PROBABILITĂŢI ŞI STATISTICĂ

Eugen Păltănea

Universitatea Transilvania din Braşov
2021

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.

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

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.*

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

3. Probabilitatea de realizare simultană a două evenimente este egală cu produsul dintre 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 *spaţiu (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

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

Proprietăţile generale ale corpurilor deduse din definiţia corpului de mulţimi sunt prezentate în propoziţia următoare.

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 $\Omega = A \cup A^c \in \mathcal{C}$ (cf. C_1 şi C_2) şi $\emptyset = A \cap A^c \in \mathcal{C}$ (cf. C_1 şi C_4). C_6 se obţine din relaţia: $A \setminus B = A \cap B^c$ (cf. C_1 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, $\Omega \in \mathcal{C}_i,\ \forall i \in I$ (cf. C_5), deci $\Omega \in \mathcal{C}$. Rezultă $\mathcal{C} \neq \emptyset$. 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 $\mathcal{M}$.

Constatăm că $\mathcal{C}(\mathcal{M})$ 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 $\mathcal{M}_3 = \{\cup_{A \in G} A\ :\ G \subset \mathcal{M}_2\}$. Se verifică faptul că $\mathcal{C}(\mathcal{M}) = \mathcal{M}_3$, deci corpul $\mathcal{M}$ 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 $\mathcal{F} \subset \mathcal{P}(\Omega)$ 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 $A_n = B,\ \forall\ n \geq 2$. Conform σ_2, avem $A \cup B = \cup_{n=1}^{\infty} A_n \in \mathcal{F}$. Deducem că σ-corpul $\mathcal{F}$ este un corp. Ca urmare, proprietăţile $C_3 - C_6$ rămân valabile pentru un σ-corp $\mathcal{F}$. Alte proprietăţi specifice unui σ-corp 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 σ_3. $\square$

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 $\emptyset \neq \mathcal{M} \subset \mathcal{P}(\Omega)$. Intersecţia σ- 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* $(\Omega, \mathcal{T})$ *un spaţiu topologic, unde* $\mathcal{T}$ *reprezintă topologia lui* Ω *(familia mulţimilor deschise). Atunci* $\sigma -$ *corpul generat de mulţimea* $\mathcal{T}$ *se numeşte borelianul spaţiului topologic* $(\Omega, \mathcal{T})$ *şi se notează (atunci când nu este pericol de confuzie) prin* $\mathcal{B}_\Omega$ *.*

Topologia $\mathcal{T}_\mathbb{R}$ a mulţimii numerelor reale $\mathbb{R}$ este **familia reuniunilor cel mult numărabile de intervale deschise disjuncte**. Are loc următorul rezultat fundamental.

Teorema 1.3.1. *Borelianul* $\mathcal{B}_\mathbb{R}$ *al spaţiului topologic* $(\mathbb{R}, \mathcal{T}_\mathbb{R})$ *conţine toate tipurile de intervale şi reuniunile cel mult numărabile ale acestora. În plus,* $\mathcal{B}_\mathbb{R}$ *este* σ *- corpul generat de mulţimea tuturor intervalelor reale de un anumit tip.*

Demonstrăm teorema anterioară într-un caz particular.

Propoziţia 1.3.3. *Considerăm mulţimea de intervale reale* $\mathcal{M} = \{(-\infty, a],\ a \in \mathbb{R}\}$.
Are loc relaţia:
$$\mathcal{B}(\mathcal{M}) = \mathcal{B}_\mathbb{R}.$$

Demonstraţie.

1. Pentru oricare $a \in \mathbb{R}$, avem $(-\infty, a] = \bigcap_{n=1}^{\infty} \left(-\infty, a + \frac{1}{n}\right) \in \mathcal{B}_\mathbb{R}$, de unde $\mathcal{M} \subset \mathcal{B}_\mathbb{R}$.

Rezultă $\mathcal{B}(\mathcal{M}) \subset \mathcal{B}_\mathbb{R}$.

2. Pentru oricare $a \in \mathbb{R}$, avem $(-\infty, a) = \bigcup_{n=1}^{\infty} \left(-\infty, a - \frac{1}{n}\right] \in \mathcal{B}(\mathcal{M})$. De asemenea, $(a, \infty) = (-\infty, a]^c \in \mathcal{B}(\mathcal{M}),\ \forall\, a \in \mathbb{R}$. Atunci, pentru $a, b \in \mathbb{R}$, cu $a < b$, obţinem $(a, b) = (a, \infty) \cap (-\infty, b) \in \mathcal{B}(\mathcal{M})$. Rezultă $\mathcal{T}_\mathbb{R} \subset \mathcal{B}(\mathcal{M})$, de unde $\mathcal{B}_\mathbb{R} \subset \mathcal{B}(\mathcal{M})$.

Ca urmare, $\mathcal{B}(\mathcal{M}) = \mathcal{B}_\mathbb{R}$. □

Pentru oricare $a \in \mathbb{R}$, avem $\{a\} = (-\infty, a] \cap [a, \infty) \in \mathcal{B}_\mathbb{R}$ (cf. Teoremei 1.3.1). Rezultă că $\mathcal{B}_\mathbb{R}$ conţine mulţimile reale finite şi mulţimile reale numărabile. Deducem $\mathbb{N}, \mathbb{Z}, \mathbb{Q} \in \mathcal{B}_\mathbb{R}$. De asemenea, avem $\mathbb{R} \setminus \mathbb{Q} = \mathbb{Q}^c \in \mathcal{B}_\mathbb{R}$.

