabilităţilor. Un câmp de probabilitate este un spaţiu măsurabil dotat cu o măsură finită numită "probabilitate".

Definiţia 1.5.1. *Fie $(\Omega, \mathcal{F})$ un spaţiu măsurabil, cu $\sigma-$ corpul $\mathcal{F}$ denumit "mulţimea evenimentelor". O funcţie $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ se numeşte probabilitate dacă satisface următoarele axiome:*

P1. $\mathbb{P}(A) \geq 0, \ \forall \ A \in \mathcal{F}$;

P2. $\mathbb{P}(\Omega) = 1$;

P3. Pentru oricare şir de evenimente $(A_n)_{n\in\mathbb{N}^*}$, cu $A_n \cap A_m = \emptyset, \ \forall \ n \neq m$, are loc relaţia

$$\mathbb{P}\left(\cup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

Tripletul $(\Omega, \mathcal{F}, \mathbb{P})$ se numeşte câmp de probabilitate.

Ultima proprietate a definiţiei de mai sus se referă la şiruri de evenimente. În continuare, ne vom referi la şirurile monotone de evenimente şi limitele de şiruri de evenimente.

Definiţia 1.5.2. *Fie $(A_n)_{n\in\mathbb{N}^*}$ un şir de evenimente din câmpul de probabilitate $(\Omega, \mathcal{F}, \mathbb{P})$.*

1. *Şirul $(A_n)_{n\in\mathbb{N}^*}$ se numeşte monoton crescător dacă $A_n \subset A_{n+1}, \ \forall \ n \geq 1$. Fie $A = \cup_{n=1}^{\infty} A_n \in \mathcal{F}$. Notăm $\lim_{n\to\infty} A_n = A$ şi $A \uparrow A$.*

2. *Şirul $(A_n)_{n\in\mathbb{N}^*}$ se numeşte monoton descrescător dacă $A_{n+1} \subset A_n, \ \forall \ n \geq 1$. Fie $A = \cap_{n=1}^{\infty} A_n \in \mathcal{F}$. Notăm $\lim_{n\to\infty} A_n = A$ şi $A \downarrow A$.*

3. *Limita superioară a şirului $(A_n)_{n\in\mathbb{N}^*}$ se defineşte prin*

$$\limsup_{n\to\infty} A_n = \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k \in \mathcal{F}.$$

4. *Limita inferioară a şirului $(A_n)_{n\in\mathbb{N}^*}$ se defineşte prin*

$$\liminf_{n\to\infty} A_n = \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k \in \mathcal{F}.$$

Probabilitatea are o serie de proprietăţi fundamentale care decurg din Definiţia 1.5.1. Aceste proprietăţi caracterizează în fapt o *măsură finită*.

Propoziţia 1.5.1. *Fie $(\Omega, \mathcal{F}, \mathbb{P})$ un câmp de probabilitate. Au loc următoarele proprietăţi*

P4. $\mathbb{P}(\emptyset) = 0$;

P5. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B), \ \forall \ A, B \in \mathcal{F}, \ A \cap B = \emptyset$;

P6. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A), \ \forall \ A \in \mathcal{F}$;

P7. $\mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(B),\ \forall\ A, B \in \mathcal{F},\ B \subset A$*;*

P8. $\mathbb{P}(A) \leq \mathbb{P}(B),\ \forall\ A, B \in \mathcal{F},\ A \subset B$*;*

P9. $\mathbb{P}(A) \in [0, 1],\ \forall\ A \in \mathcal{F}$*;*

P10. (Formula lui Poincaré)

$$\mathbb{P}(\cup_{i=1}^n A_i) = \sum_{k=1}^n (-1)^{k+1} \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \mathbb{P}(\cap_{j=1}^k A_{i_j});$$

P11. Dacă $(A_n)_{n \geq 1}$ *este un şir monoton (crescător sau descrescător) de evenimente, cu* $\lim_{n\to\infty} A_n = A$, *atunci* $\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A)$.

P12. Pentru oricare şir de evenimente $(A_n)_{n \in \mathbb{N}^*}$ *avem*

$$\mathbb{P}\left(\cup_{n=1}^\infty A_n\right) \leq \sum_{n=1}^\infty \mathbb{P}(A_n).$$

Demonstraţie.
P4. Presupunem, prin absurd, că $\mathbb{P}(\emptyset) = p > 0$. Considerăm şirul $A_n = \emptyset,\ n \in \mathbb{N}^*$. Conform $[P3]$, $\mathbb{P}(\emptyset) = \mathbb{P}(\cup_{n=1}^\infty (A_n) = \sum_{n=1}^\infty \mathbb{P}(A_n) = \infty$. Contradicţie.
P5. Alegem şirul $A_n = \emptyset,\ n \geq 3,\ A_1 = A$ şi $A_2 = B$ şi aplicăm $[P3],\ [P4]$.
P6. Relaţia rezultă din $A^c \cup A = \Omega$ şi proprietăţile $[P2],\ [P5]$.
P7. Se obţine din $A = B \cup (A \setminus B)$ şi $[P5]$.
P8. Rezultă din proprietatea anterioară, $[P7]$.
P9. Consecinţă a proprietăţilor $[P1], [P2]$ şi $[P8]$.
P10. Formula se demonstrează prin inducţie matematică.
P11. Presupunem $A_n \downarrow \emptyset$. Avem $A_1 = \cup_{n=1}^\infty (A_n \setminus A_{n+1})$, cu $(A_i \setminus A_{i+1}) \cap (A_j \setminus A_{j+1}) = \emptyset,\ \forall\ i \neq j$. Atunci $\mathbb{P}(A_1) = \sum_{n=1}^\infty \mathbb{P}(A_n \setminus A_{n+1})$. Rezultă $\lim_{n\to\infty} \sum_{k=n}^\infty \mathbb{P}(A_k \setminus A_{k+1}) = 0$. Dar $\sum_{k=n}^\infty \mathbb{P}(A_k \setminus A_{k+1}) = \mathbb{P}\left(\cup_{n=1}^\infty (A_n \setminus A_{n+1})\right) = \mathbb{P}(A_n)$. Deci $\lim_{n\to\infty} \mathbb{P}(A_n) = 0 = \mathbb{P}(\emptyset)$.
Presupunem $A_n \downarrow A$. Atunci $(A_n \setminus A) \downarrow \emptyset$, de unde $\lim_{n\to\infty} \mathbb{P}(A_n \setminus A) = 0$ şi obţinem $\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A)$.
Presupunem $A_n \uparrow A$. Atunci $A_n^c \downarrow A^c$. Rezultă $\lim_{n\to\infty} \mathbb{P}(A_n^c) = \mathbb{P}(A^c)$, de unde obţinem $\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A)$.
P12. Definim şirul de evenimente $(B_n)_{n \geq 1},\ B_1 = A_1,\ B_n = A_n \setminus \left(\cup_{k=1}^{n-1} A_k\right) \subset A_n,\ \forall\ n > 1$. Avem $\cup_{k=1}^n B_k = \cup_{k=1}^n A_k,\ \forall\ n \geq 1$ şi $B_i \cap B_j = \emptyset,\ \forall\ i \neq j$. Ca urmare, conform proprietăţilor $[P3]$ şi $[P8]$, avem

$$\mathbb{P}\left(\cup_{n=1}^\infty A_n\right) = \mathbb{P}\left(\cup_{n=1}^\infty B_n\right) = \sum_{n=1}^\infty \mathbb{P}(B_n) \leq \sum_{n=1}^\infty \mathbb{P}(A_n),$$

ceea ce trebuia demonstrat. $\square$

1.6 Evenimente independente, evenimente condiţionate

Un fenomen important studiat de teoria probabilităţilor este cel al dependenţei (respectiv independenţei) evenimentelor.

Definiţia 1.6.1. *Fie* $(\Omega, \mathcal{F}, \mathbb{P})$ *un câmp de probabilitate. Considerăm un eveniment* $A \in \mathcal{F}$, *cu* $\mathbb{P}(A) > 0$. *Pentru oricare eveniment* $B \in \mathcal{F}$ *numim probabilitatea lui* B *condiţionată de (realizarea evenimentului)* A *mărimea:*

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

Evenimentele A *şi* B *se numesc evenimente independente dacă* $\mathbb{P}(B|A) = \mathbb{P}(B)$, *adică* $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$. *Mai general,* $\mathcal{G} \subset \mathcal{F}$ *se numeşte familie de evenimente independente dacă pentru oricare evenimente distincte* $A_1, A_2, \cdots, A_n \in \mathcal{G}$ *avem*

$$\mathbb{P}\left(\cap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} \mathbb{P}(A_i).$$

Se constată astfel că două evenimente sunt independente dacă probabilitatea realizării unuia dintre acestea nu este influienţată de realizarea celuilalt. Se arată cu uşurinţă că evenimentele contrare ale unor evenimente independente sunt de asemenea independente. Pe de altă parte, conceptul de *evenimente condiţionate* permite considerarea *sistemelor complete de evenimente.*

Definiţia 1.6.2. *Evenimentele* A_i, $i \in I$, *unde* I *este o mulţime cel mult numărabilă de indici, formează un sistem complet de evenimente în câmpul de probabilitate* $(\Omega, \mathcal{F}, \mathbb{P})$ *dacă:*

1. $A_i \cap A_j = \emptyset, \ \forall\, i \neq j$
2. $\cup_{i \in I} A_i = \Omega$
3. $\mathbb{P}(A_i) > 0, \ \forall\, i \in I$

Propoziţia 1.6.1. *Fie* $(\Omega, \mathcal{F}, \mathbb{P})$ *un câmp de probabilitate, iar* A_i, $i \in I$, *un sistem complet de evenimente. Dacă* A *este un eveniment cu probabilitate nenulă, atunci loc:*

1. *Formula probabilităţii totale*

$$\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A|A_i)\mathbb{P}(A_i)$$

2. *Formula lui Bayes*

$$\mathbb{P}(A_k|A) = \frac{\mathbb{P}(A|A_k)\mathbb{P}(A_k)}{\sum_{i \in I} \mathbb{P}(A|A_i)\mathbb{P}(A_i)}, \quad \forall\, k \in I.$$

Demonstraţie.
1. Avem

$$A = A \cap \Omega = A \cap (\cup_{i\in I} A_i) = \cup_{i\in I} (A \cap A_i) .$$

Deoarece evenimentele $A \cap A_i,\ i \in I$ sunt incompatibile două câte două, obţinem

$$\mathbb{P}(A) = \sum_{i\in I} \mathbb{P}(A \cap A_i) = \sum_{i\in I} \mathbb{P}(A|A_i)\,\mathbb{P}(A_i) ,$$

deci formula probabilităţii totale este demonstrată.
2. Fie $k \in I$. Conform Definiţiei 1.6.1 şi formulei probabilităţii totale, obţinem

$$\mathbb{P}(A_k|A) = \frac{\mathbb{P}(A \cap A_k)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|A_k)\mathbb{P}(A_k)}{\sum_{i\in I} \mathbb{P}(A|A_i)\mathbb{P}(A_i)},$$

deci formula lui Bayes este demonstrată. $\square$

1.7 Scheme clasice de probabilitate

Prezentăm în continuare modelele de bază de cmpuri de probabilitate (discrete).

1. *Câmpul lui Laplace.*
   Spaţiul Ω este considerat o mulţime finită, iar $\sigma-$*corpul* $\mathcal{F}$ este considerat mulţimea $\mathcal{P}(\Omega)$ (mulţimea submulţimilor lui Ω). Probabilitatea unui eveniment $A \subset \Omega$ se defineşte prin:
   $$\mathbb{P}(A) = \frac{|A|}{|\Omega|},$$
   unde $|M|$ reprezintă numărul elementelor unei mulţimi finite M. Elementele lui Ω determină evenimentele *elementare*, *echiprobabile*. Astfel, pentru $\omega \in \Omega$, avem
   $$\mathbb{P}(\{\omega\}) = \frac{1}{n},$$
   unde $n = |\Omega|$.

2. *Câmp discret de evenimente.*
   Considerăm $\Omega = \{\omega_i : i \in I\}$, unde mulţimea de indici I este cel mult numărabilă. Definim $\mathcal{F} = \mathcal{P}(\Omega)$. Considerăm o măsură de probabilitate $(p_i)_{i\in I}$ pe mulţimea Ω, unde $p_i \geq 0,\ \forall\, i \in I$ şi $\sum_{i\in I} p_i = 1$. Probabilitatea unui eveniment $A = \{\omega_j : j \in J\}$, cu $J \subset I$, se defineşte prin
   $$\mathbb{P}(A) = \sum_{j\in J} p_j .$$

Modelele de mai sus se materializează concret în *scheme de probabilitate.* Este vorba exemple cu un grad mare de generalitate şi utilizare frecventă în *calculul probabilităţilor.* Prezentăm în continuare cele mai cunoscute scheme de probabilitate.

1. *Schema lui Bernoulli.*
Este un exemplu de câmp Laplace. Schema modelează următoarea experienţă:
O urnă conţine c bile, dintre care a sunt albe şi b sunt negre (a+b=c). Se extrage o bilă, i se înregistrează culoarea, iar apoi se reintroduce în urnă. Se repetă această experienţă de n ori. Se studiază probabilitatea evenimentului $A_{k|n}$ al apariţiei bilei albe de k ori în cele n extrageri (cu repunerea bilei în urnă).
Notăm $\mathbb{P}_{k|n} = \mathbb{P}\left(A_{k|n}\right)$ probabilitatea evenimentului menţionat. Un raţionament elementar ne conduce la formula de calcul
$$\mathbb{P}_{k|n} = C_n^k p^k q^{n-k},\ k = 0, 1, \cdots, n,$$
unde $p = a/c$ este probabilitatea apariţiei unei bile la o extragere, iar $q = b/c = 1-p$ este probabilitatea evenimentului contrar (apariţia unei bile negre).
Modelul este echivalent cu repetarea unei experienţe (în condiţii identice) de n ori şi studiul probabilităţii realizării unui anumit eveniment de k ori (din totalul de n experienţe).
Remarcăm
$$\sum_{k=0}^{n} \mathbb{P}\left(A_{k|n}\right) = \sum_{k=0}^{n} \mathbb{P}_{k|n} = \sum_{k=0}^{n} C_n^k p^k q^{n-k} = (q+p)^n = 1,$$
deci evenimentele $A_{k|n},\ k = 0, 1, \cdots, n$, reprezintă un sistem complet.

2. *Schema multinomială.*
Este o extindere a schemei lui Bernoulli.
Se consideră o urnă conţinând bile de s culori, câte a_i bile din culoarea $i \in \{1, 2, \cdots, s\}$.
Notăm
$$p_i = \frac{a_i}{\sum_{j=1}^{s} a_j}$$
probabilitatea extragerii unei bile de culoarea $i \in \{1, 2, \cdots, s\}$. Se fac n extrageri, cu repunerea bilei în urnă după fiecare extragere. Fie $(k_i)_{i=\overline{1,s}}$ un vector de dimensiune s, cu componente naturale, având proprietatea $\sum_{i=1}^{s} k_i = n$. Notăm
$$\mathbb{P}_{(k_i)_{i=\overline{1,s}}|n}$$
probabilitatea evenimentului $A_{(k_i)_{i=\overline{1,s}}|n}$ de extragere a exact k_i bile de culoarea i, unde $i = 1, 2, \cdots, s$, din totalul de n bile extrase. Se constată că
$$\mathbb{P}_{(k_i)_{i=\overline{1,s}}|n} = \left(\begin{array}{c} n \\ k_1,\ k_2, \cdots,\ k_s \end{array}\right) p_1^{k_1} p_2^{k_2} \cdots p_s^{k_s},$$
unde
$$\left(\begin{array}{c} n \\ k_1,\ k_2, \cdots,\ k_s \end{array}\right) = \frac{n!}{k_1! k_2! \cdots k_s!}$$
reprezintă combinările multinomiale de n luate câte $k_1,\ k_2, \cdots,\ k_s$.
Avem
$$\sum_{(k_i)\in\mathbb{N}^s;\ \sum_{i=1}^{s} k_i = n} \mathbb{P}_{(k_i)_{i=\overline{1,s}}|n} = \sum_{(k_i)\in\mathbb{N}^s;\ \sum_{i=1}^{s} k_i = n} \left(\begin{array}{c} n \\ k_1,\ k_2, \cdots,\ k_s \end{array}\right) p_1^{k_1} p_2^{k_2} \cdots p_s^{k_s} =$$

$$= (p_1 + p_2 + \cdots + p_s)^n = 1,$$

deci $A_{(k_i)_{i=\overline{1,s}}|n}$, cu $(k_i) \in \mathbb{N}^s$, astfel ca $\sum_{i=1}^{s} k_i = n$, formează un sistem complet de evenimente.

3. *Schema lui Poisson.*
   Reprezintă o generalizare a schemei lui Bernoulli. Se consideră n urne cu bile albe şi negre, dar în proporţii diferite. Fie p_i probabilitatea extragerii unei bile albe din urna i. Probabilitatea extragerii unei bile negre din urna i va fi $q_i = 1 - p_i,\ i = 1, 2, \cdots, n$. Se extrage câte o bilă din fiecare urnă. Se analizează probabilitatea evenimentului $A_{k|n}$ de a obţine exact k bile albe din cele n extrase. Notăm $\mathbb{N}_n = \{1, 2, \cdots, n\}$. Avem
   $$\mathbb{P}_{k|n} = \mathbb{P}\left(A_{k|n}\right) = \sum_{I \subset \mathbb{N}_n;\ |I| = k} \prod_{i \in I} p_i \prod_{j \in \mathbb{N}_n \setminus I} q_j.$$
   Observăm că $\mathbb{P}_{k|n}$ reprezintă coeficientul lui x^k al polinomului
   $$f(x) = (p_1 x + q_1)(p_2 x + q_2) \cdots (p_n x + q_n).$$
   Evident,
   $$\sum_{k=0}^{n} \mathbb{P}_{k|n} = f(1) = 1,$$
   deci evenimentele $A_{k|n},\ k = 0, 1, \cdots, n$, reprezintă un sistem complet.

4. *Schema geometrică.*
   Schema este un exemplu de câmp discret de evenimente. Astfel, se repetă continuu o experienţă, în condiţii identice, urmărindu-se de fiecare dată realizarea unui anumit eveniment A. Fie $\mathbb{P}(A) = p$. Considerăm momentul în care evenimentul A se realizează pentru prima dată. Fie $n \in \mathbb{N}^*$. Probabilitatea ca evenimentul A să se realizeze pentru prima dată în cea de a n-a experienţă este
   $$p_n = pq^{n-1},$$
   unde $q = 1 - p$. Avem
   $$\sum_{n=1}^{\infty} p_n = \sum_{n=1}^{\infty} pq^{n-1} = p \sum_{k=0}^{\infty} q^k = \frac{p}{1-q} = 1.$$

5. *Schema hipergeometrică.*
   Se consideră o urnă cu N bile, dintre care a bile sunt albe iar restul de b bile sunt negre ($a + b = N$). Se extrag n bile (fără a le repune în urmă după fiecare extragere). Notăm $p_{k|n}$ probabilitatea evenimentului $A_{k|n}$ de a avea exact k bile albe între cele n bile extrase. Presupunem $n \leq N,\ k \leq a$ şi $n - k \leq b$. Atunci
   $$p_{k|n} = \frac{C_a^k C_b^{n-k}}{C_N^n}.$$
   Conform unei formule combinatoriale binecunoscute, obţinem
   $$\sum_{k=\max\{0,n-b\}}^{\min\{a,n\}} p_{k|n} = \sum_{k=\max\{0,n-b\}}^{\min\{a,n\}} \frac{C_a^k C_b^{n-k}}{C_N^n} = 1,$$

deci $A_{k|n}$, $\max\{0, n-b\} \leq k \leq \min\{a, n\}$, formează un sistem complet de evenimente.

6. *Schema hipergeometrică generalizată.*
   Se consideră o urnă cu N bile de s culori. Fie a_i numărul de bile de culoarea i, $i = 1, 2, \cdots, s$. Avem deci $\sum_{i=1}^{s} a_i = N$. Se extrag, fără repunere, n bile. Considerăm vectorul $(k_i) \in \mathbb{N}^s$, cu $\sum_{i=1}^{s} k_i = n$, astfel încât $k_i \leq a_i$, $i = 1, 2, \cdots, s$. Atunci probabilitatea extragerii a k_i bile din culoarea i, pentru fiecare $i \in \{1, 2, \cdots, s\}$, este

$$p_{(k_i)} = \frac{C_{a_1}^{k_1} C_{a_2}^{k_2} \cdots C_{a_s}^{k_s}}{C_N^n}.$$

1.8 Rezumat

În teoria probabilităţilor, evenimentele rezultate în urma unei experienţe sunt interpretate ca o familie de submulţimi ale evenimentului sigur, mulţimea Ω. Mulţimea evenimentelor formează un $\sigma - corp$ (borelian).