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 măsurabilă dacă*

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

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

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

Din caracterizarea borelianului $\mathcal{B}_\mathbb{R}$ (Teorema 1.3.1) şi Definiţia 1.4.1 rezultă că, pentru o funcţie reală măsurabilă $f : \Omega \to \mathbb{R}$, avem:

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

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

Următoarea teoremă oferă o condiţie necesară şi suficientă de măsurabilitate a unei funcţii reale definite pe un spaţiu măsurabil.

Teorema 1.4.1. *Fie $(\Omega, \mathcal{F})$ un spaţiu măsurabil, $\mathcal{M}$ mulţimea tuturor intervalelor de un anumit tip şi $f : \Omega \to \mathbb{R}$ o funcţie reală. Funcţia f este măsurabilă dacă şi numai dacă*

$$\{f \in A\} \in \mathcal{F}, \ \forall\, A \in \mathcal{M}.$$

Demonstraţie.
Presupunem că f este măsurabilă. Din Teorema 1.3.1 rezultă $\mathcal{M} \subset \mathcal{B}_{\mathbb{R}}$. Atunci, conform Definiţiei 1.4.1, deducem $\{f \in A\} \in \mathcal{F}, \ \forall\, A \in \mathcal{M}$.
Presupunem $\{f \in A\} = f^{-1}(A) \in \mathcal{F}, \ \forall\, A \in \mathcal{M}$. Pe baza relaţiilor

$$f^{-1}\left(A^c\right) = \left(f^{-1}(A)\right)^c, \ f^{-1}\left(\cup_{i \in I} A_i\right) = \cup_{i \in I} f^{-1}\left(A_i\right) \text{ şi } f^{-1}\left(\cap_{i \in I} A_i\right) = \cap_{i \in I} f^{-1}\left(A_i\right),$$

unde $A, \ A_i \in \mathcal{M}, \ \forall\, i \in I$, iar I este o mulţime de indici cel mult numărabilă, rezultă $\{f \in A\} \in \mathcal{F}, \ \forall\, A \in \mathcal{B}(\mathcal{M})$. Dar, conform Teoremei 1.3.1, $\mathcal{B}(\mathcal{M}) = \mathcal{B}_{\mathbb{R}}$. Astfel, f este o funcţie măsurabilă. □

Teorema de mai sus permite demonstrarea unor proprietăţi importante ale funcţiilor măsurabile.

Propoziţia 1.4.1. *Fie $(\Omega, \mathcal{F})$ un spaţiu măsurabil şi $f, g : \Omega \to \mathbb{R}$ două funcţii măsurabile. Atunci*

1. $\{f < g\} \in \mathcal{F}$;
2. $\{f \geq g\} \in \mathcal{F}$;
3. $\{f = g\} \in \mathcal{F}$.

Demonstraţie.
1. Pe baza densităţii mulţimii numărabile $\mathbb{Q}$ în mulţimea $\mathbb{R}$, caracterizării σ-corpului $\mathcal{F}$ şi proprietăţilor funcţiilor măsurabile f şi g, obţinem

$$\{f < g\} = \bigcup_{r \in \mathbb{Q}} \{f < r < g\} = \bigcup_{r \in \mathbb{Q}} \left(\{f < r\} \cap \{g > r\}\right) \in \mathcal{F}.$$

2. Conform 1., obţinem $\{f \geq g\} = \{f < g\}^c \in \mathcal{F}$. Analog, $\{f \leq g\} \in \mathcal{F}$.
3. Din 2. rezultă $\{f = g\} = \{f \geq g\} \cap \{f \leq g\} \in \mathcal{F}$. □

Conservarea măsurabilităţii prin operaţii algebrice este stabilită de propoziţia următoare.

Propoziţia 1.4.2. *Fie $(\Omega, \mathcal{F})$ un spaţiu măsurabil. Dacă $f, g : \Omega \to \mathbb{R}$ sunt funcţii măsurabile şi $a \in \mathbb{R}$ următoarele funcţii sunt măsurabile:*

a) $f + a, \ a \in \mathbb{R}$; *b)* $af, \ a \in \mathbb{R}$; *c)* $f + g$; *d)* $|f|$; *e)* f^2; *f)* fg; *g)* $min\{f, g\}$; *h)* $\max\{f, g\}$.