O mulţime nevidă $\mathcal{F} \subset \Omega$ se numeşte $\sigma - corp$ a lui Ω daca satisface următoarele condiţii:

σ_1. $A^c \in \mathcal{F},\ \forall\ A \in \mathcal{F}$;

σ_2. dacă $A_n \in \mathcal{F},\ \forall n \in \mathbb{N}^*$ atunci $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Perechea $(\Omega, \mathcal{F})$ se numeşte spaţiu măsurabil sau câmp de evenimente.

Borelianul lui $\Omega = \mathbb{R}$, notat $\mathcal{B}_{\mathbb{R}}$ este $\sigma - corpul$ generat de mulţimile deschise (topologia lui $\mathbb{R}$), adică cel mai mic $\sigma - corp$ a lui $\mathbb{R}$ care include mulţimile deschise. $\mathcal{B}_{\mathbb{R}}$ conţine toate tipurile de intervale şi este generat de oricare tip de intervale reale.

Fie $(\Omega, \mathcal{F})$ şi $(\Omega', \mathcal{F}')$ o pereche de spaţii măsurabile (câmpuri de evenimente). O funcţie $f : \Omega \to \Omega'$ se numeşte *funcţie măsurabilă* dacă

$$f^{-1}(A) \in \mathcal{F}, \quad \forall\ A \in \mathcal{F}',$$

unde $f^{-1}(A) = \{\omega \in \Omega \,|\, f(\omega \in A\}$.

În particular, dacă $\Omega' = \mathbb{R}$ şi $\mathcal{F}' = \mathcal{B}_{\mathbb{R}}$, funcţia f se va numi *funcţie reală măsurabilă*. Pentru o funcţie reală măsurabilă $f : \Omega \to \mathbb{R}$, avem:

- $\{f \leq a\} := f^{-1}((-\infty, a]) \in \mathcal{F},\ \ \forall\ a \in \mathbb{R}$
- $\{f > a\} := f^{-1}((a, \infty)) \in \mathcal{F},\ \ \forall\ a \in \mathbb{R}$
- $\{a \leq f \leq b\} := f^{-1}([a, b]) \in \mathcal{F},\ \ \forall\ a, b \in \mathbb{R},\ a < b.$

Dacă $f, g : \Omega \to \mathbb{R}$ sunt funcţii reale măsurabile şi $a \in \mathbb{R}$ atunci funcţiile reale $f + a,\ af,\ |f|,\ f + g,\ fg,\ \min\{f, g\}$ şi $\max\{f, g\}$ sunt măsurabile.

Fie $(\Omega, \mathcal{F})$ un spaţiu măsurabil, cu $\sigma - corpul$ $\mathcal{F}$ denumit "mulţimea evenimentelor". O funcţie $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ se numeşte probabilitate dacă satisface următoarele axiome:

P1. $\mathbb{P}(A) \geq 0,\ \forall\ A \in \mathcal{F}$;

P2. $\mathbb{P}(\Omega) = 1$;

P3. Pentru oricare şir de evenimente $(A_n)_{n \in \mathbb{N}^*}$, cu $A_n \cap A_m = \emptyset,\ \forall\ n \neq m$, are loc relaţia

$$\mathbb{P}\left(\cup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

Tripletul $(\Omega, \mathcal{F}, \mathbb{P})$ se numeşte câmp de probabilitate.

Probabilitatea are proprietăţile:

P4. $\mathbb{P}(\emptyset) = 0$;

P5. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B),\ \forall\ A, B \in \mathcal{F},\ A \cap B = \emptyset$;

P6. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A),\ \forall\ A \in \mathcal{F}$;

P7. $\mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(B),\ \forall\ A, B \in \mathcal{F},\ B \subset A$;

P8. $\mathbb{P}(A) \leq \mathbb{P}(B),\ \forall\ A, B \in \mathcal{F},\ A \subset B$;

P9. $\mathbb{P}(A) \in [0, 1],\ \forall\ A \in \mathcal{F}$;

P10. (Formula lui Poincaré)

$$\mathbb{P}(\cup_{i=1}^n A_i) = \sum_{k=1}^n (-1)^{k+1} \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \mathbb{P}(\cap_{j=1}^k A_{i_j});$$

P11. Dacă $(A_n)_{n\geq 1}$ este un şir monoton (crescător sau descrescător) de evenimente, cu $\lim_{n\to\infty} A_n = A$, atunci $\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A)$.

P12. Pentru oricare şir de evenimente $(A_n)_{n\in\mathbb{N}^*}$ avem

$$\mathbb{P}\left(\cup_{n=1}^\infty A_n\right) \leq \sum_{n=1}^\infty \mathbb{P}(A_n).$$

Considerăm un eveniment $A \in \mathcal{F}$, cu $\mathbb{P}(A) > 0$. Pentru oricare eveniment $B \in \mathcal{F}$, numim probabilitatea lui B condiţionată de (realizarea evenimentului) A mărimea:

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

Evenimentele A şi B se numesc independente dacă $\mathbb{P}(B|A) = \mathbb{P}(B)$, adică $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.
Mai general, $\mathcal{G} \subset \mathcal{F}$ se numeşte familie de evenimente independente dacă pentru oricare evenimente distincte $A_1, A_2, \cdots, A_n \in \mathcal{G}$ avem

$$\mathbb{P}\left(\cap_{i=1}^n A_i\right) = \prod_{i=1}^n \mathbb{P}(A_i).$$

Astfel, două evenimente sunt independente dacă probabilitatea realizării unuia dintre acestea nu este influienţată de realizarea celuilalt.

Evenimentele $A_i,\ i \in I$, unde I este o mulţime cel mult numřabilă de indici, formează un sistem complet de evenimente în câmpul de probabilitate $(\Omega, \mathcal{F}, \mathbb{P})$ dacă:

1. $A_i \cap A_j = \emptyset,\ \forall\ i \neq j$
2. $\cup_{i\in I} A_i = \Omega$
3. $\mathbb{P}(A_i) > 0,\ \forall\ i \in I$

Dacă A este un eveniment cu probabilitate nenulă (neneglijabil), atunci loc:

1. Formula probabilităţii totale

$$\mathbb{P}(A) = \sum_{i\in I} \mathbb{P}(A|A_i)\mathbb{P}(A_i)$$

2. Formula lui Bayes

$$\mathbb{P}(A_k|A) = \frac{\mathbb{P}(A|A_k)\mathbb{P}(A_k)}{\sum_{i\in I}\mathbb{P}(A|A_i)\mathbb{P}(A_i)}, \quad \forall\, k \in I.$$

Modelele de bază de cm̂puri de probabilitate discrete:

1. *Câmpul lui Laplace.*
   Spaţiul Ω este considerat o mulţime finită, iar $\sigma - corpul$ $\mathcal{F}$ este considerat mulţimea $\mathcal{P}(\Omega)$ (mulţimea submulţimilor lui Ω). Probabilitatea unui eveniment $A \subset \Omega$ se defineşte prin:
   $$\mathbb{P}(A) = \frac{|A|}{|\Omega|},$$
   unde $|M|$ reprezintănumărul elementelor unei mulţimi finite M. Elementele lui Ω determină evenimentele *elementare*, *echiprobabile*. Astfel, pentru $\omega \in \Omega$, avem
   $$\mathbb{P}(\{\omega\}) = \frac{1}{n},$$
   unde $n = |\Omega|$.

2. *Câmp discret de evenimente.*
   Considerăm $\Omega = \{\omega_i : i \in I\}$, unde mulţimea de indici I este cel mult numărabilă. Definim $\mathcal{F} = \mathcal{P}(\Omega)$. Considerăm o măsură de probabilitate $(p_i)_{i\in I}$ pe mulţimea Ω, unde $p_i \geq 0$, $\forall\, i \in I$ şi $\sum_{i\in I} p_i = 1$. Probabilitatea unui eveniment $A = \{\omega_j : j \in J\}$, cu $J \subset I$, se defineşte prin
   $$\mathbb{P}(A) = \sum_{j\in J} p_j.$$

Schemele de probabilitate reprezintă modele/exemple cu un grad mare de generalitate în *calculul probabilităţilor.*

1. *Schema lui Bernoulli.*
   Este un exemplu de câmp Laplace. Schema modelează următoarea experienţă:
   O urnă conţine c bile, dintre care a sunt albe şi b sunt negre (a+b=c). Se extrage o bilă, i se înregistrează culoarea, iar apoi se reintroduce în urnă. Se repetă această experienţă de n ori. Se studiază probabilitatea evenimentului $A_{k|n}$ al apariţiei bilei albe de k ori din cele n extrageri (cu repunerea bilei în urnă).
   Notăm $\mathbb{P}_{k|n} = \mathbb{P}\left(A_{k|n}\right)$ probabilitatea evenimentului menţionat. Avem
   $$\mathbb{P}_{k|n} = C_n^k p^k q^{n-k},\ k = 0, 1, \cdots, n,$$
   unde $p = a/c$ este probabilitatea apariţiei unei bile la o extragere, iar $q = b/c = 1-p$ este probabilitatea evenimentului contrar (apariţia unei bile negre).

2. *Schema multinomială.*
   Este o extindere a schemei lui Bernoulli.
   Se consideră o urnă conţinând bile de s culori, câte a_i bile din culoarea $i \in \{1, 2, \cdots, s\}$.
   Notăm
   $$p_i = \frac{a_i}{\sum_{j=1}^{s} a_j}$$

probabilitatea extragerii unei bile de culoarea $i \in \{1, 2, \cdots, s\}$. Se fac n extrageri, cu repunerea bilei în urnă după fiecare extragere. Fie $(k_i)_{i=\overline{1,s}}$ un vector de dimensiune s, cu componente naturale, având proprietatea $\sum_{i=1}^{s} k_i = n$. Notăm

$$\mathbb{P}_{(k_i)_{i=\overline{1,s}}|n}$$

probabilitatea evenimentului $A_{(k_i)_{i=\overline{1,s}}|n}$ de extragere a exact k_i bile de culoarea i, unde $i = 1, 2, \cdots, s$, din totalul de n bile extrase. Avem

$$\mathbb{P}_{(k_i)_{i=\overline{1,s}}|n} = \binom{n}{k_1,\ k_2, \cdots,\ k_s} p_1^{k_1} p_2^{k_2} \cdots p_s^{k_s},$$

unde

$$\binom{n}{k_1,\ k_2, \cdots,\ k_s} = \frac{n!}{k_1! k_2! \cdots k_s!}$$

reprezintă combinările multinomiale de n luate câte $k_1,\ k_2, \cdots,\ k_s$.

3. *Schema lui Poisson.*
Reprezintă o generalizare a schemei lui Bernoulli. Se consideră n urne cu bile albe şi negre, dar în proporţii diferite. Fie p_i probabilitatea extragerii unei bile albe din urna i. Probabilitatea extragerii unei bile negre din urna i va fi $q_i = 1 - p_i,\ i = 1, 2, \cdots, n$. Se extrage câte o bilă din fiecare urnă. Se analizează probabilitatea evenimentului $A_{k|n}$ de a obţine exact k bile albe din cele n extrase. Notăm $\mathbb{N}_n = \{1, 2, \cdots, n\}$. Avem
$$\mathbb{P}_{k|n} = \mathbb{P}\left(A_{k|n}\right) = \sum_{I \subset \mathbb{N}_n;\ |I|=k} \prod_{i \in I} p_i \prod_{j \in \mathbb{N}_n \setminus I} q_j.$$
Observăm că $\mathbb{P}_{k|n}$ reprezintă coeficientul lui x^k al polinomului
$$f(x) = (p_1 x + q_1)(p_2 x + q_2) \cdots (p_n x + q_n).$$

4. *Schema geometrică.*
Schema este un exemplu de câmp discret de evenimente. Astfel, se repetă continuu o experienţă, în condiţii identice, urmărindu-se de fiecare dată realizarea unui anumit eveniment A. Fie $\mathbb{P}(A) = p$. Considerăm momentul în care evenimentul A se realizează pentru prima dată. Fie $n \in \mathbb{N}^*$. Probabilitatea ca evenimentul A să se realizeze pentru prima dată în cea de a n-a experienţă este
$$p_n = pq^{n-1},$$
unde $q = 1 - p$.

5. *Schema hipergeometrică.*
Se consideră o urnă cu N bile, dintre care a bile sunt albe iar restul de b bile sunt negre $(a+b = N)$. Se extrag n bile (fără a le repune în urmă după fiecare extragere). Notăm $p_{k|n}$ probabilitatea evenimentului $A_{k|n}$ de a avea exact k bile albe între cele n bile extrase. Presupunem $n \leq N,\ k \leq a$ şi $n - k \leq b$. Atunci
$$p_{k|n} = \frac{C_a^k C_b^{n-k}}{C_N^n}.$$

1.9 Test

1. Definiţi noţiunea de evenimente condiţionate. Enunţati şi demonstraţi formula probabilităţii totale.

2. Aplicaţi formula lui Poincaré pentru n evenimente independente, echiprobabile, de probabilitate $p \in (0, 1)$.

3. Fie $(\Omega, \mathcal{F}, \mathbb{P})$ un câmp de probabilitate şi evenimentele $A, B \in \mathcal{F}$. Ştiind că $\mathbb{P}(A) = 1/5$, $\mathbb{P}(B) = 3/5$ şi $\mathbb{P}(A \cap B) = 1/10$, să se calculeze probabilităţile:

   (a) $\mathbb{P}(A \cup B)$;

   (b) $\mathbb{P}(A \cap \overline{B})$;

   (c) $\mathbb{P}(A|B) = \mathbb{P}_B(A)$.

4. Un sportiv trage la o ţintă cu 5 puşti diferite, cu probabilităţile respective de ochire: $p_1 = 0,8$, $p_2 = 0,7$, $p_3 = 0,4$, $p_4 = 0,9$ şi $p_5 = 0,5$. Să se calculeze:

   (a) probabilitatea ca ţinta să fie nimerită de exact 3 ori;

   (b) probabilitatea ca ţinta să nimerită cel puţin o data.

5. Un portar apără o lovitură de la 11 metri cu probabilitatea $p = 0,3$. La finalul unui meci, se execută 5 lovituri de la 11 metri.

   (a) Care este probabilitatea ca portarul respectiv să apere exact 2 lovituri?

   (b) Care este cel mai probabil număr de goluri primite?

6. Care este probabilitatea de a nimeri exact doua numere la jocul *Loto 6 din 49*, dacă se joacă o singură variantă?

7. Un baschetbalist execută aruncări de la distanţă la un panou de baschet, până la prima reuşită. Ştiind că şansa de reuşită a fiecărei aruncări este de 20%, să se determine probabilitatea ca sportivul să înscrie primul coş din a patra aruncare. Care este cel mai probabil număr de aruncări până la prima reuşită?

Capitolul 2

Variabile aleatoare. Caracteristici numerice

2.1 Introducere

2.2 Variabile aleatoare

Variabilele aleatoare se interpretează ca o "enumerare" a rezultatelor (măsurabile) potenţiale ale unui anumit experiment. Fiecare din aceste rezultate posibile are o probabilitate de realizare. Definiţia de mai jos formalizează matematic un experiment cu rezultate aleatoare, predictibile.

Definiţia 2.2.1. *Fie* $(\Omega, \mathcal{F}, \mathbb{P})$ *un câmp de probabilitate. O funcţie măsurabilă reală* $X : \Omega \to \mathbb{R}$ *se numeşte variabilă aleatoare (reală).*

Conform definiţiei, pentru o variabilă aleatoare X definită pe câmpul de probabilitate $(\Omega, \mathcal{F}, \mathbb{P})$, avem

$$\{X \in I\} \in \mathcal{F},$$

pentru orice interval real I. In particular,

$$\{X \in (-\infty, a]\} = \{X \leq a\} \in \mathcal{F}, \ \forall \ a \in \mathbb{R}.$$

Probabilităţile evenimentelor descrise mai sus oferă o informaţie fundamentală privind variabila aleatoare. În fapt, cunoaşterea probabilităţile evenimentelor $\{X \leq a\}, \ a \in \mathbb{R}$, oferă o informaţie precisă privind distribuţia valorilor variabilei aleatore X. În secţiunea următoare, va fi introdusă şi caracterizată *funcţia de repartiţie* a a unei variabile aleatoare, construită pe baza probabilităţilor menţionate.

O mulţime finită de variabile aleatoare definite pe acelaşi câmp de probabilitate va determina un vector aleator. Mai precis, avem următoarea definiţie.

Definiţia 2.2.2. *Fie* $d > 1$ *un număr întreg şi* $(\Omega, \mathcal{F}, \mathbb{P})$ *un câmp de probabilitate. Considerăm borelianul* $\mathcal{B}_{\mathbb{R}^d}$ *spaţiului* $\mathbb{R}^d$ *(*$\sigma-$ *corpul generat de topologia lui* $\mathbb{R}^d$*). O funcţie* $X : \Omega \to \mathbb{R}^d$, $(\mathcal{F}, \mathcal{B}_{\mathbb{R}^d})$ *măsurabilă, se numeşte vector aleator.*

Fie $X_i, \ i = 1, 2, \cdots, d$, componentele scalare ale funcţiei X, deci $X = (X_1, X_2, \cdots, X_d)$. Conform definiţiei de mai sus,

$$\{X_1 \leq a_1, X_2 \leq a_2, \cdots, X_d \leq a_d\} \in \mathcal{F}, \ \forall \ (a_1, a_2, \cdots, a_d) \in \mathbb{R}^d.$$

2.3 Funcţia de repartiţie

Distribuţia valorilor unei variabile aleatoare reale este caracterizată prin noţiunea de funcţie de repartiţie. De regulă, câmpul de probabilitate pe care este definită variabila aleatoare nu este menţionat în mod explicit.

Definiţia 2.3.1. *Fie* $(\Omega, \mathcal{F}, \mathbb{P})$ *un câmp de probabilitate şi* $X : \Omega \to \mathbb{R}$ *o variabilă aleatoare. Funcţia reală* $F : \mathbb{R} \to [0, 1]$, *definită prin*

$$F(x) = \mathbb{P}\{X \leq x\} = \mathbb{P}(\{\omega \in \Omega \,:\, X(\omega) \leq x\}),\ x \in \mathbb{R},$$

se numeşte funcţia de repartiţie a variabilei aleatoare X.

Propoziţia următoare grupează proprietăţile de bază ale unei funcţii de repartiţie.

Propoziţia 2.3.1. *Fie* $F : \mathbb{R} \to [0, 1]$ *funcţia de repartiţie a unei variabilei aleatoare* X. *Au loc următoarele proprietăţi cu caracter general.*

1. F *este monoton crescătoare pe* $\mathbb{R}$*;*
2. F *este continuă la dreapta pe mulţimea* $\mathbb{R}$ *(*$F(x + 0) := \lim\limits_{t \downarrow x} F(t) = F(x),\ \forall\, x \in \mathbb{R}$*);*
3. $\lim\limits_{x\to-\infty} F(x) = 0;\quad \lim\limits_{x\to\infty} F(x) = 1$*;*
4. $F(x - 0) := \lim\limits_{t\uparrow x} F(t) = \mathbb{P}\{X < x\},\ \forall\, x \in \mathbb{R}$*;*
5. $\mathbb{P}\{X = x\} = F(x) - F(x - 0),\ \forall\, x \in \mathbb{R}$*;*
6. $\mathbb{P}\{a < X \leq b\} = F(b) - F(a),\ \forall\, a, b \in \mathbb{R},$ *cu* $a < b$*;*
7. $\mathbb{P}\{X > x\} = 1 - F(x),\ \forall\, x \in \mathbb{R}$.