Demonstraţie. Vom aplica caracterizarea funcţiilor măsurabile asigurată de Teorema 1.4.1, în cazul particular $\mathcal{M} = \{(-\infty, c),\ c \in \mathbb{R}\}$.

a) Fie $a \in \mathbb{R}$. Avem $\{f + a < c\} = \{f < c - a\} \in \mathcal{F}$, $\forall c \in \mathbb{R}$, deci $f + a$ este măsurabilă.

b) Dacă $a = 0$, atunci $af = 0$, iar funcţia nulă este evident măsurabilă; dacă $a > 0$, atunci $\{af < c\} = \{f < c/a\} \in \mathcal{F}$, $\forall c \in \mathbb{R}$, deci af este măsurabilă; dacă $a < 0$, atunci $\{af < c\} = \{f > c/a\} \in \mathcal{F}$, $\forall c \in \mathbb{R}$, deci af este măsurabilă.

c) Fie $c \in \mathbb{R}$. Conform *a)* şi *b)*, funcţia $c - g$ este măsurabilă. Atunci, pe baza Propoziţiei 1.4.1, subpunctul 1, obţinem $\{f + g < c\} = \{f < c - g\} \in \mathcal{F}$. Deducem că funcţia $f + g$ este măsurabilă.

d) $\{|f| < c\} = \emptyset \in \mathcal{F}$, pentru $c \leq 0$ şi $\{|f| < c\} = \{-c < f < c\} \in \mathcal{F}$, pentru $c > 0$. Rezultă că funcţia $|f|$ este măsurabilă.

e) $\{f^2 < c\} = \emptyset \in \mathcal{F}$, pentru $c \leq 0$ şi $\{f^2 < c\} = \{-\sqrt{c} < f < \sqrt{c}\} \in \mathcal{F}$, pentru $c > 0$. Rezultă că funcţia f^2 este măsurabilă.

f) Avem $fg = \dfrac{1}{4}\left[(f+g)^2 - (f-g)^2\right]$. Rezultă că fg este o funcţie măsurabilă, conform *a), b), c)* şi *e)*.

g) Avem $\min\{f, g\} = \dfrac{1}{2}\left(f + g - |f - g|\right)$, deci funcţia $\min\{f, g\}$ este măsurabilă, conform *a), b), c), d)*.

h) Avem $\max\{f, g\} = \dfrac{1}{2}\left(f + g + |f - g|\right)$, deci funcţia $\max\{f, g\}$ este măsurabilă, conform *a), b), c), d)*. $\square$

Dacă $(\Omega, \mathcal{T})$ este un spaţiu topologic, atunci putem defini funcţii continue $f : \Omega \to \mathbb{R}$. Continuitatea unei astfel de funcţie implică măsurabilitatea sa în raport cu spaţiul măsurabil $(\Omega, \mathcal{B}_\Omega)$, unde $\mathcal{B}_\Omega = \mathcal{B}(\mathcal{T})$.

Propoziţia 1.4.3. *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 Spaţii de probabilitate

Noţiunea de spaţiu (sau câmp) de probabilitate este fundamentală în teoria probabilităţilor. Un spaţiu de probabilitate este un spaţiu măsurabil dotat cu un anumit tip de măsură finită, denumită "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 spaţiu (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 spaţiul 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* $\lim\limits_{n\to\infty} A_n = A$ *şi* $A_n \uparrow A$.

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* $\lim\limits_{n\to\infty} A_n = A$ *şi* $A_n \downarrow A$.

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 spaţiu de probabilitate. Au loc următoarele proprietăţi*

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

P5. $\mathbb{P}(A_1 \cup A_2 \cup \cdots \cup A_n) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \cdots + \mathbb{P}(A_n)$, $\forall\ A_1, A_2, \cdots, A_n \in \mathcal{F}$, *cu* $A_i \cap A_j = \emptyset$ *pentru* $i \neq j$, *unde* $n \geq 2$;

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

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

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)

$$\mathbb{P}(\cup_{i=1}^{n} A_i) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1\leq i_1 < i_2 < \cdots < i_k \leq n} \mathbb{P}(\cap_{j=1}^{k} A_{i_j}),\ \forall\ A_1, \cdots, A_n \in \mathcal{F},\ n \in \mathbb{N},\ n \geq 2;$$

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*

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

şi

$$\mathbb{P}\left(\cup_{k=1}^{n} A_k\right) \leq \sum_{k=1}^{n} \mathbb{P}(A_k), \ n \in \mathbb{N}^*.$$

Demonstraţie.

P4. Presupunem că $\mathbb{P}(\emptyset) = p > 0$. Considerăm şirul $A_n = \emptyset, \ n \in \mathbb{N}^*$. Atunci, conform $[P3]$, $p = \mathbb{P}(\emptyset) = \mathbb{P}(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n) = \sum_{n=1}^{\infty} p = \infty$. Contradicţie. Deci $\mathbb{P}(\emptyset) = 0$.

P5. Fie $A_1, A_2, \cdots, A_n \in \mathcal{F}$, cu $A_i \cap A_j = \emptyset$ pentru $i \neq j$. Considerăm şirul de evenimente $A_k = \emptyset, \ k = n+1, n+2, \cdots$. Atunci

$$\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) = \mathbb{P}\left(\bigcup_{k=1}^{\infty} A_k\right) \overset{[P3]}{=} \sum_{k=1}^{\infty} \mathbb{P}(A_k) \overset{[P4]}{=} \sum_{k=1}^{n} \mathbb{P}(A_k).$$

P6. Din $\Omega = A^c \cup A$, cu $A \cap A^c = \emptyset$, rezultă $1 \overset{[P2]}{=} \mathbb{P}(\Omega) = \mathbb{P}(A^c \cup A) \overset{[P5]}{=} \mathbb{P}(A^c) + \mathbb{P}(A)$, de unde $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

P7. Dacă $A \subset B$, atunci $B = A \cup (B \setminus A)$, cu $A \cap (B \setminus A) = \emptyset$.
Rezultă $\mathbb{P}(B) \overset{[P5]}{=} \mathbb{P}(A) + \mathbb{P}(B \setminus A)$, deci $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$.

P8. Dacă $A \subset B$, atunci $\mathbb{P}(B) - \mathbb{P}(A) \overset{[P7]}{=} \mathbb{P}(B \setminus A) \geq 0$, deci $\mathbb{P}(A) \leq \mathbb{P}(B)$.
Observaţie. Proprietatea reflectă monotonia crescătoare a funcţiei probabilitate,

$$\mathbb{P} : (\mathcal{F}, \subset) \to (\mathbb{R}, \leq).$$

P9. Fie $A \in \mathcal{F}$. $\mathbb{P}(A) \geq 0$ (conform $[P1]$); cum $A \subset \Omega$, avem $\mathbb{P}(A) \overset{[P8]}{\leq} \mathbb{P}(\Omega) \overset{[P2]}{=} 1$.

P10. Demonstrăm formula lui Poincaré, notată $P(n)$, prin inducţie matematică după numărul natural $n \geq 2$.

Fie $A_1, A_2 \in \mathcal{F}$. Avem

$$A_1 \cup A_2 = [A_1 \setminus (A_1 \cap A_2)] \cup [A_2 \setminus (A_1 \cap A_2)] \cup (A_1 \cap A_2),$$

unde termenii reuniunii sunt mulţimi mutual disjuncte. Atunci:

$$\mathbb{P}(A_1 \cup A_2) \overset{[P5]}{=} \mathbb{P}(A_1 \setminus (A_1 \cap A_2)) + \mathbb{P}(A_2 \setminus (A_1 \cap A_2)) + \mathbb{P}(A_1 \cap A_2)$$

$$\overset{[P7]}{=} [\mathbb{P}(A_1) - \mathbb{P}(A_1 \cap A_2)] + [\mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)] + \mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2).$$

Rezultă că formula este valabilă pentru $n = 2$, deci propoziţia $P(2)$ este adevărată.

Demonstrăm implicaţia $P(n) \Rightarrow P(n+1), \ \forall \ n \in \mathbb{N}, \ n \geq 2$. Presupunem că formula este adevărată pentru un număr natural $n \geq 2$ (deci propoziţia $P(n)$ este presupusă adevărată) şi o demonstrăm pentru $n+1$ (deci demonstrăm că propoziţia $P(n+1)$ este adevărată).

Fie $A_1, A_2, \cdots, A_n, A_{n+1} \in \mathcal{F}$. Avem

$$\mathbb{P}\left(\cup_{k=1}^{n+1} A_k\right) = \mathbb{P}\left((\cup_{k=1}^{n} A_k) \cup A_{n+1}\right) \overset{P(2)}{=} \mathbb{P}\left(\cup_{k=1}^{n} A_k\right) + \mathbb{P}(A_{n+1}) - \mathbb{P}\left((\cup_{k=1}^{n} A_k) \cap A_{n+1}\right) =$$

$$= \mathbb{P}\left(\cup_{k=1}^{n} A_k\right) + \mathbb{P}(A_{n+1}) - \mathbb{P}\left(\cup_{k=1}^{n} (A_k \cap A_{n+1})\right) \overset{P(n)}{=} \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \le i_1 < \cdots < i_k \le n} \mathbb{P}(\cap_{j=1}^{k} A_{i_j}) +$$

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

$$+ \sum_{k=2}^{n} (-1)^{k-1} \left[\sum_{1 \le i_1 < \cdots < i_k \le n} \mathbb{P}(\cap_{j=1}^{k} A_{i_j}) + \sum_{1 \le i_1 < \cdots < i_{k-1} \le n} \mathbb{P}((\cap_{j=1}^{k-1} A_{i_j}) \cap A_{n+1}) \right] +$$

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

Rezultă că formula este adevărată pentru $n+1$ (deci propoziţia $P(n+1)$ este adevărată). Formula lui Poincaré este o expersie a *principiului includerii-excluderii.*

P11. Fie $(A_n)_{n \ge 1}$ un şir monoton de evenimente.

a) Presupunem $A_n \downarrow \emptyset$.

Are loc reprezentarea $A_1 = \cup_{k=1}^{\infty} (A_k \setminus A_{k+1})$, unde $(A_i \setminus A_{i+1}) \cap (A_j \setminus A_{j+1}) = \emptyset,\ i \ne j$. Atunci $\mathbb{P}(A_1) = \sum_{k=1}^{\infty} \mathbb{P}(A_k \setminus A_{k+1})$. Convergenţa seriei implică $\lim_{n \to \infty} \sum_{k=n}^{\infty} \mathbb{P}(A_k \setminus A_{k+1}) = 0$. Dar $\sum_{k=n}^{\infty} \mathbb{P}(A_k \setminus A_{k+1}) = \mathbb{P}\left(\cup_{k=n}^{\infty} (A_k \setminus A_{k+1})\right) = \mathbb{P}(A_n)$. Rezultă $\lim_{n \to \infty} \mathbb{P}(A_n) = 0 = \mathbb{P}(\emptyset)$.