Demonstraţie. Fie $(\Omega, \mathcal{F}, \mathbb{P})$ un câmpul de probabilitate pe care este definităvariabila aleatoare X.
1. Fie $x_1, x_2 \in \mathbb{R}$, cu $x_1 < x_2$. Avem $\{X \leq x_1\} \subset \{X \leq x_2\}$, de unde $\mathbb{P}\{X \leq x_1\} \leq \mathbb{P}\{X \leq x_2\}$, sau $F(x_1) \leq F(x_2)$.
2. Fiind monoton crescătoare pe $\mathbb{R}$, funcţia F admite limite laterale finite pe $\mathbb{R}$. Mai precis, pentru $x \in \mathbb{R}$, *limita la stânga în* x este $F(x - 0) = \sup_{t<x} F(t)$, iar *limita la dreapta în* x este $F(x + 0) = \inf_{t>x} F(t)$. Evident,

$$F(x - 0) \leq F(x) \leq F(x + 0),\ \forall\, x \in \mathbb{R}.$$

Fixăm $x \in \mathbb{R}$, arbitrar. Notăm $A_n = \left\{X \leq x + \frac{1}{n}\right\} \in \mathcal{F}$, $\mathbb{N} \in \mathbb{N}^*$. Fie $A = \{X \leq x\}$. Avem $A = \cap_{n=1}^{\infty} A_n$ şi $A_n \downarrow A$. Conform Propoziţiei 1.5.1, avem $\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A)$. Dar, $\mathbb{P}(A) = F(x)$ şi $\lim_{n\to\infty} \mathbb{P}(A_n) = \lim_{n\to\infty} F\left(x + \frac{1}{n}\right) = F(x+0)$. Obţinem $F(x+0) = F(x)$, deci F este continuă la dreapta în x.
3. Conform monotoniei, F admite limite finite la ∞ şi $-\infty$. Astfel, $\lim\limits_{x\to-\infty} F(x) = \inf\limits_{x\in\mathbb{R}} F(x) \geq 0$ şi $\lim\limits_{x\to\infty} F(x) = \sup\limits_{x\in\mathbb{R}} F(x) \leq 1$. Fie $B_n = \{X \leq -n\}$, $C_n\{X \leq n\}$, $n \in \mathbb{N}^*$. Avem $B_n \downarrow \emptyset$, respectiv $C_n \uparrow \Omega$. Ca urmare,

$$\lim_{x\to-\infty} F(x) = \lim_{n\to\infty} F(-n) = \lim_{n\to\infty} \mathbb{P}(B_n) = \mathbb{P}(\emptyset) = 0,$$

respectiv
$$\lim_{x\to\infty} F(x) = \lim_{n\to\infty} F(n) = \lim_{n\to\infty} \mathbb{P}(C_n) = \mathbb{P}(\Omega) = 1.$$

4. Fie $x \in \mathbb{R}$, arbitrar, fixat. Conform proprietăţilor unei funcţii monoton crescătoare şi Propoziţiei 1.5.1, avem

$$F(x-0) = \lim_{t\uparrow x} F(t) = \lim_{n\to\infty} F\left(x - \frac{1}{n}\right) = \lim_{n\to\infty} \mathbb{P}\left\{X \le x - \frac{1}{n}\right\} = \mathbb{P}\{X < x\}.$$

5. Pentru $x \in \mathbb{R}$, avem

$$\mathbb{P}\{X = x\} = \mathbb{P}(\{X \le x\} \setminus \{X < x\}) = \mathbb{P}\{X \le x\} - \mathbb{P}\{X < x\} = F(x) - F(x-0).$$

6. Fie $a, b \in \mathbb{R},\ a < b$. Avem $\{a < X \le b\} = \{X \le b\} \setminus \{X \le a\}$ şi $\{X \le a\} \subset \{X \le b\}$. Atunci, conform Propoziţiei 1.5.1,

$$\mathbb{P}\{a < X \le b\} = \mathbb{P}\{X \le b\} - \mathbb{P}\{X \le a\} = F(b) - F(a).$$

7. Fie $x \in \mathbb{R}$. Avem $\mathbb{P}\{X > x\} = \mathbb{P}\{X \le x\}^c = 1 - \mathbb{P}\{X \le x\} = 1 - F(x)$. $\square$

Din propoziţia de mai sus rezultă că, dacă F este continuă în $x_0 \in \mathbb{R}$, atunci $\mathbb{P}\{X = x_0\} = 0$. Menţionăm că, reciproc, orice funcţie $F : \mathbb{R} \to [0,1]$ care satisface proprietăţile 1-3 din Propoziţia 2.3.1 este funcţia de repartiţie a unei anumite variabile aleatoare. Două sau mai multe variabile aleatore care admit o funcţie de repartiţie comună vor fi denumite *identic distribuite*.

De asemenea, putem defini funcţia de repartiţie a unui vector aleator, având proprietăţi generale apropiate de cele ale funcţiei de repartiţie a unei variabile aleatoare.

Definiţia 2.3.2. *Fie $X = (X_1, X_2, \cdots, X_d)$ un vector aleator. Definim funcţia sa de repartiţie prin $F : \mathbb{R}^d \to [0,1]$,*

$$F(x1, x_2, \cdots, x_d) = \mathbb{P}\{X_1 \le x_1, X_2 \le x_2, \cdots, X_d \le x_d\},\ (x_1, x_2, \cdots, x_d) \in \mathbb{R}^d.$$

.

O noţiune importantă este *independenţa variabilelor aleatoare*, definită mai jos.

Definiţia 2.3.3. *Variabilele aleatoare $X_i,\ i \in I$, definite pe un câmp de probabilitate $(\Omega, \mathcal{F}, \mathbb{P})$, se numesc independente dacă, pentru oricare oricare indici distincţi $i_1, i_2, \cdots, i_n$ din mulţimea I şi oricare $A_1, A_2, \cdots, A_n \in \mathcal{B}_{\mathbb{R}}$ are loc*

$$\mathbb{P}\{X_{i_1} \in A_1, X_{i_2} \in A_2, \cdots, X_{i_n} \in A_n\} = \mathbb{P}\{X_{i_1} \in A_1\}\mathbb{P}\{X_{i_2} \in A_2\} \cdots \mathbb{P}\{X_{i_n} \in A_n\}.$$

În particular, dacă variabilele aleatoare $X_1, X_2, \cdots, X_n$, având, respectiv, funcţiile de repartiţie $F_1, F_2, \cdots, F_n$, sunt independente, atunci

$$\mathbb{P}\{X_1 \le a_1, X_2 \le a_2, \cdots, X_n \le a_n\} = F_1(a_1)F_2(a_2)\cdots F_n(a_n),\ \forall\ a_1, a_2, \cdots, a_n \in \mathbb{R}.$$

Are loc următorul rezultat de bază privind suma a două variabile aleatoare independente.

Propoziţia 2.3.2. *Fie variabilele aleatoare independente X şi Y, având funcţiile de repartiţie F şi respectiv G. Atunci funcţia de repartiţie a variabilei aleatoare $Z = X + Y$ este $H = F * G$, unde*

$$(F * G)(x) = \int_{-\infty}^{\infty} F(x-t)dG(t), \ \forall \ x \in \mathbb{R}.$$

Variabilele aleatoare cu valori într-o mulţime cel mult numărabilă se numesc *discrete*. Variabilele aleatoare cu funcţia de repartiţie derivabilă, cu derivata continuă, se numesc *continue*. In secţiunea următoare vom discuta aceste tipuri particulare, dar deosebit de importante, de variabile aleatoare. Vom ilustra cele două tipuri de variabile aleatoare (discrete şi continue) prin exemple clasice, semnificative.

2.4 Variabile aleatoare discrete. Exemple clasice

În cele ce urmează, definim, caracterizăm şi exemplificăm variabilele aleatoare de tip discret.

Definiţia 2.4.1. *Fie $(\Omega, \mathcal{F}, \mathbb{P})$ un câmp de probabilitate şi $X : \Omega \to \mathbb{R}$ o variabilă aleatoare. X se numeşte variabilă aleatoare discretă dacă mulţimea $X(\Omega)$ este cel mult numărabilă.*

Fie $X(\Omega) = \{x_i, \ i \in I\} \subset \mathbb{R}$, unde I este o mulţime de indici cel mult numărabilă. Notăm $A_i = \{X = x_i\} \in \mathcal{F}, \ i \in I$. Remarcăm că familia de evenimente $\{A_i, i \in I\}$ reprezintă o partiţie a spaţiului Ω.

Pentru $A \in \mathcal{F}$, notăm prin $c_A : \Omega \to \mathbb{R}$ funcţia caracteriatică a mulţimii A, definită prin

$$c_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \in \Omega \setminus A = A^c \end{cases} .$$

Este elementar de verificat faptul că funcţia caracteristică a unui eveniment este o variabilă aleatoare. Are loc următoarea reprezentare a variabilei aleatoare discrete X:

$$X = \sum_{i \in I} x_i c_{A_i}.$$

Notuăm $p_i = \mathbb{P}(A_i), \ i \in I$. Avem $p_i \geq 0, \ \forall \ i \in I$ şi $\sum_{i \in I} p_i = 1$.

Distribuţia variabilei aleatoare X se reprezintă convenţional astfel:

$$X : \begin{pmatrix} x_i \\ p_i \end{pmatrix}_{i \in I}.$$

Fie variabilele aleatoare discrete, independente, X şi Y, cu distribuţiile

$$X : \begin{pmatrix} x_i \\ p_i \end{pmatrix}_{i \in I}, \ Y : \begin{pmatrix} y_j \\ p'_j \end{pmatrix}_{j \in J}.$$

Atunci variabila aleatoare $X + Y$ va avea distribuţia

$$X + Y : \begin{pmatrix} x_i + y_j \\ p_i p'_j \end{pmatrix}_{(i,j) \in I \times J}.$$

Exemplul 2.4.1. *Variabile aleatoare Bernoulli.*

O variabilă aleatoare Bernoulli X ia doar două valori: valoarea 1, cu probabilitatea $p \in (0,1)$ şi valoarea 0, cu probabilitatea $q = 1 - p$. Distribuţia lui X se reprezintă astfel:

$$X : \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}.$$

Notăm $X \in Bin(1,p)$.

Exemplul 2.4.2. *Variabile aleatoare binomiale.*

O variabilă aleatoare binomială X, de parametrii $n \in \mathbb{N}^*$ şi $p \in (0,1)$ ia valorile $k \in \{0, 1, \cdots, n\}$ cu probabilităţile $p_k = C_n^k p^k q^{n-k}$. Distribuţia lui X se reprezintă astfel:

$$X : \begin{pmatrix} 0 & 1 & 2 & \cdots & n \\ C_n^0 q^n & C_n^1 p q^{n-1} & C_n^2 p^2 q^{n-2} & \cdots & C_n^n p^n \end{pmatrix}.$$

Variabila aleatoare binomială indică numărul de realizări ale unui eveniment având probabilitatea p în n experienţe identice, independente (schema binomială - Bernoulli). Notăm $X \in Bin(n,p)$.

Este important de remarcat că o variabilă aleatoare binomială de parametrii (n,p) este suma a n variabile aleatoare Bernoulli independente, identic distribuite (de parametru p).

Exemplul 2.4.3. *Variabile aleatoare Poisson.*

O variabilă aleatoare Poisson de parametru λ ia valoarea n cu probabilitatea $\frac{\lambda^n}{n!} e^{-\lambda}$, pentru $n \in \mathbb{N}$, având deci distribuţia:

$$X : \begin{pmatrix} n \\ \frac{\lambda^n}{n!} e^{-\lambda} \end{pmatrix}_{n \in \mathbb{N}}.$$

Notăm $X \in Poiss(\lambda)$.

Exemplul 2.4.4. *Variabile aleatoare geometrice.*

O variabilă aleatoare geometrică X de parametru p ia valoarea n cu probabilitatea pq^{n-1}, pentru $n \in \mathbb{N}^*$, unde $q = 1 - p$. X are distribuţia:

$$X : \begin{pmatrix} n \\ pq^{n-1} \end{pmatrix}_{n \in \mathbb{N}^*}.$$

Notăm $X \in Geom(p)$. Variabila X indică numărul de experienţe, independente, identice, necesar realizării (apariţiei) unui eveniment care are probabilitatea p de realizare în cadrul fiecărei experienţe (schema geometrică).

2.5 Variabile aleatoare continue. Exemple clasice

Un tip important de variabile aleatoare sunt cele "continue".

Definiţia 2.5.1. *Fie* $(\Omega, \mathcal{F}, \mathbb{P})$ *un câmp de probabilitate şi* $X : \Omega \to \mathbb{R}$ *o variabilă aleatoare.* X *se numeşte variabilă aleatoare continuă dacă există o funcţie continuă* $f : \mathbb{R} \to [0, \infty)$ *cu proprietatea*

$$F(x) = \int_{-\infty}^{x} f(t)\, dt, \quad \forall\ x \in \mathbb{R}.$$

Funcţia f *se numeşte densitatea de probabilitate a variabilei aleatoare* X *(sau funcţia de densitate a v.a.* X*).*

Proprietăţile densităţii unei variabile aleatoare reale sunt evidenţite în propoziţia de mai jos.

Propoziţia 2.5.1. *Fie* $f : \mathbb{R} \to [0, \infty)$ *funcţia de densitate asociată unei variabile aleatoare de tip continuu* $X : \Omega \to \mathbb{R}$*. Notăm* $F : \mathbb{R} \to [0, 1]$ *funcţia sa de repartiţie. Atunci:*

1. F *este de clasă* $\mathcal{C}^1$ *pe* $\mathbb{R}$*, cu* $F'(x) = f(x),\ \forall\ x \in \mathbb{R}$*;*
2. $\int_{\mathbb{R}} f(t)\, dt = 1$*;*
3. $\mathbb{P}\{X = x\} = 0,\ \forall\ x \in \mathbb{R}$*;*
4. $\mathbb{P}\{a \leq X \leq b\} = \int_a^b f(t)\, dt,\ \forall\ [a, b] \subset \mathbb{R}.$

Demonstraţie.
1. Conform definiţiei, F este o primitivă a lui f pe $\mathbb{R}$.
2. $\int_{\mathbb{R}} f(t)\, dt = \lim\limits_{x\to\infty} \int\limits_{-\infty}^{x} f(t)\, dt = \lim\limits_{x\to\infty} F(x) = 1.$
3. F este continuă pe $\mathbb{R}$, deci $\mathbb{P}\{X = x\} = F(x) - F(x - 0) = 0,\ \forall\ x \in \mathbb{R}$.
4. $\mathbb{P}\{a \leq X \leq b\} = \mathbb{P}\{X \leq b\} - \mathbb{P}\{X < a\} = \mathbb{P}\{X \leq b\} - \mathbb{P}\{X \leq a\} = F(b) - F(a) =$
$= \int_a^b f(t)\, dt,\ \forall\ [a, b] \subset \mathbb{R}.$ □

Fie variabilele aleatoare continue, independente, X şi Y, cu densităţiile $f, g : \mathbb{R} \to [0, 1]$. Din Propoziţia 2.3.2 se deduce că variabila aleatoare $X + Y$ are densitatea $f * g$, unde

$$f * g : \mathbb{R} \to [0, \infty),\ (f * g)(x) = \int_{\mathbb{R}} f(x - t) g(t)\, dt,\ \forall\ x \in \mathbb{R},$$

este convoluţia densităţilor f şi g.

Vom prezenta în continuare câteva exemple clasice de repartiţii continue.

Exemplul 2.5.1. *Distribuţia normală.*

Este cel mai important exemplu de distribuţie de tip continuu. Astfel, spunem că o variabilă aleatoare reală X are o *repartiţie normală (gaussiană)* de parametrii 0 şi 1 dacă are funcţia de densitate:

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}, \quad \forall\ t \in \mathbb{R}.$$

Variabila normală (standard) X admite funcţie de repartiţie:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt, \quad \forall\ x \in \mathbb{R}.$$

Notăm $X \in N(0,1)$. Distribuţia normală standard este *legea limită universală*, conform *Teoremei limită centrală*.

Generalizare. $X \in N(\mu, \sigma^2)$, unde $\mu \in \mathbb{R},\ \sigma > 0$, dacă admite densitatea de probabilitate

$$\varphi(t) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(t-\mu)^2}{2\sigma^2}}, \quad \forall\ t \in \mathbb{R}.$$

Exemplul 2.5.2. *Distribuţia exponenţială.*

Variabila aleatoare X are o distribuţie exponenţială de parametru $\lambda > 0$, notat prin $X \in Exp(\lambda)$, dacă ia valori în intervalul $[0, \infty)$ (deci este o *variabilă aleatoare pozitivă*) şi admite funcţia de repartiţie

$$F(x) = 1 - e^{-\lambda x},\ \forall\ x \geq 0.$$

Distribuţia exponenţială este fundamentală în teoria fiabilităţii, caracterizând modelele de tip Markov.

Exemplul 2.5.3. *Distribuţia Gama.*

Variabila aleatoare X are o distribuţie Gama de parametrii $\alpha, \beta > 0$, notat prin $X \in Gama(\alpha, \beta)$, dacă ia valori în intervalul $(0, \infty)$ şi admite funcţia de densitate

$$f(x) = \frac{x^{\alpha-1} e^{-\frac{x}{\beta}}}{\beta^{\alpha}\Gamma(\alpha)},\ \forall\ x > 0,$$

unde

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$$

este funcţia lui Euler de prima speţă.

În statistică, distribuţia de tip $Gam(n/2, 2)$, cu $n \in \mathbb{N}^*$, notată prin χ_n^2, are o importanţă majoră.

Exemplul 2.5.4. *Distribuţia Beta.*

Variabila aleatoare X are o distribuţie Beta de parametrii $\alpha, \beta > 0$, notat prin $X \in Beta(\alpha, \beta)$, dacă ia valori în intervalul $(0, 1)$ şi admite funcţia de densitate

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)},\ \forall\ x \in (0, 1),$$

unde

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

este funcţia lui Euler de a doua speţă.

2.6 Media. Medii de ordin superior

Pentru o variabilă aleatoare reală, media şi dispersia reprezintă cele mai importante caracteristici numerice. Astfel, media unei variabile aleatoare (dacă există) indicăcea mai aşteptată valoare a variabilei aleatoare.

Definiţia 2.6.1. *Fie $(\Omega, \mathcal{F}, \mathbb{P})$ un câmp de probabilitate şi $X : \Omega \to \mathbb{R}$ o variabilă aleatoare având funcţia de repartiţie $F : \mathbb{R} \to [0, 1]$. Mărimea:*

$$\mathbb{E}(X) := \int_{-\infty}^{\infty} x\, dF(x)$$

se numeşte media lui X. Media este definită în ipoteza că integrala de tip Stieltjes-Riemann de mai sus este convergentă.

Calculul mediei se explicitează astfel:

1. pentru o variabilă aleatoare X de tip discret, care ia valorile x_i cu probabilităţile p_i, pentru $i \in I$, media devine
$$\mathbb{E}(X) = \sum_{i\in I} x_i p_i;$$

2. pentru o variabilă aleatoare X de tip continuu, care admite densitatea de probabilitate $f : \mathbb{R} \to [0, \infty)$, media devine
$$\mathbb{E}(X) := \int_{-\infty}^{\infty} x f(x)\, dx.$$

Următoarea propoziţie stabileşte proprietăţie de bază ale mediei variabilelor aleatoare.

Propoziţia 2.6.1. *Fie X şi Y două variabile aleatoare cu medie finită.*

1. $\mathbb{E}(aX) = a\mathbb{E}(X),\ \forall\, a \in \mathbb{R}$*;*

2. $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$*;*

3. *dacă X şi Y sunt independente, atunci $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.*

Demonstraţie. Vom trata doar cazul discret. Presupunem că X şi Y au distribuţiile

$$X : \begin{pmatrix} x_i \\ p_i \end{pmatrix}_{i\in I},\ Y : \begin{pmatrix} y_j \\ p'_j \end{pmatrix}_{j\in J},$$

unde I şi J sunt mulţimi de indici cel mult numărabile.
1. Pentru $a \in \mathbb{R}$, avem:

$$\mathbb{E}(aX) = \sum_{i\in I}(ax_i)p_i = a\sum_{i\in I} x_i p_i = a\mathbb{E}(X).$$

2. Fie $A_{ij} = \{X = x_i,\ Y = y_j\}$ şi $p_{ij} = \mathbb{P}(A_{ij}),\ (i, j) \in I \times J$. Atunci

$$\sum_{j\in J} p_{ij} = \sum_{j\in J} \mathbb{P}(A_{ij}) = \mathbb{P}(\cup_{j\in J} A_{ij}) = \mathbb{P}\{X = x_i\} = p_i,\ \forall\, i \in I,$$

respectiv
$$\sum_{i\in I} p_{ij} = \sum_{i\in I} \mathbb{P}(A_{ij}) = \mathbb{P}(\cup_{i\in I} A_{ij}) = \mathbb{P}\{Y = y_j\} = p'_j,\ \forall\ j \in J.$$
Deaoarece $X+Y$ are distribuţia
$$X+Y : \begin{pmatrix} x_i + y_j \\ p_{ij} \end{pmatrix}_{(i,j)\in I\times J},$$
obţinem
$$\mathbb{E}(X+Y) = \sum_{(i,j)\in I\times J} (x_i + y_j)p_{ij} = \sum_{i\in I}\left(x_i \sum_{j\in J} p_{ij}\right) + \sum_{j\in J}\left(y_j \sum_{i\in I} p_{ij}\right)$$
$$= \sum_{i\in I}(x_i p_i) + \sum_{j\in J}(y_j p'_j) = \mathbb{E}(X) + \mathbb{E}(Y).$$
3. Dacă X şi Y sunt independente, atunci $p_{ij} = p_i p'_j,\ \forall\ (i,j) \in I \times J$. Rezultă
$$\mathbb{E}(XY) = \sum_{(i,j)\in I\times J} (x_i y_j)p_{ij} = \sum_{(i,j)\in I\times J} (x_i p_i)(y_j p'_j)$$
$$= \sum_{i\in I}(x_i p_i) \sum_{j\in J}(y_j p'_j) = \mathbb{E}(X)\mathbb{E}(Y).$$
Astfel, proprietăţile mediei sunt demonstrate în cazul discret. $\square$

Fie $X : \Omega \to \mathbb{R}$ o variabilă aleatoare iar $\varphi : \mathbb{R} \to \mathbb{R}$ o funcţie reală măsurabilă (în particular, φ continuă). Funcţia compusă $\varphi(X) := \varphi \circ X : \Omega \to \mathbb{R}$ este de asemenea o variabilă aleatoare definită pe Ω. Dacă X admite funcţia de repartiţie F, atunci
$$\mathbb{E}(\varphi(X)) = \int_{-\infty}^{\infty} \varphi(x)\, dF(x),$$
în ipoteza că integrala este convergentă. În particular, în cazul când φ este o funcţie putere sau modulul unei funcţii putere, obţinem mediile de ordin superior, respectiv mediile absolute de ordin superior.