b) 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 $\lim_{n \to \infty} \mathbb{P}(A_n) = \mathbb{P}(A)$.

c) 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. Fie şirul de evenimente $(A_n)_{n \ge 1}$. Definim şirul $(B_n)_{n \ge 1},\ B_1 = A_1$, $B_n = A_n \setminus \left(\cup_{k=1}^{n-1} A_k\right)\ \forall\ n > 1$. Avem $\cup_{n=1}^{\infty} B_n = \cup_{n=1}^{\infty} A_n$, iar $B_i \cap B_j = \emptyset,\ \forall\ i \ne j$. Conform proprietăţilor $[P3]$ şi $[P8]$, obţinem

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

Cea de a doua relaţie se obţine analog. Astfel, demonstraţia propoziţiei este încheiată. $\square$

1.6 Evenimente independente, evenimente condiţionate

Un fenomen important evidenţiat de teoria probabilităţilor este cel al *independenţei evenimentelor.*

Definiţia 1.6.1. *Fie* $(\Omega, \mathcal{F}, \mathbb{P})$ *un spaţiu de probabilitate. Evenimentele* A *şi* B *se numesc independente dacă* $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$*. Mai general, o mulţime cel mult numărabilă* $\mathcal{G} \subset \mathcal{F}$ *reprezintă o familie de evenimente independente dacă pentru oricare evenimente distincte* $A_1, A_2, \cdots, A_n \in \mathcal{G}$ *avem*

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

Din definiţie rezultă că, pentru oricare eveniment A, A şi $\emptyset$, respectiv A şi Ω reprezintă perechi de evenimente independente. Are loc de asemenea următorul rezultat elementar.

Propoziţia 1.6.1. *Dacă* A *şi* B *sunt evenimente independente, atunci* A^c *şi* B *sunt de asemenea evenimente independente.*

Demonstraţie. Fie evenimentele independente A şi B. Avem deci $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$. Atunci

$$\mathbb{P}(A^c \cap B) = \mathbb{P}(B \setminus (A \cap B)) = \mathbb{P}(B) - \mathbb{P}(A \cap B) = \mathbb{P}(B) - \mathbb{P}(A) \cdot \mathbb{P}(B) = \mathbb{P}(B)(1 - \mathbb{P}(A)) = \mathbb{P}(A^c) \cdot \mathbb{P}(B).$$

Rezultă că evenimentele A^c şi B sunt independente. □

Vom defini în continuare *evenimentele condiţionate.*

Definiţia 1.6.2. *Fie* $(\Omega, \mathcal{F}, \mathbb{P})$ *un spaţiu de probabilitate. Considerăm* $A \in \mathcal{F}$*, cu* $\mathbb{P}(A) > 0$ *(A se numeşte eveniment neneglijabil). Pentru un eveniment* $B \in \mathcal{F}$ *numim probabilitatea lui* B *condiţionată de (realizarea evenimentului)* A *mărimea:*

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

Probabilitatea condiţionată $\mathbb{P}(B|A)$ *se mai notează* $\mathbb{P}_A(B)$.

Probabilitatea condiţionată $\mathbb{P}(B|A)$ reprezintă deci *probabilitatea realizării evenimentului* B*, ştiind ca evenimentul* A *s-a realizat.* Formula

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

a fost menţionată ca principiu de bază al teoriei probabilităţilor. Proprietăţile elementare ale probabilităţii condiţionate sunt evidenţiate în propoziţia următoare.

Propoziţia 1.6.2. *Fie* $(\Omega, \mathcal{F}, \mathbb{P})$ *un spaţiu de probabilitate.*

1. *Dacă* A *şi* B *sunt două evenimente neneglijabile, independente, atunci*

$$\mathbb{P}(B|A) = \mathbb{P}(B) \text{ şi } \mathbb{P}(A|B) = \mathbb{P}(A).$$

2. Dacă A şi B sunt două evenimente neneglijabile atunci

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

3. Pentru evenimentele $A_1, A_2, \cdots, A_n$ *(*$n \geq 2$*), cu* $\mathbb{P}(\cap_{i=1}^{n-1} A_i) > 0$*, are loc formula intersecţiei finite:*

$$\mathbb{P}(\cap_{i=1}^{n} A_i) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2|A_1) \cdot \mathbb{P}(A_3|A_1 \cap A_2) \cdots \mathbb{P}(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1}).$$

4. Dacă A este un eveniment neneglijabil, atunci funcţia $\mathbb{P}_A : \mathcal{F} \to \mathbb{R}$*, definită prin*

$$\mathbb{P}_A(X) = \mathbb{P}(X|A), \ \forall\ X \in \mathcal{F},$$

este o probabilitate pe spaţiul măsurabil $(\Omega, \mathcal{F})$*.*

Demonstraţie.

1. Conform definiţiei probabilităţii condiţionate şi definiţiei independenţei evenimentelor, obţinem

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

Similar deducem $\mathbb{P}(A|B) = \mathbb{P}(A)$. Se constată astfel că două evenimente sunt independente dacă probabilitatea realizării unuia dintre acestea nu este influienţată de realizarea celuilalt.