Definiţia 2.6.2. *Fie $(\Omega, \mathcal{F}, \mathbb{P})$ un câmp de probabilitate şi $X : \Omega \to \mathbb{R}$ o variabilă aleatoare.*

1. *Media de ordin $k \in \mathbb{N}^*$ a variabilei aleatoare X se defineţe prin*
$$\mathbb{E}\left(X^k\right) = \int_{-\infty}^{\infty} x^k dx,$$
dacă integrala este convergentă.

2. *Media absolută de ordin $k \in \mathbb{N}^*$ a variabilei aleatoare X se defineţe prin*
$$\mathbb{E}\left(|X|^k\right) = \int_{-\infty}^{\infty} |x|^k dx,$$
dacă integrala este convergentă. Notăm $L^k = L^k(\Omega)$ mulţimea variabilelor aleatoare (definite pe Ω) cu medie absolută finită de ordinul k.

Menţionăm că $L^k(\Omega)$ este un spaţiu vectorial normat. Din inegalitatea lui Hölder se deduc incluziunile

$$L^{k+1}(\Omega) \subset L^k(\Omega),\ k \in \mathbb{N}^*.$$

În plus, dacă $X \in L^k(\Omega)$, atunci X are medie finită de ordinele $i = 1, \cdots, n$.

Exemplul 2.6.1. *Mediile unor distribuţii clasice.*

1. *Distribuţia Bernoulli:* $X \in Bin(1,p) \ \Rightarrow\ \mathbb{E}(X) = p$.
2. *Distribuţia binomială:* $X \in Bin(n,p) \ \Rightarrow\ \mathbb{E}(X) = np$.
3. *Distribuţia Poisson:* $X \in Poiss(\lambda) \ \Rightarrow\ \mathbb{E}(X) = \lambda$.
4. *Distribuţia geometrică:* $X \in Geom(p) \ \Rightarrow\ \mathbb{E}(X) = 1/p$.
5. *Distribuţia normală:* $X \in N(\mu, \sigma^2) \ \Rightarrow\ \mathbb{E}(X) = \mu$.
6. *Distribuţia exponenţială:* $X \in Exp(\lambda) \ \Rightarrow\ \mathbb{E}(X) = 1/\lambda$.

2.7 Dispersia

Dispersia unei variabile aleatoare (din spaţiul L^2) este un indicator al gradului de împrăştiere a valorilor sale în jurul mediei.

Definiţia 2.7.1. *Fie $X \in L^2(\Omega)$, cu $\mathbb{E}(X) = \mu$. Mărimea:*

$$V(X) = \int_{-\infty}^{\infty} (x-\mu))^2\, dF(x) = \mathbb{E}\left[(X - \mathbb{E}(X))^2\right]$$

se numeşte dispersia lui X. Mărimea

$$\sigma(X) = \sqrt{V(X)}$$

se numeşte abaterea medie pătratică a variabilei aleatoare X.

Dispersia are următoarele proprietăţi elementare.

Propoziţia 2.7.1. *Fie $X, Y \in L^2(\Omega)$.*

1. $V(X) \geq 0$;
2. $V(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$ *(formula de calcul a dispersiei);*
3. $V(aX) = a^2 V(X),\ \forall\ a \in \mathbb{R}$;
4. *dacă X şi Y sunt v.a. independente, atunci $V(X+Y) = V(X) + V(Y)$.*

Demonstraţie. Proprietăţile 1 şi 3 sunt evidente. Fie $\mu = \mathbb{E}(X)$ şi $\eta = \mathbb{E}(Y)$. Formula de calcul a dispersiei se obţine astfel:

$$V(X) = \mathbb{E}\left[(X-\mu)^2\right] = \mathbb{E}\left(X^2 - 2\mu X + \mu^2\right) = \mathbb{E}\left(X^2\right) - 2\mu\mathbb{E}(X) + \mu^2 = \mathbb{E}\left(X^2\right) - \mu^2.$$

Conform Propoziţiei 2.6.1 şi formulei demonstrate avem:

$$V(X+Y) = \mathbb{E}\left[(X+Y)^2\right] - [\mathbb{E}(X+Y)]^2 = \mathbb{E}\left(X^2 + 2XY + Y^2\right] - (\mu+\eta)^2$$

$$= \mathbb{E}\left(X^2\right) + \mathbb{E}\left(Y^2\right) + 2\mu\eta - \mu^2 - 2\mu\eta - \eta^2 = \left[\mathbb{E}\left(X^2\right) - \mu^2\right] + \left[\mathbb{E}\left(Y^2\right) - \eta^2\right] = V(X) + V(Y).$$

□

Exemplul 2.7.1. *Dispersiile unor distribuţii clasice.*

1. *Distribuţia Bernoulli:* $X \in Bin(1,p) \;\Rightarrow\; V(X) = pq$.
2. *Distribuţia binomială:* $X \in Bin(n,p) \;\Rightarrow\; V(X) = npq$.
3. *Distribuţia Poisson:* $X \in Poiss(\lambda) \;\Rightarrow\; V(X) = \lambda$.
4. *Distribuţia geometrică:* $X \in Geom(p) \;\Rightarrow\; V(X) = q/p^2$.
5. *Distribuţia normală:* $X \in N(\mu, \sigma^2) \;\Rightarrow\; V(X) = \sigma^2$.

2.8 Corelaţia

Corelaţia a două variabile aleatoare este un indicator al tipului de dependenţă al acestora.

Definiţia 2.8.1. *Fie* $X, Y \in L^2(\Omega)$. *Corelaţia variabilelor aleatoare* X *şi* Y *se defineşte prin:*

$$\lambda(X,Y) = \mathbb{E}\{[X - \mathbb{E}(X)][Y - \mathbb{E}(Y)]\}.$$

Mărimea

$$\rho(X,Y) = \frac{\lambda(X,Y)}{\sigma(X)\sigma(Y)}$$

se numeşte coeficientul de corelaţie al variabilelor aleatoare X *şi* Y.

Propoziţia următoare stabileşte formula de calcul a corelaţiei şi o proprietate a coeficientului de corelaţie.

Propoziţia 2.8.1. *Cu notaţiile anterioare, avem:*

1. $\lambda(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$;
2. *dacă* X *şi* Y *sunt independente, atunci* $\lambda(X,Y) = 0$;
3. $\rho(X,Y) \in [-1,1]$.

Demonstraţie.
1. Avem

$$\lambda(X,Y)=\mathbb{E}\left(XY-\mu Y-\eta X+\mu\eta\right)=\mathbb{E}(XY)-\mu\mathbb{E}(Y)-\eta\mathbb{E}(X)+\mu\eta=\mathbb{E}(XY)-\mu\eta.$$

2. Din inegalitatea

$$\mathbb{E}\left\{[t(X-\mu)+(Y-\eta)]^2\right\}\geq 0,\ \forall\ t\in\mathbb{R},$$

sau

$$V(X)t^2+2\lambda(X,Y)t+V(Y)\geq 0,\ \forall\ t\in\mathbb{R},$$

obţinem

$$\Delta:=4\left(\lambda^2(X,Y)-V(X)V(Y)\right)\leq 0.$$

Rezultă $|\rho(X,Y)|\leq 1$. □

2.9 Funcţia caracteristică

Funcţia caracteristică asociată unei variabile aleatoare reprezintă un instrument matematic important. Pentru variabilele aleatoare de tip continuu, funcţia caracteristică reprezintă transformata Fourier a funcţiei de densitate. Transformarea integrală Fourier este importantă în *calculul operaţional.*

Definiţia 2.9.1. *Fie X o variabilă aleatoare având funcţia de repartiţie F. Funcţia φ_X : $\mathbb{R}\to\mathbb{C}$, definită prin*

$$\varphi_X(t)=\mathbb{E}\left(e^{itX}\right)=\int_{-\infty}^{\infty}e^{itx}dF(x),\ t\in\mathbb{R},$$

se numeşte functia caracteristică a variabilei aleatoare X.

Fie X o variabilă aleatoare de tip continuu, cu densitatea de probabilitate f. În acest caz, functia sa caracteristică se reprezintă astfel

$$\varphi_X(t)=\int_{-\infty}^{\infty}e^{itx}f(x)dx,\ t\in\mathbb{R},$$

adică φ_X reprezintă transformata Fourier a funcţiei f.

Dacă X este o variabilă aleatoare discretă, cu distribuţia

$$X:\begin{pmatrix}x_k\\ p_k\end{pmatrix}_{k\in I},$$

atunci funcţia sa caracteristică se calculează astfel

$$\varphi_X(t)=\sum_{k\in I}e^{itx_k}p_k.$$

Prezentăm în continuare principalele proprietăţi ale funcţiei caracteristice.

Teorema 2.9.1. *Fie X o variabilă aleatoare având funcţia de reaprtiţie F şi funcţia caracteristică $\varphi_X:\mathbb{R}\to\mathbb{C}$. Au loc următoarele proprietăţi:*

1. $\varphi_X(0) = 1$;

2. $\varphi(-t) = \overline{\varphi_X(t)},\ \forall\ t \in \mathbb{R}$;

3. $|\varphi_X(t)| \leq 1,\ \forall\ t \in \mathbb{R}$;

4. $\varphi_{aX+b}(t) = e^{ibt}\varphi_X(at), \forall\ t \in \mathbb{R},\ \forall\ a, b \in \mathbb{R}$;

5. φ_X *este uniform continuă (deci şi continuă) pe* $\mathbb{R}$;

6. dacă $X \in L^n(\Omega)$, *atunci* φ_X *este derivabilă de ordinul* n, *cu*

$$\varphi_X^{(n)}(t) = i^n \int_{-\infty}^{\infty} x^n e^{itx} dF(x),\ t \in \mathbb{R};$$

7. dacă $X \in \cap_{n=1}^{\infty} L^n(\Omega)$, *atunci* φ_X *este dezvoltabilă în serie Taylor în jurul originii:*

$$\varphi_X(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \varphi_X^{(n)}(0),$$

cu $\varphi_X^{(n)}(0) = i^n \mathbb{E}(X^n),\ n \in \mathbb{N}$.

Demonstraţie.
1. $\varphi_X(0) = \mathbb{E}(e^0) = \mathbb{E}(1) = 1$.
2. $\varphi_X(-t) = \mathbb{E}\left(e^{-itX}\right) = \mathbb{E}\left(\overline{e^{itX}}\right) = \overline{\varphi_X(t)},\ \forall\ t \in \mathbb{R}$.
3. $|\varphi_X(t)| = \left|\int_{-\infty}^{\infty} e^{itx} dF(x)\right| \leq \int_{-\infty}^{\infty} |e^{itx}|\, dF(x) = \int_{-\infty}^{\infty} 1 dF(x) = 1,\ \forall\ t \in \mathbb{R}$.
4. Fie $a, b \in \mathbb{R}$. $\varphi_{aX+b}(t) = \int_{-\infty}^{\infty} e^{it(ax+b)} dF(x) = e^{ibt} \int_{-\infty}^{\infty} e^{i(at)x} dF(x) e^{ibt} \varphi_X(at)$.
5. Fie $t \in \mathbb{R}$ şi $h > 0$. Avem

$$|\varphi_X(t+h) - \varphi_X(t)| = \left|\int_{-\infty}^{\infty} e^{itx}\left(e^{ihx} - 1\right) dF(x)\right|$$

$$\leq \int_{-\infty}^{\infty} \left|e^{itx}\right| \left|\left(e^{ihx} - 1\right)\right| dF(x) = 2 \int_{-\infty}^{\infty} \left|\sin \frac{hx}{2}\right| dF(x).$$

Fie $\varepsilon > 0$. Cum $\int_{-\infty}^{\infty} \left|\sin \frac{hx}{2}\right| dF(x) \leq \int_{-\infty}^{\infty} dF(x) = 1$, deducem că există $a > 0$ astfel ca

$$\int_{(-\infty,-a] \cup [a,\infty)} \left|\sin \frac{hx}{2}\right| dF(x) < \varepsilon/4.$$

Apoi, pentru $h < \varepsilon/2a$, avem

$$2 \int_{[-a,a]} \left|\sin \frac{hx}{2}\right| dF(x) \leq \int_{[-a,a]} h\,|x|\, dF(x) \leq ha \int_{-a}^{a} dF(x) \leq ha < \varepsilon/2.$$

Rezultă

$$|\varphi_X(t+h) - \varphi_X(t)| < \varepsilon,\ \forall\ h \in \left(0, \frac{\varepsilon}{2a}\right).$$

Ca urmare, φ_X este uniform continuă pe $\mathbb{R}$.
6. Se aplică teorema de derivare a integralelor improprii cu parametru.
7. Conform formulei anterioare, avem

$$\varphi_X^{(n)}(0) = i^n \int_{-\infty}^{\infty} x^n e^0 dF(x) = i^n \mathbb{E}\left(X^n\right).$$

Atunci, dezvoltând în serie Taylor exponenţiala, obţinem

$$\varphi_X(t) = \int_{-\infty}^{\infty} e^{itx} dF(x) = \int_{-\infty}^{\infty} \sum_{n=0}^{\infty} \frac{(itx)^n}{n!} dF(x) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \left[i^n \mathbb{E}\left(X^n\right)\right] = \sum_{n=0}^{\infty} \frac{t^n}{n!} \varphi_X^{(n)}(0),$$

pentru oricare $t \in \mathbb{R}$. □

Teorema 2.9.2. *(Teorema multiplicării) Dacă $X_1, X_2, \cdots, X_n$ sunt variabile aleatoare independente, atunci*

$$\varphi_{X_1+X_2+\cdots+X_n} = \varphi_{X_1}\varphi_{X_2}\cdots\varphi_{X_n}.$$

Demonstraţie. Media produsului unui număr finit de variabile aleatoare independente este egală cu produsul mediilor acestor variabile aleatoare. Atunci

$$\varphi_{\sum_{k=1}^n X_k}(t) = \mathbb{E}\left(e^{it\sum_{k=1}^n X_k}\right) = \mathbb{E}\left(\prod_{k=1}^n e^{itX_k}\right) = \prod_{k=1}^n \mathbb{E}\left(e^{itX_k}\right) = \prod_{k=1}^n \varphi_{X_k}(t),\ t \in \mathbb{R}.\ \square$$

Teorema 2.9.3. *(Teorema de unicitate) Fie X o variabilă aleatoare cu funcţia de repartiţie F şi funcţia caracteristică φ_X. Atunci*

1.

$$F(x_2) - F(x_1) = \mathbb{P}\{x_1 < X \le x_2\} = \frac{1}{\pi} \lim_{c\to\infty} \int_{-c}^{c} \frac{e^{-itx_1} - e^{-itx_2}}{it} \varphi_X(t)\, dt,\ \forall\ x_1 < x_2;$$

2.

$$F(x) = \frac{1}{2\pi} \lim_{y\to-\infty} \lim_{c\to\infty} \int_{-c}^{c} \frac{e^{-ity} - e^{-itx}}{it} \varphi_X(t)\, dt,\ \forall\ x \in \mathbb{R}.$$

Demonstraţie. Se aplică transformarea Fourier inversă. □

Exemplul 2.9.1. *Funcţiile caracteristice ale unor distribuţii clasice.*

1. *Distribuţia Bernoulli:* $X \in Bin(1,p) \ \Rightarrow\ \varphi_X(t) = 1 + p\left(e^{it} - 1\right),\ t \in \mathbb{R}$.
2. *Distribuţia binomială:* $X \in Bin(n,p) \ \Rightarrow\ \varphi_X(t) = \left[1 + p\left(e^{it} - 1\right)\right]^n,\ t \in \mathbb{R}$.
3. *Distribuţia Poisson:* $X \in Poiss(\lambda) \ \Rightarrow\ \varphi_X(t) = e^{\lambda\left(e^{it}-1\right)},\ t \in \mathbb{R}$.
4. *Distribuţia geometrică:* $X \in Geom(p) \ \Rightarrow\ \varphi_X(t) = \frac{pe^{it}}{pe^{it}-e^{it}+1},\ t \in \mathbb{R}$.
5. *Distribuţia normală:* $X \in N(\mu,\sigma^2) \ \Rightarrow\ \varphi_X(t) = e^{i\mu t - \frac{t^2\sigma^2}{2}},\ t \in \mathbb{R}$.
6. *Distribuţia exponenţială:* $X \in Exp(\lambda) \ \Rightarrow\ \varphi_X(t) = \frac{\lambda}{\lambda - it},\ t \in \mathbb{R}$.

2.10 Rezumat

Variabilele aleatoare se interpretează ca o "enumerare" a rezultatelor (măsurabile) potenţiale ale unui anumit experiment. Fiecare din aceste rezultate posibile are o probabilitate de realizare.

Fie $(\Omega, \mathcal{F}, \mathbb{P})$ un câmp de probabilitate. O funcţie măsurabilă reală $X : \Omega \to \mathbb{R}$ se numeşte *variabilă aleatoare*. Conform definiţiei, $\{X \in I\} \in \mathcal{F}$ pentru orice interval real I. In particular, $\{X \in (-\infty, a]\} = \{X \leq a\} \in \mathcal{F}, \ \forall\, a \in \mathbb{R}$.

O funcţie $X : \Omega \to \mathbb{R}^d$, $(\mathcal{F}, \mathcal{B}_{\mathbb{R}^d})$ măsurabilă, se numeşte *vector aleator*. Fie $X_i,\ i = 1, 2, \cdots, d$, componentele scalare ale funcţiei X, deci $X = (X_1, X_2, \cdots, X_d)$. Atunci

$$\{X_1 \leq a_1, X_2 \leq a_2, \cdots, X_d \leq a_d\} \in \mathcal{F},\ \forall\ (a_1, a_2, \cdots, a_d) \in \mathbb{R}^d.$$

Fie $X : \Omega \to \mathbb{R}$ o variabilă aleatoare. Funcţia reală $F : \mathbb{R} \to [0, 1]$, definită prin

$$F(x) = \mathbb{P}\{X \leq x\} = \mathbb{P}(\{\omega \in \Omega | X(\omega) \leq x\}),\ x \in \mathbb{R},$$

se numeşte *funcţia de repartiţie* a variabilei aleatoare X.

Funcţia de repartiţie are următoarele proprietăţi:

1. F este monoton crescătoare pe $\mathbb{R}$;
2. F este continuă la dreapta pe mulţimea $\mathbb{R}$ $(F(x + 0) := \lim\limits_{t \downarrow x} F(t) = F(x),\ \forall\ x \in \mathbb{R})$;
3. $\lim\limits_{x \to -\infty} F(x) = 0;\quad \lim\limits_{x \to \infty} F(x) = 1$;
4. $F(x - 0) := \lim\limits_{t \uparrow x} F(t) = \mathbb{P}\{X < x\},\ \forall\ x \in \mathbb{R}$;
5. $\mathbb{P}\{X = x\} = F(x) - F(x - 0),\ \forall\ x \in \mathbb{R}$;
6. $\mathbb{P}\{a < X \leq b\} = F(b) - F(a),\ \forall\ a, b \in \mathbb{R}$, cu $a < b$;
7. $\mathbb{P}\{X > x\} = 1 - F(x),\ \forall\ x \in \mathbb{R}$.