2. Din definiţia probabilităţii condiţionate constatăm că are loc relaţia:

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

3. Fie evenimentele $A_1, A_2, \cdots, A_n$ ($n \geq 2$), cu $\mathbb{P}(\cap_{i=1}^{n-1} A_i) > 0$. Din ipoteză deducem $\mathbb{P}(\cap_{i=1}^{k} A_i) > 0,\ k = 1, \cdots, n-1$ (conform monotoniei crescătoare a probabilităţii). Atunci

$$\mathbb{P}(A_1) \cdot \mathbb{P}(A_2|A_1) \cdot \mathbb{P}(A_3|A_1 \cap A_2) \cdots \mathbb{P}(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})$$

$$= \mathbb{P}(A_1) \cdot \frac{\mathbb{P}(A_1 \cap A_2)}{\mathbb{P}(A_1)} \cdot \frac{\mathbb{P}(A_1 \cap A_2 \cap A_3)}{\mathbb{P}(A_1 \cap A_2)} \cdots \frac{\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n)}{\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_{n-1})} = \mathbb{P}(\cap_{i=1}^{n} A_i).$$

4. Verificăm axiomele probabilităţii.

P1. Evident, $\mathbb{P}_A(X) \geq 0, \ \forall\ X \in \mathcal{F}$.

P2. Avem $\mathbb{P}_A(\Omega) = \dfrac{\mathbb{P}(\Omega \cap A)}{\mathbb{P}(A)} = \dfrac{\mathbb{P}(A)}{\mathbb{P}(A)} = 1$.

P3. Fie un şir $(X_n)_{n \geq 1}$ de evenimente incompatibile mutual. Atunci

$$\mathbb{P}_A(\cup_{n=1}^{\infty} X_n) = \frac{\mathbb{P}((\cup_{n=1}^{\infty} X_n) \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(\cup_{n=1}^{\infty} (X_n \cap A))}{\mathbb{P}(A)} = \frac{\sum_{n=1}^{\infty} \mathbb{P}(X_n \cap A)}{\mathbb{P}(A)} = \sum_{n=1}^{\infty} \mathbb{P}_A(X_n).$$

P_A se numeşte probabilitatea condiţionată de evenimentul A. □

Pe de altă parte, conceptul de *evenimente condiţionate* permite considerarea *sistemelor complete de evenimente.*

Definiţia 1.6.3. *O mulţime de evenimente* $\mathcal{S} = \{A_i,\ i \in I\}$, *unde* I *este o mulţime cel mult numărabilă de indici, se numeşte un sistem complet de evenimente în spaţiul 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 1.6.3. *Fie* $(\Omega, \mathcal{F}, \mathbb{P})$ *un spaţiu de probabilitate, iar* $\mathcal{S} = \{A_i,\ i \in I\}$, *un sistem complet de evenimente. Au loc formulele:*

1. *Formula probabilităţii totale*
$$\mathbb{P}(A) = \sum_{i\in I} \mathbb{P}(A|A_i)\mathbb{P}(A_i),\ A \in \mathcal{F};$$
2. *Formula lui Bayes*
$$\mathbb{P}(A_k|A) = \frac{\mathbb{P}(A|A_k)\mathbb{P}(A_k)}{\sum_{i\in I}\mathbb{P}(A|A_i)\mathbb{P}(A_i)},\quad \forall\ k \in I,\ A \in \mathcal{F},\ \textit{astfel ca } \mathbb{P}(A) > 0.$$

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 1.6.2 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. □

În finalul secţiunii, prezentăm două rezultate clasice importante.

Propoziţia 1.6.4. *Fie* $(\Omega, \mathcal{F}, \mathbb{P})$ *un spaţiu de probabilitate. Pentru evenimentele* $A_1, A_2, \cdots, A_n$ *are loc următoarea inegalitate:*
$$\mathbb{P}(\cap_{i=1}^n A_i) \geq 1 - \sum_{i=1}^n \mathbb{P}(A_i^c).$$

Demonstraţie. Fie evenimentele $A_1, A_2, \cdots, A_n$. Avem
$$\mathbb{P}(\cap_{i=1}^n A_i) = 1 - \mathbb{P}((\cap_{i=1}^n A_i)^c) = 1 - \mathbb{P}(\cup_{i=1}^n A_i^c) \overset{P12}{\geq} 1 - \sum_{i=1}^n \mathbb{P}(A_i^c).$$
Rezultatul poartă denumirea de *inegalitatea lui Boole*. □

Teorema 1.6.1 (Legea 0-1 (Borel-Cantelli)). *Fie* $(\Omega, \mathcal{F}, \mathbb{P})$ *un spaţiu de probabilitate. Considerăm un şir* $(A_n)_{n\in\mathbb{N}^*}$ *de evenimente. Notăm*

$$A = \limsup_{n\to\infty} A_n = \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k$$

şi $s = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$. *Au loc proprietăţile:*

a) dacă $s < \infty$, *atunci* $\mathbb{P}(A) = 0$*;*

b) dacă $s = \infty$ *şi evenimentele şirului sunt mutual independente, atunci* $\mathbb{P}(A) = 1$.