Variabilele aleatoare $X_i,\ i \in I$, se numesc independente dacă, pentru oricare oricare indici distincţi $i_1, i_2, \cdots, i_n$ din mulţimea I şi oricare $A_1, A_2, \cdots, A_n \in \mathcal{B}_{\mathbb{R}}$ are loc

$$\mathbb{P}\{X_{i_1} \in A_1, X_{i_2} \in A_2, \cdots, X_{i_n} \in A_n\} = \mathbb{P}\{X_{i_1} \in A_1\}\mathbb{P}\{X_{i_2} \in A_2\} \cdots \mathbb{P}\{X_{i_n} \in A_n\}.$$

X se numeşte variabilă aleatoare discretă dacă mulţimea $X(\Omega)$ este cel mult numărabilă. Fie $X(\Omega) = \{x_i,\ i \in I\} \subset \mathbb{R}$, unde I este o mulţime de indici cel mult numărabilă. Notăm $p_i = \mathbb{P}\{X = x_i\},\ i \in I$. Avem $p_i \geq 0,\ \forall\ i \in I$ şi $\sum_{i \in I} p_i = 1$. Distribuţia variabilei aleatoare X se reprezintă prin:

$$X : \begin{pmatrix} x_i \\ p_i \end{pmatrix}_{i \in I}.$$

Variabila aleatoare X se numeşte *de tip continuu* dacă există o funcţie continuă $f : \mathbb{R} \to [0, \infty)$ cu proprietatea

$$F(x) = \int_{-\infty}^{x} f(t)\, dt,\ \ \forall\ x \in \mathbb{R}.$$

Funcţia f se numeşte *densitatea de probabilitate* a variabilei aleatoare X (sau *funcţia de densitate* a v.a. X).

Densitatea de probabilitate are proprietăţile:

1. F este de clasă $\mathcal{C}^1$ pe $\mathbb{R}$, cu $F'(x) = f(x),\ \forall\ x \in \mathbb{R}$;

2. $\int_{\mathbb{R}} f(t)\, dt = 1$;

3. $\mathbb{P}\{X = x\} = 0,\ \forall\ x \in \mathbb{R}$;

4. $\mathbb{P}\{a \le X \le b\} = \int_a^b f(t)\, dt,\ \forall\ [a,b] \subset \mathbb{R}$.

Fie $X : \Omega \to \mathbb{R}$ o variabilă aleatoare având funcţia de repartiţie $F : \mathbb{R} \to [0,1]$. Mărimea:

$$\mathbb{E}(X) := \int_{-\infty}^{\infty} x\, dF(x)$$

se numeşte *media* lui X. Media este definită în ipoteza că integrala de tip Stieltjes-Riemann de mai sus este convergentă.

Calculul mediei:

1. pentru o variabilă aleatoare X de tip discret, care ia valorile x_i cu probabilităţile p_i, pentru $i \in I$:
$$\mathbb{E}(X) = \sum_{i \in I} x_i p_i;$$

2. pentru o variabilă aleatoare X de tip continuu, care admite densitatea de probabilitate $f : \mathbb{R} \to [0, \infty)$:
$$\mathbb{E}(X) := \int_{-\infty}^{\infty} x f(x)\, dx.$$

Media are următoarele proprietăţi:

1. $\mathbb{E}(aX) = a\mathbb{E}(X),\ \forall\ a \in \mathbb{R}$;

2. $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$;

3. dacă X şi Y sunt independente, atunci $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

Dacă X admite funcţia de repartiţie F, iar φ este o fucţie reală măsurabilă, atunci

$$\mathbb{E}(\varphi(X)) = \int_{-\infty}^{\infty} \varphi(x)\, dF(x),$$

în ipoteza că integrala este convergentă.

Definim:

1. *media de ordin* $k \in \mathbb{N}^*$ a variabilei aleatoare X:
$$\mathbb{E}\left(X^k\right) = \int_{-\infty}^{\infty} x^k dx,$$
dacă integrala este convergentă.

2. *media absolută de ordin* $k \in \mathbb{N}^*$ a variabilei aleatoare X:
$$\mathbb{E}\left(|X|^k\right) = \int_{-\infty}^{\infty} |x|^k dx,$$
dacă integrala este convergentă. Notăm $L^k = L^k(\Omega)$ mulţimea variabilelor aleatoare (definite pe Ω) cu medie absolută finită de ordinul k.

- Mărimea:

$$V(X) = \int_{-\infty}^{\infty} (x-\mu))^2\, dF(x) = \mathbb{E}(X-\mathbb{E}(X))^2 = \mathbb{E}(X^2) - \mathbb{E}^2(X)$$

  se numeşte *dispersia* lui X.

- Mărimea

$$\sigma(X) = \sqrt{V(X)}$$

  se numeşte *abaterea medie pătratică* a variabilei aleatoare X.

Dispersia are următoarele proprietăţi:

1. $V(X) \geq 0$;
2. $V(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$ (formula de calcul a dispersiei);
3. $V(aX) = a^2 V(X),\ \forall\ a \in \mathbb{R}$;
4. dacă X şi Y sunt v.a. independente, atunci $V(X+Y) = V(X) + V(Y)$.

Corelaţia a două variabile aleatoare este un indicator al tipului de dependenţă al acestora. Astfel, *corelaţia* variabilelor aleatoare X şi Y se defineşte prin:

$$\lambda(X,Y) = \mathbb{E}\{[X-\mathbb{E}(X)][Y-\mathbb{E}(Y)]\}.$$

Mărimea

$$\rho(X,Y) = \frac{\lambda(X,Y)}{\sigma(X)\sigma(Y)}$$

se numeşte *coeficientul de corelaţie* al variabilelor aleatoare X şi Y. Formula de calcul a corelaţiei $\lambda(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$.
Dacă X şi Y sunt v.a. independente, atunci $\lambda(X,Y) = 0$;
Avem $\rho(X,Y) \in [-1,1]$.

Fie X o variabilă aleatoare având funcţia de repartiţie F. Funcţia $\varphi_X : \mathbb{R} \to \mathbb{C}$, definită prin

$$\varphi_X(t) = \mathbb{E}\left(e^{itX}\right) = \int_{-\infty}^{\infty} e^{itx} dF(x),\ t \in \mathbb{R},$$

se numeşte functia caracteristică a variabilei aleatoare X.

Dacă X este o variabilă aleatoare de tip continuu, cu densitatea de probabilitate f, atunci

$$\varphi_X(t) = \int_{-\infty}^{\infty} e^{itx} f(x) dx,\ t \in \mathbb{R}.$$

Dacă X este o variabilă aleatoare discretă, cu distribuţia

$$X : \begin{pmatrix} x_k \\ p_k \end{pmatrix}_{k\in I},$$

atunci $\varphi_X(t) = \sum_{k\in I} e^{itx_k} p_k$.

2.11 Test

1. Definiţi funcţia de repartiţie a unei variabile aleatoare şi descrieţi proprietăţile acesteia. Discutaţi cazul variabilelor aleatoare independente.

2. Calculaţi media, dispersia şi funcţia caracteristică pentru o variabilă aleatoare:

   (a) Poisson;

   (b) exponenţială;

   (c) geometrică.

3. Variabila aleatoare X are distribuţia discretă:

$$X : \begin{pmatrix} 1 & 2 & 3 \\ p_1 & p_2 & p_3 \end{pmatrix},$$

   unde $p_1, p_2, p_3 > 0$, cu $p_1 + p_2 + p_3 = 1$. Ştiind că X are media $E(X) = \frac{7}{4}$ şi dispersia $V(X) = \frac{11}{16}$, să se determine probabilităţile p_1, p_2, p_3.

4. Să se determine constantele reale a şi b astfel ca funcţia $F(x) = a + b \operatorname{arctg} x,\ x \in \mathbb{R}$, să fie funcţia de repartiţie a unei variabile aleatoare X şi apoi să se calculeze $E(X^2)$.

5. Fie funcţia $f : \mathbb{R} \to \mathbb{R}$,

$$f(x) = \begin{cases} \frac{a}{\sqrt{1-x^2}}, & x \in (-1, 1) \\ 0, & x \in (-\infty, -1] \cup [1, \infty) \end{cases}.$$

   (a) Să se determine $a > 0$ astfel încât f să reprezinte densitatea de probabilitate a unei variabile aleatoare X.

   (b) Să se determine media $E(X)$ şi dispersia $V(X)$ a variabilei aleatoare X.

   (c) Să se calculeze $\mathbb{P}\left\{|X| \le \frac{1}{2}\right\}$.

6. Timpul de aşteptare într-o staţie de autobuz este o variabilă aleatoare cu funcţia de repartiţie

$$F(x) = \begin{cases} 0, & x < 0; \\ x/2, & 0 \le x < 1; \\ 1/2, & 1 \le x < 2; \\ x/4, & 2 \le x < 4; \\ 1, & x \ge 4. \end{cases}$$

   Să se determine:

   (a) probabilitatea ca o persoană să aştepte în staţie mai mult de 3 minute;

   (b) probabilitatea ca o persoană să mai aştepte în staţie cel puţin de 2 minute, după ce a aşteptat 1 minut.

7. Să se calculeze funcţia caracteristică a variabilei aleatoare X cu densitatea de probabilitate:

$$f(x) = \begin{cases} 0, & |x| \ge 2; \\ \frac{1}{2}\left(1 - \frac{|x|}{2}\right), & |x| < 2 \end{cases}.$$

Capitolul 3

Convergenţa şirurilor de variabile aleatoare

3.1 Introducere

Studiul convergenţa şirurilor de variabilelor aleatoare, în accepţiuni specifice teoriei probabilităţilor, evidenţiază fenomene asimptotice utile. Un şir de variabile aleatoare poate converge la o variabilă aleatoare (distribuţie limită) în sens "tare" sau în sens "slab". Convergenţa frecvenţei de realizare a unui eveniment, într-un şir de experienţe identice şi independente, la probabilitatea acelui eveniment este asigurată de *Legea numerelor mari.* Legea slabă a numerelor mari se deduce din inegalitatea lui Cebîşev. *Teorema limită centrală*, considerat cel mai important rezultat al teoriei probabilităţilor, stabileşte convergenţa la legea normală a distribuţiei normalizate a sumelor parţiale ale unui şir de variabile aleatoare independente şi identic distribuite. Rezultatele asimptotice menţionate au o importanţă majoră în *statistica matematică.*

3.2 Inegalităţi fundamentale în teoria probabilităţilor

Următoarele inegalităţi elementare joacă un rol fundamental în teoria probabilităţilor.

Teorema 3.2.1. *(Inegalitatea lui Markov) Fie $X \in L^n(\Omega)$, unde $n \in \mathbb{N}^*$. Atunci*

$$\mathbb{P}\{|X| > a\} \leq \frac{\mathbb{E}\left(|X|^n\right)}{a^n}, \ \forall\ a > 0.$$

Demonstraţie. Fie F funcţia de repartiţie a variabilei aleatoare $X \in L^n(\Omega)$. Pentru $a > 0$, avem

$$\mathbb{E}\left(|X|^n\right) = \int_{[-a,a]} |x|^n dF(x) + \int_{(-\infty,-a]\cup[a,\infty)} |x|^n dF(x) \geq a^n \int_{(-\infty,-a]\cup[a,\infty)} 1\, dF(x)$$

$$= a^n[F(-a) + 1 - F(a)] \geq a^n \mathbb{P}\{|X| > a\},$$

de unde concluzia. $\square$

Teorema 3.2.2. *(Inegalitatea lui Cebîşev (Chebyshev)) Fie $X \in L^2(\Omega)$, având media $\mathbb{E}(X) = \mu$ şi dispersia $V(X) = \sigma^2$. Atunci*

$$\mathbb{P}\{|X - \mu| > a\} \leq \frac{\sigma^2}{a^2},\ \forall\ a > 0.$$

Demonstraţie. Aplicăm inegalitatea lui Markov variabilei aleatoare $Z = X - \mu$, pentru $n = 2$. □
Inegalitatea lui Cebîşev se utilizează la demonstrarea *legii slabe a numerelor mari.*

Teorema 3.2.3. *(Inegalitatea lui Hölder) Fie $p, q > 1$ astfel ca $\frac{1}{p} + \frac{1}{q} = 1$. Pentru două variabile aleatoare X şi Y are loc inegalitatea*

$$\mathbb{E}(|XY|) \leq [\mathbb{E}(|X|^p)]^{\frac{1}{p}} \cdot [\mathbb{E}(|Y|^q)]^{\frac{1}{q}}.$$

În particular, pentru $Y = 1$, avem

$$\mathbb{E}(|X|) \leq [\mathbb{E}(|X|^p)]^{\frac{1}{p}},\ \forall\ p > 1.$$

Pentru $p = 2$, obţinem inegalitatea lui Cauchy-Buniakovski:

$$\mathbb{E}(|X|) \leq \sqrt{\mathbb{E}(X^2)}.$$

Inegalitatea de mai sus exprimă faptul că dispersia variabilei aleatoare $X \in L^2(\Omega)$ este pozitivă. Astfel,

$$V(X) = \mathbb{E}\left(X^2\right) - [\mathbb{E}(X)]^2 \geq \mathbb{E}\left(X^2\right) - [\mathbb{E}(|X|)]^2 \geq 0.$$

Teorema 3.2.4. *(Inegalitatea lui Lyapunov) Fie X o variabilă aleatoare şi $r, s \in \mathbb{N}^*$, cu $r < s$. Are loc inegalitatea*

$$[\mathbb{E}(|X|^r)]^{\frac{1}{r}} \leq [\mathbb{E}(|X|^s)]^{\frac{1}{s}}.$$

Demonstraţie. Fie $p = s/r > 1$ şi $q = p/(p-1) > 1$. Avem $\frac{1}{p} + \frac{1}{q} = 1$. Atunci, conform inegalităţii lui Hölder, obţinem

$$[\mathbb{E}(|X|^r)]^{\frac{1}{r}} \leq [\mathbb{E}(|X|^{pr})]^{\frac{1}{rp}} = [\mathbb{E}(|X|^s)]^{\frac{1}{s}}. \ \square$$

Consecinţă. $L^n(\Omega) \subset L^{n+1}(\Omega),\ \forall\ n \in \mathbb{N}^*$.

3.3 Tipuri de convergenţă. Relaţii

Vom defini *convergenţa şirurilor de variabile aleatoare*, în diverse accepţiuni. Studiul comportării asimptotice a şirurilor de variabile aleatoare reprezintă o preocupare majoră a teoriei probabilităţilor.

Definiţia 3.3.1. *Fie X o variabilă aleatoare şi $(X_n)_{n\geq 1}$ un şir de variabile aleatoare definite pe câmpul de probabilitate $(\Omega, \mathcal{F}, \mathbb{P})$. Fie F funcţia de repartiţie a variabilei aleatoare X, iar F_n funcţia de repartiţie a variabilei aleatoare X_n, $n \geq 1$. Definim următoarele tipuri de convergenţă ale şirului $(X_n)_{n\geq 1}$ către limita X:*

1. *şirul* $(X_n)_{n\geq 1}$ *converge în distribuţie (repartiţie) la* X*, notat prin*

$$X_n \xrightarrow{d} X,$$

dacă $\lim\limits_{n\to\infty} F_n(x_0) = F(x_0)$*, pentru orice punct de continuitate* $x_0 \in \mathbb{R}$ *al funcţiei* F*;*

2. *şirul* $(X_n)_{n\geq 1}$ *converge în probabilitate la* X*, notat prin*

$$X_n \xrightarrow{p} X,$$

dacă $\lim\limits_{n\to\infty} \mathbb{P}\{|X_n - X| > \varepsilon\} = 0,\ \forall\, \varepsilon > 0$*;*

3. *şirul* $(X_n)_{n\geq 1}$ *converge aproape sigur la* X*, notat prin*

$$X_n \xrightarrow{a.s.} X,$$

dacă $\mathbb{P}\{X_n \to X\} = 1$*;*

4. *şirul* $(X_n)_{n\geq 1}$ *converge în spaţiul* $L^k(\Omega)$ *la* X*, unde* $k \in \mathbb{N}^*$*, notat prin*

$$X_n \xrightarrow{L^k} X,$$

dacă $\lim\limits_{n\to\infty} \mathbb{E}\left(|X_n - X|^k\right) = 0.$

Convergenţele "în distribuţie" şi "în probabilitate" sunt *convergenţe slabe*, în timp ce convergenţele "aproape sigură" şi "în spaţiul $L^k(\Omega)$" sunt *convergenţe tari.* Această clasificare formală este susţinută de următoarea teoremă de ierarhizare a tipurilor de convergenţă definite anterior.

Teorema 3.3.1. *Fie* X *o variabilă aleatoare şi* $(X_n)_{n\geq 1}$ *un şir de variabile aleatoare definite pe câmpul de probabilitate* $(\Omega, \mathcal{F}, \mathbb{P})$*. Au loc implicaţiile*

1.
$$X_n \xrightarrow{p} X \implies X_n \xrightarrow{d} X;$$

2.
$$X_n \xrightarrow{a.s.} X \implies X_n \xrightarrow{p} X;$$

3.
$$X_n \xrightarrow{L^k} X \implies X_n \xrightarrow{p} X.$$

Implicaţiile reciproce sunt false.

Demonstraţie.
1. Presupunem $X_n \xrightarrow{p} X$. Fie $x_0 \in \mathbb{R}$ un punct de continuitate al funcţiei F. Considerăm $\varepsilon > 0$, arbitrar. Deoarece, pentru $n \geq 1$, avem

$$\{X \leq x_0 - \varepsilon\} \setminus \{|X_n - X| > \varepsilon\} \subset \{X \leq x_0 - \varepsilon\} \setminus \{X < X_n - \varepsilon\} \subset \{X_n \leq x_0\},$$

obţinem

$$F(x_0 - \varepsilon) - \mathbb{P}\{|X_n - X| > \varepsilon\} \leq \mathbb{P}(\{X \leq x_0 - \varepsilon\} \setminus \{|X_n - X| > \varepsilon\}) \leq F_n(x_0).$$

Pe baza ipotezei, $\lim_{n\to\infty} \mathbb{P}\{|X_n - X| > \varepsilon\} = 0$. Atunci

$$F(x_0 - \varepsilon) \leq \liminf_{n\to\infty} F_n(x_0). \tag{3.1}$$

Apoi, pentru $n \geq 1$, avem

$$\{X_n \leq x_0\} \setminus \{|X_n - X| > \varepsilon\} \subset \{X_n \leq x_0\} \setminus \{X_n < X - \varepsilon\} \subset \{X \leq x_0 + \varepsilon\},$$

de unde

$$F_n(x_0) - \mathbb{P}\{|X_n - X| > \varepsilon\} \leq \mathbb{P}(\{X_n \leq x_0\} \setminus \{|X_n - X| > \varepsilon\}) \leq F(x_0 + \varepsilon).$$

Pe baza ipotezei, rezultă

$$\limsup_{n\to\infty} F_n(x_0) \leq F(x_0 + \varepsilon). \tag{3.2}$$

Din (3.1), (3.2) şi continuitatea lui F în x_0, obţinem $\lim_{n\to\infty} F_n(x_0) = F(x_0)$. Deci $X_n \xrightarrow{d} X$.
2. Presupunem $X_n \xrightarrow{a.s.} X$. Avem

$$\omega \in \{X_n \to X\} \;\Leftrightarrow\; \forall\, \varepsilon > 0,\ \exists\, n \geq 1,\ \forall\, k \geq n:\ |X_n(\omega) - X(\omega)| < \varepsilon.$$

Conform ipotezei, rezultă

$$\mathbb{P}\left(\cup_{n=1}^{\infty} \cap_{p=0}^{\infty} \{|X_{n+p} - X| < \varepsilon\}\right) = 1,\ \forall\, \varepsilon > 0,$$

de unde

$$\lim_{n\to\infty} \mathbb{P}\left(\cap_{p=0}^{\infty}\{|X_{n+p} - X| < \varepsilon\}\right) = 1,\ \forall\, \varepsilon > 0.$$

Cum $\cap_{p=0}^{\infty}\{|X_{n+p} - X| < \varepsilon\} \subset \{|X_n - X| < \varepsilon\}$, vom deduce

$$\lim_{n\to\infty} \mathbb{P}\{|X_n - X| < \varepsilon\} = 1,\ \forall\, \varepsilon > 0,$$

sau, trecând la evenimentele contrare,

$$\lim_{n\to\infty} \mathbb{P}\{|X_n - X| \geq \varepsilon\} = 0,\ \forall\, \varepsilon > 0.$$

Rezultă $X_n \xrightarrow{p} X$.
3. Implicaţia rezultă din inegalitatea lui Markov. Astfel, pentru oricare $a > 0$, avem

$$\mathbb{P}\{|X_n - X| > a\} \leq \frac{\mathbb{E}\left(|X_n - X|^k\right)}{a^k} \to 0,\ \text{pentru } n \to \infty.$$

□

3.4 Legile numerelor mari

Legea numerelor mari reflectă următoare proprietate: *frecvenţa de apariţie a unui eveniment într-un şir de experienţe identice, independente, tinde către probabilitatea de realizare a evenimentului în fiecare experienţă.* Proprietatea a fost descrisă pentru prima dată la sfârşitul sec. al XVI-lea de matematicianul italian Gerolamo Cardano (1501-1576). Formalizarea matematică (pentru variabile aleatoare binare) ca o *lege a numerelor mari* i se datorează lui Jacob Bernoulli (publicată în lucrarea *Ars conjectandi*, 1713). În 1837, S.D. Poisson a descris-o sub numele de ”la Loi des grands nombres”. Contribuţii importante la demonstrarea riguroasă şi extinderea rezultatului sunt datorate lui Chebyshev (demonstrarea legii slabe a numerelor mari, 1837), Markov, Borel (demonstrarea legii tari a numerelor mari, 1909), Cantelli, Kolmogorov şi Khinchin.