Demonstraţie.

a) Presupunem $s < \infty$. Atunci $\lim_{n\to\infty} \sum_{k=n}^{\infty} \mathbb{P}(A_k) = 0$. Are loc convergenţa $\cup_{k=n}^{\infty} A_k \downarrow A$. Conform proprietăţilor P11 şi P12, obţinem

$$\mathbb{P}(A) = \lim_{n\to\infty} \mathbb{P}(\cup_{k=n}^{\infty} A_k) \le \lim_{n\to\infty} \sum_{k=n}^{\infty} \mathbb{P}(A_k) = 0.$$

b) Presupunem $s = \infty$. Avem

$$\mathbb{P}(A) = 1 - \mathbb{P}(A^c) = 1 - \mathbb{P}((\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k)^c) = 1 - \mathbb{P}(\cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k^c).$$

Cum $\cap_{k=n}^{\infty} A_k^c \uparrow A^c$ şi $\cap_{k=n}^{n+p} A_k^c \downarrow \cap_{k=n}^{\infty} A_k^c$, obţinem

$$\mathbb{P}(\cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k^c) = \lim_{n\to\infty} \mathbb{P}(\cap_{k=n}^{\infty} A_k^c) = \lim_{n\to\infty} \lim_{p\to\infty} \mathbb{P}(\cap_{k=n}^{n+p} A_k^c) \overset{indep.}{=} \lim_{n\to\infty} \lim_{p\to\infty} \prod_{k=n}^{n+p} \mathbb{P}(A_k^c).$$

Deoarece $s = \infty$, avem

$$\lim_{p\to\infty} \sum_{k=n}^{n+p} \mathbb{P}(A_k) = \sum_{k=n}^{\infty} \mathbb{P}(A_k) = \infty, \ \forall\, n \ge 1. \tag{1.1}$$

Din inegalitatea binecunoscută $e^x \ge 1 + x, \ \forall\, x \in \mathbb{R}$, obţinem

$$0 \le \mathbb{P}(A_k^c) = 1 - \mathbb{P}(A_k) \le e^{-\mathbb{P}(A_k)}, \ \forall\, k \in \mathbb{N}.$$

Rezultă

$$\lim_{p\to\infty} \prod_{k=n}^{n+p} \mathbb{P}(A_k^c) \le \lim_{p\to\infty} \prod_{k=n}^{n+p} e^{-\mathbb{P}(A_k)} = \lim_{p\to\infty} e^{-\sum_{k=n}^{n+p} \mathbb{P}(A_k)} = e^{-\sum_{k=n}^{\infty} \mathbb{P}(A_k)} \overset{(1.1)}{=} 0.$$

Ca urmare, $\mathbb{P}(A) = 1$. $\square$

1.7 Scheme clasice de probabilitate

Prezentăm în continuare modelele de bază de spaţii discrete de probabilitate.

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. *Spaţiu 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 $\sum_{i\in I} p_i = 1$. Definim $\mathbb{P}(\{\omega_i\}) = p_i$, $i \in I$. Atunci, 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 $P_{k|n} = \mathbb{P}\left(A_{k|n}\right)$ probabilitatea evenimentului menţionat. Un raţionament elementar ne conduce la formula de calcul
   $$P_{k|n} = C_n^k p^k q^{n-k},\ k = 0, 1, \cdots, n,$$
   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
   $$\sum_{k=0}^{n} \mathbb{P}\left(A_{k|n}\right) = \sum_{k=0}^{n} P_{k|n} = \sum_{k=0}^{n} C_n^k p^k q^{n-k} = (q+p)^n = 1,$$
   deci evenimentele $A_{k|n}$, $k = 0, 1, \cdots, n$, reprezintă un sistem complet.

2. *Schema multinomială.*
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 $\overline{k} = (k_i)_{i=\overline{1,s}}$ un vector de dimensiune s, cu componente naturale, având proprietatea $\sum_{i=1}^{s} k_i = n$. Notăm $A_{\overline{k}|n}$ evenimentul extragerii a n bile având compoziţia $\overline{k}$, adică extragerea a exact k_i bile de culoarea i, unde $i = 1, 2, \cdots, s$, din totalul de n bile extrase (cu repunere). Fie $P_{\overline{k}|n} = \mathbb{P}(A_{\overline{k}|n})$. Se constată că
$$P_{\overline{k}|n} = \binom{n}{k_1,\, k_2, \cdots,\, k_s} p_1^{k_1} p_2^{k_2} \cdots p_s^{k_s},$$
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 $k_1,\, k_2, \cdots,\, k_s$.
Avem
$$\sum_{\overline{k}=(k_i)\in\mathbb{N}^s,\ \sum_{i=1}^{s} k_i = n} P_{\overline{k}|n} = \sum_{(k_i)\in\mathbb{N}^s;\ \sum_{i=1}^{s} k_i=n} \binom{n}{k_1,\, k_2, \cdots,\, k_s} p_1^{k_1} p_2^{k_2} \cdots p_s^{k_s} =$$
$$= (p_1 + p_2 + \cdots + p_s)^n = 1,$$
deci $A_{\overline{k}|n}$, cu $\overline{k} = (k_i) \in \mathbb{N}^s$, astfel ca $\sum_{i=1}^{s} k_i = n$, formează un sistem complet de evenimente. *Schema multinomială*, denumită şi *schema polinomială* este o extindere a schemei lui Bernoulli.