Formalizarea Bernoulli a fenomenului observat de Cardano este următoarea. Să considerăm că, într-o experienţă, un anumit eveniment A se produce cu probabilitatea $p = \mathbb{P}(A)$. Fie şirul $(X_n)_{n\geq 1}$ de variabile aleatoare Bernoulli, independente şi identic distribuite (iid), asociate şirului de experienţe considerat, având distribuţia comună

$$X_n : \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix},\ n \in \mathbb{N}^*.$$

Notăm $S_n = X_1+X_2+\cdots+X_n$ care indică numărul de apariţii (realizări) ale evenimentului A în n experienţe. Frecvenţa de realizare a evenimentului A este deci S_n/n. Fenomenul descris de legea numerelor mari este prin urmare

$$\frac{S_n}{n} \to p = \mathbb{E}(X_1).$$

Conergenţa frecvenţei de apariţie a evenimentului la probabilitatea sa se poate realiza în sensul unui tip de convergenţă slab (lege slabă a numerelor mari) sau în sensul unui tip de convergenţă tare (lege tare a numerelor mari). Rezultatul este extins la variabile aleatoare cu medie şi dispersie finite.

Teorema 3.4.1. *(Legea slabă a numerelor mari)*
Fie $(X_n)_{n\geq 1}$ un şir de variabile aleatoare independente şi identic distribuite (iid), cu $\mathbb{E}(X_1) = \mu$ şi $V(X_1) = \sigma^2$. Notăm $S_n = X_1 + X_2 + \cdots + X_n$, $n \in \mathbb{N}^*$. Atunci are loc convergenţa în probabilitate

$$\frac{S_n}{n} \xrightarrow{p} \mu.$$

Demonstraţie. Vom interpreta $\mu = \mathbb{E}(X_n),\ n \in \mathbb{N}^*$, ca o variabilă aleatoare constantă, care ia valoarea μ cu probabilitatea 1. Avem $\mathbb{E}(S_n) = \sum_{k=1}^n \mathbb{E}(X_k) = n\mu$ şi $V(S_n) = \sum_{k=1}^n V(X_k) = n\sigma^2$. Fie $\varepsilon > 0$, arbitrar. Conform inegalităţii lui Chebyshev,

$$\mathbb{P}\left\{\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right\} = \mathbb{P}\left\{|S_n - \mathbb{E}(S_n)| > n\varepsilon\right\} \leq \frac{V(S_n)}{n^2\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}.$$

Rezultă

$$\lim_{n\to\infty} \mathbb{P}\left\{\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right\} = 0.$$

Cum $\varepsilon > 0$ este arbitrar, obţinem concluzia. $\square$

Menţinăm că legea slabă a numerelor poate fi formulată pentru şiruri de variabile aleatoare independente (dar neidentic distribuite) cu dispersii uniform marginite (Chebyshev), respectiv pentru siruri de variabile de v.a. cu proprietatea $V(S_n)/n^2 \to 0$ (Markov). Prezentăm (fără) demonstraţie şi o versiune a *legii tari a numerelor mari.*

Teorema 3.4.2. *(Legea tare a numerelor mari)*
Fie $(X_n)_{n\geq 1}$ un şir de variabile aleatoare independente şi identic distribuite (iid), din spaţiul L^4 ($\mathbb{E}(|X_1|^4) < \infty$) cu $\mathbb{E}(X_1) = \mu$. Fie $S_n = X_1 + X_2 + \cdots + X_n,\ n \in \mathbb{N}^*$. Atunci

$$\frac{S_n}{n} \xrightarrow{a.s.} \mu.$$

3.5 Teorema limită centrală

Teorema limită centrală, în versiunea clasică, stabileşte că sumele parţiale "normalizate" ale unui şir de variabile aleatoare iid, cu dispersie finită, tind în distribuţie către legea normală (gaussiană). Rezultatul admite numeroase extinderi relativ la şiruri de variabile aleatoare heterogene, neindependente, supuse unor anumite condiţii. De asemenea, rezultatul se extinde la şiruri de vectori aleatori.

Prima versiune a acestei teoreme a fost formulată de Abraham de Moivre (1733) pentru variabile aleatoare de tip Bernoulli, cu $p = 1/2$). Pierre-Simon Laplace generalizează rezultatul în lucrarea *Theorie Analytique des Probabilités* (1812), constatând apropierea distribuţiei binomiale de distribuţia normală. În 1901, matematicianul rus Aleksandr Lyapunov a definit în termeni generali şi a demonstrat riguros *teorema limită centrală*. Teorema limită centrală este considerată rezultatul principal al teoriei probabilităţilor.

Enunţăm pentru început rezultatul datorat lui de Moivre şi Laplace relativ la şiruri iid de v. a. Bernoulli.

Teorema 3.5.1. *(Teorema de Moivre-Laplace) Fie $(X_n)_{n\geq 1}$ un şir de variabile aleatoare Bernoulli independente şi identic distribuite (iid), de parametru p, având media comună p şi dispersia comună pq (unde q=1-p). Pentru $n \in \mathbb{N}^*$, notăm $S_n = \sum_{k=1}^n X_k$ şi*

$$I_n = \left\{ \frac{k - np}{\sqrt{npq}},\ k = 0, 1, \cdots, n \right\}.$$

Atunci, pentru oricare interval $[a, b] \subset \mathbb{R}$, are loc:

$$\lim_{n\to\infty} \left(\sup_{x\in I_n\cap[a,b]} \left| \sqrt{2\pi npq} \cdot \mathbb{P}\left\{ \frac{S_n - np}{\sqrt{npq}} = x \right\} - e^{-\frac{x^2}{2}} \right| \right) = 0$$

Demonstraţia se bazează pe aproximarea factorialului dată de formula lui Stirling

$$n! = \sqrt{2n\pi} \left(\frac{n}{e}\right)^n e^{\theta_n},$$

unde $0 < \theta_n < \frac{1}{12n},\quad n \geq 1$. Din Teorema de Moivre-Laplace se deduce o versiune particulară a Teoremei limită centrală. Enunţul clasic general al acestei teoreme este prezentat în continuare.

Teorema 3.5.2. *(Teorema limită centrală) Fie $(X_n)_{n\geq 1}$ un şir de variabile aleatoare independente şi identic distribuite (iid), cu $\mathbb{E}(X_1) = \mu$ şi $V(X_1) = \sigma^2$. Considerăm şirul de variabile aleatoare $S_n = X_1 + X_2 + \cdots + X_n,\ n \in \mathbb{N}^*$. Atunci:*

1.

$$\lim_{n\to\infty} \mathbb{P}\left\{ a \leq \frac{S_n - n\mu}{\sigma\sqrt{n}} \leq b \right\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}},\ \forall\ a, b \in \mathbb{R},\ a < b;$$

2. *(versiunea Lindeberg-Lévy) Are loc următoarea convergenţa în distribuţie către o variabilă aleatoare gaussiană:*

$$\frac{\sqrt{n}}{\sigma}\left(\frac{S_n}{n}-\mu\right)\xrightarrow{d}X,$$

unde $X\in N(0,1)$.

Demonstraţie. (schiţă)
Considerăm şirul de v. a. iid $Y_n=\frac{X_n-\mu}{\sigma}$, $n\geq 1$, de medie comună 0 şi dispersie comună 1. Atunci, conform Teoremei 2.9.1, 7, funcţia caracteriatică a variabilelor aleatoare Y_n este de forma

$$\varphi_{Y_n}(t)=1-\frac{t^2}{2}+0\left(t^2\right),\ t\to 0.$$

Fie

$$Z_n=\frac{\sum_{k=1}^{n}Y_k}{\sqrt{n}}=\frac{S_n-n\mu}{\sigma\sqrt{n}},\ n\geq 1.$$

Pe baza proprietăţilor funcţiei caracteristice (Teorema 2.9.1), obţinem

$$\varphi_{Z_n}(t)=\left[\varphi_{Y_1}\left(\frac{t}{\sqrt{n}}\right)\right]^n=\left[1-\frac{t^2}{2n}+0\left(\frac{t^2}{n}\right)\right]^n\overset{n\to\infty}{\longrightarrow}e^{-\frac{t^2}{2}}.$$

Dar, pentru $X\in N(0,1)$, avem $\varphi_X(t)=e^{-\frac{t^2}{2}}$, $t\in\mathbb{R}$ (a se vedea Exemplul 3.10.1). Conform teoremei de unicitate (Teorema 2.9.2) se obţine $Z_n\xrightarrow{d}X$. Astfel, teorema limită centrală este demonstrată. $\square$

3.6 Rezumat

Inegalitatea lui Markov Fie $X \in L^n(\Omega)$, unde $n \in \mathbb{N}^*$. Atunci

$$\mathbb{P}\{|X| > a\} \leq \frac{\mathbb{E}\left(|X|^n\right)}{a^n}, \ \forall \ a > 0.$$

Inegalitatea lui Cebîşev (Chebyshev) Fie $X \in L^2(\Omega)$, având media $\mathbb{E}(X) = \mu$ şi dispersia $V(X) = \sigma^2$. Atunci

$$\mathbb{P}\{|X - \mu| > a\} \leq \frac{\sigma^2}{a^2}, \ \forall \ a > 0.$$

Definim următoarele tipuri de convergenţă ale şirului $(X_n)_{n\geq 1}$ către limita X:

1. şirul $(X_n)_{n\geq 1}$ converge în distribuţie (repartiţie) la X, notat prin

$$X_n \stackrel{d}{\to} X,$$

   dacă $\lim\limits_{n\to\infty} F_n\left(x_0\right) = F\left(x_0\right)$, pentru orice punct de continuitate $x_0 \in \mathbb{R}$ al funcţiei F;

2. şirul $(X_n)_{n\geq 1}$ converge în probabilitate la X, notat prin

$$X_n \stackrel{p}{\to} X,$$

   dacă $\lim\limits_{n\to\infty} \mathbb{P}\left\{|X_n - X| > \varepsilon\right\} = 0, \ \forall \ \varepsilon > 0;$

3. şirul $(X_n)_{n\geq 1}$ converge aproape sigur la X, notat prin

$$X_n \stackrel{a.s.}{\to} X,$$

   dacă $\mathbb{P}\left\{X_n \to X\right\} = 1;$

4. şirul $(X_n)_{n\geq 1}$ converge în spaţiul $L^k(\Omega)$ la X, unde $k \in \mathbb{N}^*$, notat prin

$$X_n \stackrel{L^k}{\to} X,$$

   dacă $\lim\limits_{n\to\infty} \mathbb{E}\left(|X_n - X|^k\right) = 0.$

Au loc implicaţiile

- $X_n \stackrel{p}{\to} X \implies X_n \stackrel{d}{\to} X;$
- $X_n \stackrel{a.s.}{\to} X \implies X_n \stackrel{p}{\to} X;$
- $X_n \stackrel{L^k}{\to} X \implies X_n \stackrel{p}{\to} X.$

Implicaţiile reciproce sunt false.

Legea slabă a numerelor mari Fie $(X_n)_{n\geq 1}$ un şir de variabile aleatoare independente şi identic distribuite (iid), cu $\mathbb{E}(X_1) = \mu$ şi $V(X_1) = \sigma^2$. Notăm $S_n = X_1 + X_2 + \cdots + X_n,\ n \in \mathbb{N}^*$. Atunci are loc convergenţa în probabilitate

$$\frac{S_n}{n} \xrightarrow{p} \mu.$$

Legea tare a numerelor mari Fie $(X_n)_{n\geq 1}$ un şir de variabile aleatoare independente şi identic distribuite (iid), din spaţiul L^4 ($\mathbb{E}(|X_1|^4) < \infty$) cu $\mathbb{E}(X_1) = \mu$. Fie $S_n = X_1 + X_2 + \cdots + X_n,\ n \in \mathbb{N}^*$. Atunci

$$\frac{S_n}{n} \xrightarrow{a.s.} \mu.$$

Teorema limită centrală Fie $(X_n)_{n\geq 1}$ un şir de variabile aleatoare independente şi identic distribuite (iid), cu $\mathbb{E}(X_1) = \mu$ şi $V(X_1) = \sigma^2$. Considerăm şirul de variabile aleatoare $S_n = X_1 + X_2 + \cdots + X_n,\ n \in \mathbb{N}^*$. Atunci:

1. $$\lim_{n\to\infty} \mathbb{P}\left\{a \leq \frac{S_n - n\mu}{\sigma\sqrt{n}} \leq b\right\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}},\ \forall\ a, b \in \mathbb{R},\ a < b;$$

2. (versiunea Lindeberg-Lévy) Are loc următoarea convergenţa în distribuţie către o variabilă aleatoare gaussiană:

   $$\frac{\sqrt{n}}{\sigma}\left(\frac{S_n}{n} - \mu\right) \xrightarrow{d} X,$$

   unde $X \in N(0,1)$.

3.7 Test

1. Să se enunţe şi să se demonstreze Legea slabă a numerelor mari.

2. Să se enunţe Teorema limită centrală, adaptată pentru variabile aleatoare de tip geometric.

3. Să se determine funcţia de repartiţie a variabilei aleatoare X care admite funcţia caracteristică
$$\varphi_X(t) = \frac{1}{4}\left(1 + e^{it}\right)^2.$$

4. Inegalitatea lui Cebîşev aplicată variabilei aleatoare X, cu media $E(X) = 3$ şi dispersia $V(X) = \sigma^2$, stabileşte:
$$\mathbb{P}\left\{|X - 3| \geq 2\right\} \leq 0,16.$$
Să se determine σ.

5. La un concurs de biatlon, un sportiv nimereste ţinta cu probabilitatea de $\frac{4}{5}$. Ştiind că sportivul execută în concurs 25 de trageri, să se arate că şansa ca acesta să nimerească ţinta de cel puţin 15 ori este mai mare de 85%.

6. Probabilitatea ca un monitor de calculator să nu aibă rezoluţia acceptabilă este de 0,1. S-au cumpărat 1000 de monitoare. Care este probabilitatea ca mai mult de 100 monitoare să nu fie acceptabile? (Notă: se va utiliza aproximarea oferită de Teorema limită centrală.)

Capitolul 4

Elemente de statistică matematică

4.1 Introducere

Statistica matematica este o ramură importantă a matematicii, care utilizează rezultatele oferite de teoria probabilităţilor. Progresul în societate, în particular în ştiinţă, este adeseori datorat *experimentului.* Cercetătorul realizează o experienţă şi obţine o serie de date, pe baza cărora generealizează rezultatele experienţei la o clasă de experienţe similare. Acest tip de extindere se datorează unui raţionament de tip *inductiv* (obţinerea unor informaţii generale din analiza unor informaţii particulare). Statistica analizează date concrete, locale, obţinute experimental, pentru a prognoza date cu caracter general, a descoperi legile care guvernează fenomemul aleator studiat. În acelaşi timp, statistica se preocupă de verosimilitatea matematică a prognozei, măsurând cu mijloace matematice gradul de încredere în rezultatele generale propuse.

Prezentarea de faţă are ca obiectiv familiarizarea cu câteva noţiuni de bază din teoria statisticii matematice, precum: populaţie statistică, sondaj, eşantion, estimator şi caracteristicile sale. Pentru simplificarea expunerii, ne vom rezuma la descrierea unor elemente de *statistica eşantioanelor Bernoulli.*

4.2 Noţiuni de bază

Statistica operează cu noţiuni specifice. Astfel, ***populaţia statistică*** reprezintă mulţimea tuturor elementelor avute în vedere în cadrul unui ***studiu statistic***. De regulă, pentru o populaţie mare, studiul statistic se realizează prin ***sondaj***. Sondajul constă în observarea (examinarea, chestionarea) unei părţi a populaţiei, numită ***eşantion***. Numărul elementelor unui eşantion (obţinut prin sondaj şi examinat) se numeşte ***volumul eşantionului***. De regulă, volumul eşantionului se stabileşte (pe baza unor evaluări matematice) astfel încât să poată furniza date generale concludente/veridice. Scopul studiului statistic este ***estimarea*** unei proprietăţii/particularităţi a populaţiei statistice. Funcţia matematică care realizează estimarea se numeşte ***estimator***. Apelând la modelul particular al eşantioanelor Bernoulli, vom trece în revistă calităţile impuse unui estimator de calitate.

4.3 Estimatori. Proprietăţi

Cel mai simplu studiu statistic se referă la analiza cazului binar: fiecare individ (element) al populaţiei statistice are valoarea 1 sau 0 (răspunde cu 1 sau 0 la un chestionar). Ne interesează proporţia indivizilor cu valoarea 1 în cadrul întregii populaţii statistice. Examinarea/chestionarea indivizilor prin sondaj se modelează printr-o variabilă aleatoare "bservabilă", care indică numărul indivizilor care valorează 1 într-un eşantion de volum n. La o populaţie "mică", este logic să apelăm la modelul oferit de legea hipergeometrică, întrucât extragerea unui eşantion dintr-o populaţie mică afectează în mod cert proporţia valorii 1 în cadrul populaţiei nesondate. În schimb, dacă volumul populaţiei este "mare" în raport cu volumul eşantionului atunci este rezonabil să consideră că fiecare element al populaţiei statistice se reprezintă printr-o variabilă aleatoare Bernoulli X, cu distribuţia

$$X : \begin{pmatrix} 0 & 1 \\ 1-\theta & \theta \end{pmatrix},$$

unde parametrul θ (necunoscut) reprezintă proporţia valorii 1 în populaţia statistică. Astfel, $\mathbb{P}\{X = 1\} = \theta$ reprezintă şansa ca elementul X să aibă valoarea 1. Evident, $\mathbb{P}\{X = 0\} = 1 - \theta$ reprezintă şansa ca X să ia valoarea 0. Obiectul studiului statistic întreprins constă în *estimarea* parametrului θ, din observarea unui eşantion de volum n, $(X_1, X_2, \cdots, X_n)$, reprezentând un vector aleator cu componentele variabile aleatoare iid, de tip $Bin(1, \theta)$. ***Distribuţia eşantionului*** se defineşte distribuţia comună a v. a. $X_1, X_2, \cdots, X_n$.

O ***statistică*** (sau un ***estimator***) se defineşte ca o funcţie a variabilelor aleatoare "observabile" (cele ale n-eşantionului), deci care nu depinde de parametrii necunoscuţi.

Pentru evaluarea apriorică a parametrului necunoscut θ se propune estimatorul ***media empirică***, definit prin:

$$\overline{X} = \frac{S_n}{n}, \text{ unde } S_n = X_1 + X_2 + \cdots + X_n.$$

În fapt, media empirică exprimă frecvenţa de apariţie a valorii 1, deci, intuitiv, pare a oferi o bună evaluare a parametrului θ. Descriem în continuare proprietăţile principale ale mediei empirice $\overline{X}$, ca estimator a lui θ.

1. *$\overline{X}$ este un estimator fără abatere a lui θ: $\mathbb{E}_\theta(\overline{X}) = \theta, \ \forall\, \theta \in (0, 1)$.*

Demonstraţie. $\mathbb{E}_\theta(\overline{X}) = \left[\sum_{k=1}^n \mathbb{E}_\theta(X_k)\right]/n = (n\theta)/n = \theta.$

2. *$\overline{X}$ este unicul estimator fără abatere a lui θ, care depinde de S_n.*

Demonstraţie. Fie $g : \{0, 1, 2, \cdots, n\} \to \mathbb{R}$, astfel încât

$$\theta = \mathbb{E}_\theta\left[g(S_n)\right] = \theta, \ \forall\, \theta \in (0, 1). \tag{4.1}$$

Avem

$$\mathbb{E}_\theta\left[g(S_n)\right] = \sum_{k=0}^n g(k)\mathbb{P}\{S_n = k\} = \sum_{k=0}^n g(k) C_n^k \theta^k (n-\theta)^{1-k},$$

iar pe de altă parte

$$\theta = \mathbb{E}_\theta(\overline{X}) = \mathbb{E}_\theta\left(\frac{S_n}{n}\right) = \sum_{k=0}^{n} \frac{k}{n}\mathbb{P}\{S_n = k\} = \sum_{k=0}^{n} \frac{k}{n} C_n^k \theta^k (1-\theta)^{n-k}.$$

Atunci, conform (4.1), obţinem

$$\sum_{k=0}^{n} \left(g(k) - \frac{k}{n}\right) C_n^k \left(\frac{\theta}{1-\theta}\right)^k = 0, \ \forall\, \theta \in (0,1).$$

Rezultă că polinomul $f = \sum_{k=0}^{n} \left(g(k) - \frac{k}{n}\right) C_n^k X^k$ este identic nul.
Prin urmare $g(k) = k/n, \ k = 0, 1, \cdots, n$, deci $g(S_n) = \overline{X}$.

3. *$\overline{X}$ este estimatorul lui θ de maximă verosimilitate.*

Demonstraţie. Să presupunem că au fost observate valorile $(x_1, x_2, \cdots, x_n) \in \{0,1\}^n$. Notăm $s_n = \sum_{n=1}^{n} x_k$. Probabilitatea de a observa aceste valori este $\theta^{s_n}(1-\theta)^{n-s_n}$. Atunci, cea mai *verosimilă* valoare pe care o poate lua θ este cea care maximizează funcţia

$$L : [0,1] \to [0,1], \ L(u) = u^{s_n}(1-u)^{n-s_n}.$$

Presupunem $0 < s_n < n$. Avem

$$L'(u) = L(u)\frac{s_n - nu}{u(1-u)}.$$

Rădăcina derivatei lui L este $u_0 = \frac{s_n}{n}$. Urmărind semnul derivatei, deducem că L îşi atinge maximul în u_0. Rezultă că $\overline{X} = S_n/n$ este estimatorul de verosimilitate maximă.

4. *Când talia eşantionului tinde la infinit, $\overline{X}$ converge la θ cu o rată exponenţială.*

Demonstraţie. Vom remarca în primul rând că, pe baza inegalităţii lui Chebyshev,

$$\overline{X} \xrightarrow{p} \theta \text{ (legea slabă numerelor mari).}$$

Pe de altă parte, conform *Teoremei limită centrală* (Teorema 3.5.2),

$$\sqrt{\frac{n}{\theta(1-\theta)}}\left(\overline{X} - \theta\right) \xrightarrow{d} X,$$

unde $X \in N(0,1)$.

Faptul că rata convergenţei este exponenţială se deduce din *Teorema marilor deviaţii*, pe care nu o vom include în această prezentare.

5. *Riscul pătratic mediu al estimatorului $\overline{X}$ a lui θ converge la 0.*

Demonstraţie. Riscul pătratic mediu al estimatorului $\overline{X}$ este defint prin $\mathbb{E}_\theta\left(\overline{X}-\theta\right)^2$. Avem

$$\mathbb{E}_\theta\left(\overline{X}-\theta\right)^2=\mathbb{E}_\theta\left(\frac{S_n}{n}-\theta\right)^2=\frac{1}{n^2}\mathbb{E}_\theta\left(S_n-n\theta\right)^2=\frac{n}{n^2}V(X_1)=\frac{\theta(1-\theta)}{n}\leq\frac{1}{4n},$$

de unde $\mathbb{E}_\theta\left(\overline{X}-\theta\right)^2\overset{n\to\infty}{\to}0$.

6. *Abaterea lui $\overline{X}$ faţă de θ poate fi evaluată prin Teorema limită centrală.*

Demonstraţie. Fie $a>0$. Pentru $n\geq\frac{4\theta(1-\theta)}{a^2}$, avem $\frac{a\sqrt{n}}{\sqrt{\theta(1-\theta)}}\geq 2$, de unde

$$\mathbb{P}_\theta\{|\overline{X}-\theta|\geq a\}=\mathbb{P}\left\{\left|\frac{S_n-n\theta}{\sqrt{n\theta(1-\theta)}}\right|\geq\frac{a\sqrt{n}}{\sqrt{\theta(1-\theta)}}\right\}$$

$$\leq\mathbb{P}\left\{\left|\frac{S_n-n\theta}{\sqrt{n\theta(1-\theta)}}\right|\geq 2\right\}\approx\frac{1}{\sqrt{2\pi}}\int_{|x|\geq 2}e^{-\frac{x^2}{2}}dx<0,05.$$

4.4 Estimatori bayesieni

Media empirică $\overline{X}$ este o estimare judicioasă a parametrului θ în absenţa oricăror informaţii suplimentare. Dar problema se modifică în cazul unor informaţii apriorice despre θ. Dacă aceste informaţii există, mai precis, dacă se cunoaşte o măsură de probabilitate pe $[0,1]$ reprezentând distribuţia valorilor posibile ale lui θ, atunci se va cerceta estimatorul $g(S_n)$ a lui θ care minimizează riscul pătratic mediu, ţinând cont de această distribuţie.

Fie D o mulţime numărabilă din $[0,1]$, iar $\rho:D\to[0,1]$ o măsură de probabilitate pe D, cu $\sum_{x\in D}\rho(x)=1$. Pentru o funcţie g definită pe $\{0,1,\cdots,n\}$ cu valori în $[0,1]$, riscul pătratic mediu al estimatorului $g(S_n)$ este

$$\sum_{\theta\in D}\rho(\theta)\mathbb{E}_\theta\left[g(S_n)-\theta\right]^2=\sum_{\theta\in D}\rho(\theta)\sum_{k=0}^{n}(g(k)-\theta)^2C_n^k\theta^k(1-\theta)^{n-k}$$

$$=\sum_{k=0}^{n}C_n^k\left(\sum_{\theta\in D}\rho(\theta)(g(k)-\theta)^2\theta^k(1-\theta)^{n-k}\right).$$

Pentru $k\in\{0,1,\cdots,n\}$, considerăm funcţia $\psi_k:[0,1]\to[0,\infty)$, definită prin

$$\psi_k(x)=\sum_{\theta\in D}\rho(\theta)\theta^k(1-\theta)^{n-k}(x-\theta)^2,\ x\in[0,1].$$

Avem

$$\psi_k'(x)=2\sum_{\theta\in D}\rho(\theta)\theta^k(1-\theta)^{n-k}(x-\theta),\ x\in[0,1].$$

Studiind semnul derivatei, deducem că ψ_k atinge minimul în punctul

$$x_k=\frac{\sum_{\theta\in D}\rho(\theta)\theta^{k+1}(1-\theta)^{n-k}}{\sum_{\theta\in D}\rho(\theta)\theta^k(1-\theta)^{n-k}}\in[0,1].$$

Vom defini deci *estimatorul bayesian* a lui θ de risc pătratic mediu minim:

$$g(S_n) = \frac{\sum_{\theta \in D} \rho(\theta)\theta^{S_n+1}(1-\theta)^{n-S_n}}{\sum_{\theta \in D} \rho(\theta)\theta^{S_n}(1-\theta)^{n-S_n}}.$$

Similar, dacă se cunoaşte "a priori" densitatea de probabilitate f a lui θ pe intervalul $[0,1]$, atunci *estimatorul bayesian* a lui θ, dependent de S_n, având *riscul pătratic minim*, va fi

$$g(S_n) = \frac{\int_0^1 \theta^{S_n+1}(1-\theta)^{n-S_n} f(\theta) d\theta}{\int_0^1 \theta^{S_n}(1-\theta)^{n-S_n} f(\theta) d\theta}.$$

4.5 Rezumat

Statistica matematica utilizează rezultatele oferite de teoria probabilităţilor pentru a generaliza rezultatele unei experienţ la o clasă de experienţe similare. Acest tip de extindere se datorează unui raţionament de tip *inductiv* (obţinerea unor informaţii generale din analiza unor informaţii particulare). Statistica analizează date concrete, locale, obţinute experimental, pentru a prognoza date cu caracter general, a descoperi legile care guvernează fenomemul studiat.
Noţiuni specifice:

- ***populaţia statistică*** reprezintă mulţimea tuturor elementelor avute în vedere în cadrul unui ***studiu statistic***;
- *sondajul* reprezintă examinarea (chestionarea) unei părţi a populaţiei statistice, numită ***eşantion***;
- numărul elementelor unui eşantion (obţinut prin sondaj şi examinat) se numeşte ***volumul eşantionului***;
- scopul unui studiu statistic este ***estimarea*** unei proprietăţii/particularităţi a populaţiei statistice. Funcţia matematică care realizează estimarea se numeşte ***estimator***;
- o ***statistică*** (sau un ***estimator*** se defineţe ca o funcţie de variabilelor aleatoare observabile (cele ale n-eşantionului) şi nu depinde de parametrii necunoscuţi.

Statistica eşantioanelor Bernoulli.
Modelare: fiecare element al populaţiei statistice se reprezintă printr-o variabilă aleatoare Bernoulli X, cu distribuţia

$$X : \begin{pmatrix} 0 & 1 \\ 1-\theta & \theta \end{pmatrix},$$

Pentru evaluarea apriorică a prametrului necunoscut θ se propune estimatorul ***media empirică***, definit prin:

$$\overline{X} = \frac{S_n}{n}, \text{ unde } S_n = X_1 + X_2 + \cdots + X_n.$$

Media empirică exprimă frecvenţa de apariţie a valorii 1, deci oferă o evaluare a parametrului θ. Proprietăţile principale ale mediei empirice $\overline{X}$, ca estimator a lui θ.
1. *$\overline{X}$ este un estimator fără abatere a lui θ: $\mathbb{E}_\theta(\overline{X}) = \theta, \ \forall \ \theta \in (0,1)$.*
2. *$\overline{X}$ este unicul estimator fără abatere a lui θ, care depinde de S_n.*
3. *$\overline{X}$ este estimatorul lui θ de maximă verosimilitate*
4. *Când talia eşantionului tinde la infinit, $\overline{X}$ converge la θ cu o rată exponenţială.*
5. *Riscul pătratic mediu al estimatorului $\overline{X}$ a lui θ converge la 0.*
6. *Abaterea lui $\overline{X}$ faţă de θ poate fi evaluată prin Teorema limită centrală.*
Definim *estimatorul bayesian* a lui θ, de risc pătratic mediu minim:

$$g(S_n) = \frac{\sum_{\theta \in D} \rho(\theta)\theta^{S_n+1}(1-\theta)^{n-S_n}}{\sum_{\theta \in D} \rho(\theta)\theta^{S_n}(1-\theta)^{n-S_n}}.$$

Dacă θ admite densitatea de probabilitate f pe $[0,1]$, atunci *estimatorul bayesian* a lui θ este

$$g(S_n) = \frac{\int_0^1 \theta^{S_n+1}(1-\theta)^{n-S_n} f(\theta)d\theta}{\int_0^1 \theta^{S_n}(1-\theta)^{n-S_n} f(\theta)d\theta}.$$

4.6 Noţiuni de statistică descriptivă

4.6.1 Concepte de bază

- **Statistica**, ramură a matematicii, studiază metodele de înregistrare, descriere şi analiză a datelor experimentale referitoare la un anumit fenomen în scopul formulării unor previziuni.

- **Populaţia statistică** este mulţimea (în sens matematic) care face obiectul unui anumit **studiu statistic**.

- **Unitate statistică** (individ) este un element component al unei populaţii statistice.

- **Caracteristică statistică** este una din însuşirile unităţii statistice.
  O caracteristică cantitativă este *măsurabilă*, iar una calitativă este *atributivă*.
  Un studiu statistic a vizează una sau mai multe caracteristici statistice ale indivizilor unei populaţii statistice.

- **Eşantionul** este o submulţime a populaţiei statistice, obţinută printr-o **selecţie**, pentru care se realizează evaluarea caracteristicii statistice. Pentru unele studii statistice, eşantionul reprezintă întreaga populaţie statistică. Numărul de elemente al eşantionului se numeşte **volumul** (efectivul) eşantionului.

- **Sondajul** este acţiunea de prevalare şi înregistrare a valorilor uneia sau mai multor caracteristici statistice ale indivizilor unui eşantion fixat.

- **Tabelul statistic** conţine înregistrarea valorilor numerice (datelor experimentale), numite **valori observate** obţinute prin măsurarea caracteristicii tuturor unităţilor statistice ale eşantionului selecţionat.

- **Frecvenţa relativă** a unei valori a caracteristicii este raportul dintre numărul indivizilor eşantionului pentru care s-a înregistrat valoarea respectivă şi volumul eşantionului.

- **Distribuţia statistică empirică** reprezintă reprezintă repartiţia valorilor observate ale caracteristicii pentru eşantionul selecţionat (înregistrate în tabelul statistic). Din punct de vedere matematic, distribuţia statistică (empirică) este funcţia care realizează corespondenţa între numerele reale, obţinute ca valori observate ale caracteristicii, şi **frecvenţa relativă** a înregistrării lor.

- **Distribuţiile (repartiţiile) statistice teoretice** reprezintă modele matematice idealizate ale distribuţiilor empirice. Concret, o distribuţie statistică teoretică este descrisă de o funcţie $f : \mathbb{R} \to [0, 1]$ cu proprietatea $\int_{-\infty}^{\infty} f(x)\,dx = 1$, numită **densitate de probabilitate**. Pentru anumite fenomene analizate statistic, distribuţia empirică **tinde** la o anumită distribuţie teoretică, dacă (ne imaginăm că) volumul eşantionului **tinde la infinit**.

- **Teoria probabilităţilor** operează cu noţiunile de *spaţiu măsurabil, probabilitate* şi *variabilă aleatoare.* O variabilă aleatoare este o funcţie X, definită pe un spaţiu măsurabil cu probabilitate, cu valori reale, având proprietatea că pentru orice număr real x se poate evalua probabilitatea $F(x)$ ca X să ia valori în intervalul $(-\infty, x]$, şi anume:

$$F(x) = P\{X \leq x\} \in [0, 1].$$

Spunem că variabila aleatoare X are densitatea de probabilitate f dacă:

$$F(x) = \int_{-\infty}^{x} f(t)\, dt, \quad \text{pentru orice număr real } x$$

Media unei variabile aleatoare X cu densitatea de probabilitate f se calculează prin:

$$M(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

Dispersia unei variabile aleatoare X cu densitatea de probabilitate f se calculează prin:

$$\sigma^2(X) = \int_{-\infty}^{\infty} (x - m)^2 f(x)\, dx, \quad \text{unde } m = M(X)$$

Un caz particular îl constituie *variabilele aleatoare finite.* O astfel de variabilă aleatoare X are un număr finit de valori reale $x_1 < x_2 < \cdots < x_k$ cu probabilităţile $p_1, p_2, \cdots, p_k \in [0, 1]$. Avem deci:

$$p_i = P\{X = x_i\}, \quad i = 1, 2, \cdots, k; \qquad p_1 + p_2 + \cdots + p_k = 1.$$

În acest caz, media şi respectiv dispersia variabilei aleatoare X se calculează după următoarele formule:

$$m = M(X) = x_1 p_1 + x_2 p_2 + \cdots + x_k p_k = \sum_{i=1}^{k} x_i p_i$$

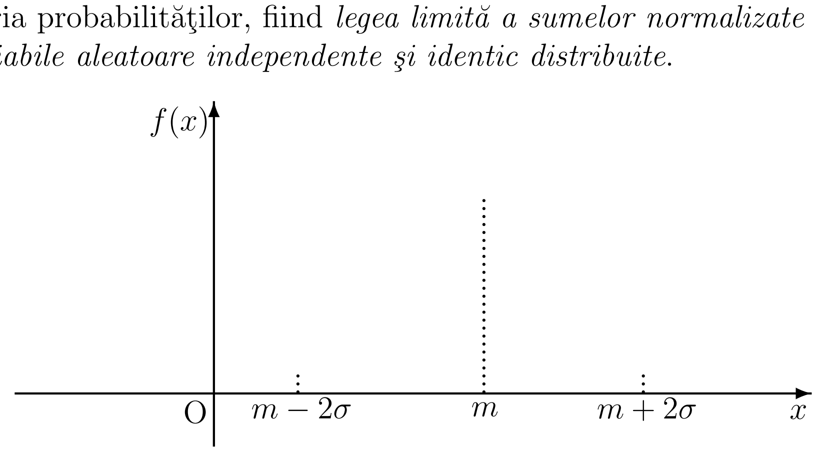
$$\sigma^2(X) = (x_1 - m)^2 p_1 + (x_2 - m)^2 p_2 + \cdots + (x_k - m)^2 p_k = \sum_{i=1}^{k} (x_i - m)^2 p_i.$$

- **Exemple de distribuţii (repartiţii, legi) teoretice clasice**:

1. **Legea normală** de **medie** m şi **dispersie** σ^2 are densitatea de probabilitate:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad x \in R$$

Graficul lui f are forma unui clopot, cunoscut sub denumirea de *clopotul lui Gauss*. Legea normală (sau legea lui *Gauss-Laplace*) joacă un rol central în teoria probabilităţilor, fiind *legea limită a sumelor normalizate ale şirurilor de variabile aleatoare independente şi identic distribuite.*



Valorile unei

variabile aleatoare X cu distribuţie normală de medie m şi dispersie σ^2 se ”concentrează” cu o probabilitate de 95% în intervalul $[m-2\sigma, m+2\sigma]$, centrat în valoarea medie m. Mai precis avem:

$$P\{m-2\sigma \leq X \leq m+2\sigma\} = 0,9544\cdots$$

unde $\sigma = \sqrt{\sigma^2}$ se numeşte **abaterea standard** a variabilei aleatoare X.

2. **Repartiţia** χ^2 cu m **grade de libertate** are densitatea de probabilitate:

$$f(x) = \frac{1}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)} x^{\frac{m}{2}-1}e^{-\frac{x}{2}}, \quad x > 0$$

unde:

$$\Gamma\left(\frac{m}{2}\right) = \int_0^{\infty} x^{\frac{m}{2}-1}e^{-x}\,dx.$$

Dacă o variabilă aleatoare X are repartiţia χ^2 cu m grade de libertate, atunci media şi respectiv dispersia variabilei aleatoare sunt următoarele:

$$M(X) = m$$

$$\sigma^2(X) = 2m$$

Repartiţia χ^2 are un rol important în statistica matematică în deducerea repartiţei statistice a unor *estimatori* şi în verificarea *ipotezelor statistice.*

4.6.2 Tipuri de distribuţii statistice empirice. Tabele statistice. Reprezentări grafice.

In cadrul unui studiu statistic, datele experimentale obţinute prin măsurarea valorică a unei caracteristici a unui eşantion considerat sunt înregistrate într-un tabel statistic. Acesta evidenţiază o anumită distribuţie statistică. După natura repartiţiei valorilor înregistrate, distingem două tipuri de distribuţii statistice empirice (experimentale).

1) **Tipul finit (sau discret)** de distribuţie corespunde cazului în care, ipotetic, caracteristica masurată poate lua un număr finit (sau numărabil) de valori. In continuare vom trata numai cazul finit.

Considerăm un eşantion fixat de volum n.. Aceasta înseamnă că numărul de unităţi statistice pentru care se determină şi se înregistrează valoarea caracteristicii este egal cu n. Să presupunem că valorile înregistrate ale caracteristicii sunt numerele reale $x_1 < x_2 < \cdots < x_k$ unde k este un număr natural cuprins între 1 şi n.

Fiecarei valori realizate (înregistrate) a caracteristicii îi determinăm ”frecvenţa” de apariţie.

- **Frecvenţa absolută a valorii** x_i este numărul n_i de subiecţi pentru care valoarea caracteristicii este x_i . Din definiţie rezultă $0 \leq n_i \leq n$ şi $\sum_{i=1}^k n_i = n$, unde $\sum_{i=1}^k n_i$ este notaţia simbolică a sumei $n_1 + n_2 + \cdots + n_k$.

- **Frecvenţa relativă a valorii** x_i este numărul $f_i = \frac{n_i}{n}$. Constatăm că au loc relaţiile: $f_i \in [0,1]$ şi $\sum_{i=1}^k f_i = 1$.

- **Frecvenţa procentuală a valorii** x_i este numărul $p_i = 100 \cdot \frac{n_i}{n} = 100 \cdot f_i$. Avem $p_i \in [0, 100]$ şi $\sum_{i=1}^{k} p_i = 100$. Pentru a evidenţia că p_i reprezină o mărime procentuală, se foloseşte curent notaţia $p_i\,\%$.

Frecvenţele determinate ale valorilor, indicând distribuţia statistică empirică, pot fi centralizate în *tabelul frecvenţelor* sau pot fi reprezentate grafic prin *diagrama de structură, histograma* şi respectiv *poligonul frecvenţelor.*

- **Diagrama de structură** ilustrează distribuţia statistică empirică prin partiţia unui disc în sectoare de cerc cu ariile proporţionale cu frecvenţele relative ale valorilor. Sectorul reprezentând valoarea x_i are unghiul de la centru de măsură $\alpha_i = f_i \cdot 360^o$.

- **Poligonul frecvenţelor** se obţine, într-un sistem de axe de coodonate, unind succesiv prin segmente punctele de coordonate (x_1, f_1), $(x_2, f_2), \cdots, (x_k, f_k)$.

- **Histograma** ilustrează, într-un sistem de axe de coodonate, frecvenţele relative (sau procentuale) ale valorilor x_i prin dreptunghiuri de înălţime f_i (respectiv p_i) dispuse cu baza pe axa absciselor, în dreptul absciselor x_i.

Exemplul 1 *La o probă sportivă fiecare participant a efectuat 6 aruncări la coşul de baschet, realizând între 0 şi 6 aruncări reuşite. Rezultatele obţinute de cei 40 de participanţi sunt prezentate sintetic în tabelul următor:*

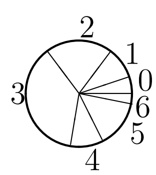
numărul de aruncări reuşite	0	1	2	3	4	5	6	
numărul de participanţi	2	4	8	15	4	6	1	total 40

Volumul eaşntionului considerat este $n = 40$ iar numărul de valori ale caracteristicii este $k = 7$. Prezentăm tabelul frecvenţelor, diagrama de structură, histograma şi respectiv poligonul frecvenţelor asociate studiului statistic considerat.

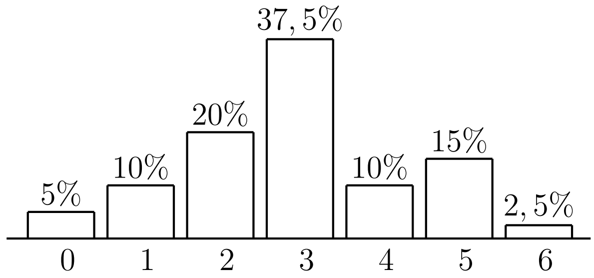
A) Tabelul frecvenţelor

valorile x_i	frecvenţele		
	absolute (n_i)	relative (f_i)	procentuale ($p_i\,\%$)
0	2	0,050	5,0 %
1	4	0,100	10,0 %
2	8	0,200	20,0 %
3	15	0,375	37,5 %
4	4	0,100	10,0 %
5	6	0,150	15,0 %
6	1	0,025	2,5 %
	$\sum_{i=1}^{7} n_i = 40$	$\sum_{i=1}^{7} f_i = 1$	$\sum_{i=1}^{7} p_i = 100$

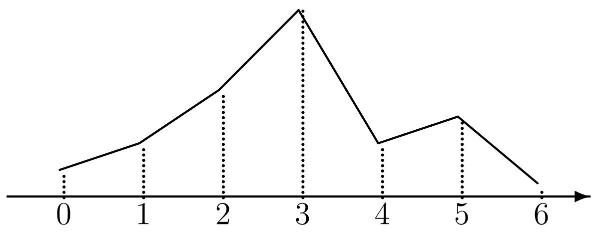
B) Diagrama de structură



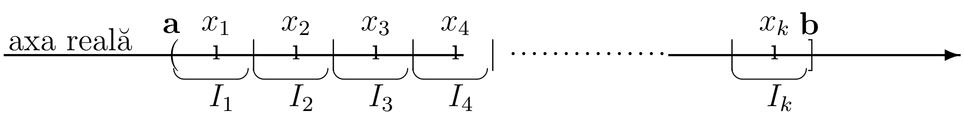
C) Histograma



D) Poligonul frecvenţelor



2) **Tipul continuu** de distribuţie corespunde cazului când valorile caracteristicii sunt numere reale situate într-un anumit interval [a,b]. Pentru a grupa valorile caracteristicii, se fixează, ţinând cont de specificul studiului statistic efectuat, un număr întreg $k > 1$ şi se împarte intervalul $(a, b]$ în k subintervale de aceeaşi lungime $l = \frac{b-a}{k}$. Notăm, în ordine, aceste intervale $I_1, I_2, \cdots, I_k$. Astfel, $I_i = (a + \frac{(i-1)(b-a)}{k}, a + \frac{i(b-a)}{k}]$, $i = 1, 2, \cdots, k$. Fie un eşantion fixat de volum n cu $n > k$. Fiecărui interval I_i i se asociază "frecvenţa de apariţie"; vom nota cu x_i valoarea "centrală" a intevalului, adică $x_i = a + \frac{(2i-1)(b-a)}{2k}$, obţinută făcând media aritmetică a capetelor intervalului I_i (a se vedea figura următoare).



- **Frecvenţa absolută a intervalului** I_i este numărul n_i de subiecţi pentru care valoarea caracteristicii aparţine intervalului I_i.
- **Frecvenţa relativă a intervalului** I_i este numărul $f_i = \frac{n_i}{n}$.
- **Frecvenţa procentuală a intervalului** I_i este numărul $p_i = 100 \cdot f_i$.

Frecvenţele determinate ale valorilor, indicând distribuţia statistică empirică, pot fi centralizate în *tabelul frecvenţelor* sau pot fi reprezentate grafic prin *diagrama de structură, histograma* şi respectiv *poligonul frecvenţelor* care nu diferă de cele descrise pentru distribuţii finite.

Exemplul 2 *La o probă contracronometru de ciclism au participat 148 de concurenţi. După în registrarea timpilor de parcurs, s-au grupat rezultatele în* 6 *intervale de câte* 30" *şi s-au calculat frecvenţele intervalelor, obţinându-se următorul tabel:*

Tabelul frecvenţelor

intervalele I_i	valorile centrale x_i	frecvenţele absolute (n_i)	frecvenţele relative (f_i)	frecvenţele procentuale ($p_i\,\%$)
$(48', 48'30"]$	$48'15"$	7	0,047	4,7 %
$(48'30", 49']$	$48'45"$	28	0,189	18,9 %
$(49', 49'30"]$	$49'15"$	53	0,358	35,8 %
$(49'30", 50']$	$49'45"$	42	0,283	28,3 %
$(50', 50'30"]$	$50'15"$	10	0,067	6,7 %
$(50'30", 51']$	$50'45"$	8	0,054	5,4 %
		$\sum_{i=1}^{6} n_i = 148$	$\sum_{i=1}^{6} f_i = 1$	$\sum_{i=1}^{6} p_i = 100$

4.6.3 Indicatori statistici

Pentru o distribuţie statistică empirică, consemnată în tabelul frecvenţelor, determinăm o serie de **indicatori statistici** în scopul obţinerii unor informaţii cu caracter general privind repartizarea valorilor carateristicii studiate. Cei mai importanţi indicatori statistici sunt: media, mediana şi respectiv dispersia distribuţiei. Media şi dispersia indică, fiecare în parte, valorea "centrală" sau valoarea "de mijloc" a distribuţiei respective, într-o o accepţiune specifică fiecărei noţiuni. Dispersia si respectiv abaterea standard indică gradul de "împrăştiere" a valorilor faţă de valoarea "centrală" a distribuţiei. Coeficientul de variaţie, calculat prin raportarea abaterii standard la medie, oferă posibilitatea stabilirii gradului de omogenitate al eşantionului.

Considerăm o distribuţie statistică empirică, asociată unei caracteristici (variabile empirice) X, care ia valorile $x_1 < x_2 < \cdots < x_k$ cu frecvenţele absolute $n_1, n_2, \cdots, n_k$ şi respectiv frecvenţele relative $f_1, f_2, \cdots, f_k$.

- **Media** $\overline{x} = M(X)$ a variabilei X se calculează prin:

$$\overline{x} = x_1 f_1 + x_2 f_2 + \cdots + x_k f_k,$$

  sau, utilizând scrierea prescurtată, $\overline{x} = M(X) = \sum_{i=1}^{k} x_i f_i$.

- **Mediana** $Med = Med(X)$ a variabilei X, corespunzând unui eşantion de volum n, se defineşte astfel: se ordonează crescător valorile caracteristicii tuturor subiecţilor eşantionului, obţinându-se şirul de valori

$$v_1 \leq v_2 \leq \cdots \leq v_n.$$

  Distingem două situaţii:

  1. dacă $n = 2p + 1$, cu p - număr natural (deci n este un număr natural impar), atunci $Med = v_{p+1}$;
  2. dacă $n = 2p$, cu p - număr natural (deci n este un număr natural par), atunci $Med = \frac{v_p + v_{p+1}}{2}$.

Observăm că din definiţia medianei reiese că aceasta este *valoarea care "împarte" eşantionul în în două părţi de volume egale*

Mediana se mai poate determina şi utilizând frecvenţele absolute $n_1, n_2, \cdots, n_k$ ale valorilor $x_1 < x_2 < \cdots < x_k$, după cum urmează:

$$Med = \begin{cases} x_i, & \text{dacă} \quad n_1 + n_2 + \cdots + n_{i-1} < \frac{n}{2} < n_1 + n_2 + \cdots + n_{i-1} + n_i \\ \frac{x_i + x_{i+1}}{2}, & \text{dacă} \quad \frac{n}{2} = n_1 + n_2 + \cdots + n_{i-1} + n_i \end{cases}$$

- **Amplitudinea** $W = W(X)$ a variabilei X se defineşte prin:

$$W = x_{max} - x_{min}$$

Astfel, dacă X are o distribuţie de *tip finit* atunci $W = x_k - x_1$, iar dacă X are o distribuţie de *tip continuu* cu valorileîn intervalul $(a, b]$, atunci $W(X) = b - a$.

- **Dispersia** $\sigma^2 = \sigma^2(X)$ a variabilei X se defineşte prin:

$$\sigma^2 = \sum_{i=1}^{k} (x_k - \overline{x})^2 f_i,$$

unde $\overline{x} = M(X)$ este media variabilei X.

- **Abaterea standard** $\sigma = \sigma(X)$ a variabilei X se defineşte prin:

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

- **Coeficientul de variaţie** $C_v = C_v(X)$ al variabilei X, se defineşte, numai în cazul când *toate valorile x_i ale distribuţiei X sunt pozitive*, prin:

$$C_v = \frac{\sigma}{\overline{x}}$$

- **Expresia procentuală coeficientului de variaţie** $C_{v\%} = C_{v\%}(X)$ este:

$$C_{v\%} = 100 \cdot C_v \ \%$$

Tabelul următor prezintă interpretarea curentă a **omogenităţii eşantionului în raport cu variabila X** după valorile coeficientului de variaţie exprimat procentual:

Coeficientul de variaţie $C_{v\%}$	Omogenitatea
0-10%	omogenitate mare
10-20%	omogenitate medie
20-35%	omogenitate mică
peste 35%	eşantion neomogen

Să urmărim calculul indicatorilor statistici definiţi anterior pentru distribuţiile din exemplele anterioare. Vom nota cu X variabila empirică cu distribuţia prezentată în Exemplul 1 şi respectiv cu Y variabila empirică cu distribuţia descrisă în Exemplul 2.

4.6.4 Calculul indicatorilor statistici ai distribuţiei variabilei X (Exemplul 1).

Media:
$\overline{x} = 0 \cdot 0,05 + 1 \cdot 0,1 + 2 \cdot 0,2 + 3 \cdot 0,375 + 4 \cdot 0,1 + 5 \cdot 0,15 + 6 \cdot 0,025 = 2,925$
Rezultă că numărul mediu de reuşite este:

$$\overline{x} = M(X) \cong 3$$

Mediana:
Avem $n : 2 = 20$ şi $n_1+n_2+n_3 = 14 < 20 < 29 = n_1+n_2+n_3+n_4$, deci $Med(X) = x_4$, sau:

$$Med(X) = 3.$$

Amplitudinea:

$$W(X) = 6 - 0 = 6$$

Dispersia:
$\sigma^2(X) = (0-2,925)^2 \cdot 0,05 + (1-2,925)^2 \cdot 0,1 + (2-2,925)^2 \cdot 0,2 + (3-2,925)^2 \cdot 0,375 + (4-2,925)^2 \cdot 0,1 + (5-2,925)^2 \cdot 0,15 + (6-2,925)^2 \cdot 0,025 = 1,969 \cdots$ Dispersia valorilor (abaterea medie pătratică) este deci:

$$\sigma^2(X) \cong 1,97$$

Abaterea standard:

$$\sigma(X) = \sqrt{1,969} \cong 1,403$$

Coeficientul de variaţie:

$$C_v(X) = \frac{1,403}{2,925} \cong 0,4797$$

Expresia procentuală a coeficientul de variaţie:

$$C_{v\%}(X) \cong 48\,\%$$

Interpretare: *eşantion neomogen.*

4.6.5 Calculul indicatorilor statistici ai distribuţiei variabilei Y (Exemplul 2).

Media
$\overline{x} = 48,25 \cdot 0,047 + 48,75 \cdot 0,189 + 49,25 \cdot 0,358 + 49,75 \cdot 0,283 + 50,25 \cdot 0,067 + 50,75 \cdot 0,054 = 49,29 \cdots$ (minute)

Rezultă că timpul mediu înregistrat al probei este:

$$M(Y) = \overline{x} \cong 49'18"$$

Mediana:
Avem: $n : 2 = 74$ şi $n_1 + n_2 = 35 < 74 < 88 = n_1 + n_2 + n_3$, deci $Med(Y) = x_3$, adică:

$$Med(Y) = 49'15"$$

Amplitudinea:

$$W(Y) = 51 - 48 = 3 \text{ (minute)}$$

Dispersia:
$\sigma^2(Y) = (48,25 - 49,29)^2 \cdot 0.047 + (48,75 - 49,29)^2 \cdot 0,189 + (49,25 - 49,29)^2 \cdot 0,358 + (49,75 - 49,29)^2 \cdot 0,283 + (50,25 - 49,29)^2 \cdot 0,067 + (50,75 - 49,29)^2 \cdot 0,054 = 0,34325 \cdots$

$$\sigma^2(Y) \cong 0,343$$

Abaterea standard:

$$\sigma(Y) = \sqrt{0,343} \cong 0,586$$

Coeficientul de variaţie:

$$C_v(Y) = \frac{0,586}{49,29} \cong 0,012$$

Exprimarea procentuală a coeficientului de variaţie:

$$C_{v\%}(Y) \cong 1,2\ \%$$

Interpretarea: *eşantion foarte omogen.*

4.6.6 Compararea a două distribuţii empirice

Să presupunem că pentru un eşantion fixat de volum n se sondează în paralel două caracteristici (variabile empirice) X şi Y şi se pune problema existenţei unei conexiuni între distribuţiile acestora. Pentru obţinerea unei concluzii se recurge la calculul unor indici de corelaţie asociaţi perechii de variabile X şi Y.

4.6.7 Corelaţia şi coeficientul de corelaţie

Fie $\overline{x} = M(X)$, $\overline{y} = M(Y)$ mediile variabilelor X şi Y, iar σ_X, σ_Y abaterile lor standard.

Presupunem că X ia valorile $x_1 < x_2 < \cdots < x_k$ iar Y ia valorile $y_1 < y_2 < \cdots y_l$. Notăm $n_{i,j}$ numărul unităţilor statistice ale eşantionului pentru care caracteristica X ia valoarea x_i iar caracteristica Y ia valoarea y_j, unde $i = 1, 2, \cdots, k$ şi $j = 1, 2, \cdots, l$. Fie $f_{i,j} = \frac{n_{i,j}}{n}$ frecvenţa relativă a perechii de valori (x_i, y_j).

Corelaţia $c_{X,Y}$ **a variabilelor** X **şi** Y se defineţe prin:

$$c_{X,Y} = M((X - \overline{x})(Y - \overline{y})) = \sum_{\substack{i=\overline{1,k} \\ j=\overline{1,l}}} (x_i - \overline{x})(y_j - \overline{y}) f_{i,j}$$

Coeficientul de corelaţie $r = r_{X,Y}$ **al variabilelor** X **şi** Y **al variabilelor** X **şi** Y se defineşte prin:

$$r = \frac{c_{X,Y}}{\sigma_X \cdot \sigma_Y}.$$

Pe baza unei inegalităţi matematice fundamentale (*inegalitatea lui Cauchy*) se demonstrează că, oricare ar fi distribuţiile variabilelor X şi Y, avem:

$$-1 \leq r \leq 1.$$

Interpretarea valorii coeficientului de corelaţie r este următoarea:

1. dacă r este pozitiv, atunci variabilele X şi Y sunt *corelate*, adică tendinţa este ca atunci când variabila X ia o "valoare mare" şi variabila Y să ia tot o "valoare mare". Variabilele X şi Y sunt cu atât mai puternic corelate cu cât coeficientul lor de corelaţie este mai apropiat de valoarea 1.

2. dacă r este negativ, atunci variabilele X şi Y sunt *invers corelate*, adică tendinţa este ca atunci când variabila X ia o "valoare mare" variabila Y să ia o "valoare mică".

3. dacă r este apropiat de valoarea 0, atunci variabilele X şi Y sunt *independente* sau cu o *dependenţă slabă.*

4.6.8 Metoda coreraţiei rangurilor

Această metodă, datorată lui Spearman, se utilizează atunci când volumul eşantinului este mic (de regulă $n < 30$). Fiecărui individ $i \in \{1, 2, \cdots n\}$ al eşantionului i se asociază o pereche de **ranguri** $(r_i(X), r_i(Y))$, cu $r_i(X), r_i(Y) \in [1, n]$. Aceste ranguri se determină în urma realizării a câte unui "clasament" crescător al valorilor înregistrate pentru fiecare din caracteristicile (variabilele) X şi Y. Rangurile $r_i(X)$ şi respectiv $r_i(Y)$ exprimă "poziţia" (eventual intermediară) a individului i în fiecare din cele două "clasamente" de valori crescătoare.

Notăm $d_i = r_i(X) - r_i(Y), \quad i = 1, 2, \cdots, n.$

Coeficientul de corelaţie al rangurilor $\rho = \rho_{X,Y}$ **asociat variabilelor** X **şi** Y se defineşte prin:

$$\rho = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2 - 1)}$$

Se poate verifica prin calcul că, indiferent de distribuţia rangurilor, avem:

$$0 \leq \rho \leq 1$$

Dacă ρ este apropiat de 1, atunci variabilele X şi Y sunt *corelate* iar dacă ρ este apropiat de 0, atunci variabilele X şi Y sunt *invers corelate.*

4.7 Test

1. Definiţi media empirică şi descrieţi proprietăţile acestui estimator în cazul eşantioanelor Bernoulli.

2. Determinaţi estimatorul bayesian al unui eşantion Bernoulli asociat unei distribuţii $Beta(\alpha, \beta)$ a parametrului de estimat θ.

3. Se observă două variabile aleatoare Bernoulli indepedente X_1 şi X_2, de parametru $\theta \in (0, 1)$. Să se arate că nu se poate găsi un estimator a fără abatere a lui $\frac{\theta}{1-\theta}$ care să depindă de X_1 şi X_2.

4. Cu ajutorul unui n-eşantion Bernoulli de parametru θ se caută estimarea dispersiei $\theta(1-\theta)$. Să se arate că estimatorul $T = \overline{X}(1-\overline{X})$ nu este fără abatere, dar există un multiplu al său având această proprietate.

Referinţe bibliografice

[1] G. Ciucu, C. Tudor, *Probabilităţi şi procese stochastice*, Ed. Academiei Române, 1979 (vol 1-2).

[2] I. Cuculescu, *Teoria probabilităţilor*, Ed. All, 1998.

[3] D. Dacunha-Castelle, M. Duflo, *Probabilités et statistique*, Masson, 1994 (vol 1).

[4] W. Feller, *Introduction to probability theory and its applications*, Wiley, New York, 1968.

[5] P.G. Hoel, *Introduction to mathematical statistics*, Wiley, 1984.

[6] M. Iosifescu, Gh. Mihoc, R. Theodorescu, *Teoria probabilităţilor şi statistică matematică*, Ed. Tehnică, 1966.

[7] O. Onicescu, Gh. Mihoc, C. Ionescu-Tulcea, *Calculul probabilităţilor şi aplicaţii*, Ed. Academiei Române, 1965.