3. *Schema lui Poisson.*
Se consideră n urne cu bile albe şi negre, dar în proporţii diferite. Fie $p_i \in (0, 1)$ 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
$$P_{k|n} = \mathbb{P}\left(A_{k|n}\right) = \sum_{I\subset\mathbb{N}_n;\ |I|=k} \prod_{i\in I} p_i \prod_{j\in\mathbb{N}_n\setminus I} q_j.$$
Observăm că $P_{k|n}$ reprezintă coeficientul lui X^k al polinomului
$$f(X) = (p_1 X + q_1)(p_2 X + q_2) \cdots (p_n X + q_n).$$
Astfel, $f(X) = P_{n|n} X^n + P_{n-1|n} X^{n-1} + \cdots + P_{1|n} X + P_{0|n}$. Rezultă
$$\sum_{k=0}^{n} P_{k|n} = f(1) = \prod_{i=1}^{n} (p_i + q_i) = 1,$$
deci evenimentele $A_{k|n},\ k = 0, 1, \cdots, n$, reprezintă un sistem complet. *Schema lui Poisson* reprezintă o generalizare a schemei lui Bernoulli.

4. *Schema geometrică.*
   Schema este un exemplu de spaţiu 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 $\overline{k} = (k_i) \in \mathbb{N}^s$, 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

$$P_{\overline{k}} = \frac{C_{a_1}^{k_1} C_{a_2}^{k_2} \cdots C_{a_s}^{k_s}}{C_N^n}.$$

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 spaţiu 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

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

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 spaţiu 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 numărabilă de indici, formează un sistem complet de evenimente în spaţiul 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 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 $|M|$ 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. *Spaţiu 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 spaţiu 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 posibile, exprimabile prin numere reale, ale unui anumit experiment. Fiecare dintre 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 spaţiu de probabilitate. O funcţie reală măsurabilă* $X : \Omega \to \mathbb{R}$ *se numeşte variabilă aleatoare.*

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

$$X^{-1}(I) = \{\omega \in \Omega : \; X(\omega) \in I\} \overset{notat}{=} \{X \in I\} \in \mathcal{F},$$

pentru orice interval real I. In particular,

$$\{X \in (-\infty, a]\} = \{X \leq a\} \in \mathcal{F}, \; \forall \, a \in \mathbb{R}. \tag{2.1}$$

Probabilităţile evenimentelor descrise în relaţia (2.1) oferă informaţii importante despre repartiţia valorilor variabilei aleatoare X. În secţiunea următoare, va fi introdusă şi caracterizată *funcţia de repartiţie* a a unei variabile aleatoare, construită pe baza probabilităţilor evenimentelor menţionate mai sus.

O mulţime finită de variabile aleatoare definite pe acelaşi spaţiu 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 spaţiu de probabilitate. Considerăm borelianul* $\mathcal{B}_{\mathbb{R}^d}$ *spaţiului* $\mathbb{R}^d$ *(*$\sigma -$ *corpul generat de topologia lui* $\mathbb{R}^d$*). O funcţie* $(\mathcal{F}, \mathcal{B}_{\mathbb{R}^d})$ *măsurabilă* $X : \Omega \to \mathbb{R}^d$ *se numeşte vector aleator.*

Fie $X_i, \; i = 1, 2, \cdots, d$, componentele scalare ale vectorului aleator $X : \Omega \to \mathbb{R}^d$, deci $X = (X_1, X_2, \cdots, X_d)$. Deoarece σ-corpul $\mathcal{B}_{\mathbb{R}^d}$ conţine toate tipurile de intervale

d-dimensionale din spaţiul $\mathbb{R}^d$, avem

$$(-\infty, a_1] \times (-\infty, a_2] \times \cdots \times (-\infty, a_d] \in \mathcal{B}_{\mathbb{R}^d}, \ \forall\, a_1, a_2, \cdots, a_d \in \mathbb{R}.$$

Ca urmare, din Definiţia 2.2.2, rezultă

$$\{X_1 \leq a_1, X_2 \leq a_2, \cdots, X_d \leq a_d\} = \left\{ X \in \prod_{i=1}^{d} (-\infty, a_i] \right\} \in \mathcal{F}, \ \forall\ (a_1, a_2, \cdots, a_d) \in \mathbb{R}^d.$$

2.3 Funcţia de repartiţie

Distribuţia valorilor unei variabile aleatoare reale este caracterizată prin noţiunea de funcţie de repartiţie. De regulă, spaţiul 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 spaţiu 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. *Funcţia de repartiţie $F : \mathbb{R} \to [0, 1]$ a unei variabilei aleatoare X are următoarele proprietăţi:*

1. *F este monoton crescătoare pe $\mathbb{R}$;*
2. *F este continuă la dreapta în orice număr real, deci $F(x) = F(x + 0), \ \forall\ x \in \mathbb{R}$;*
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 spaţiu de probabilitate şi $X : \Omega \to \mathbb{R}$ o variabilă aleatoare.
1. Fie $x_1, x_2 \in \mathbb{R}$, cu $x_1 < x_2$. Atunci $\{X \leq x_1\} \subset \{X \leq x_2\}$. Cum funcţia $\mathbb{P}$ este monoton crescătoare, deducem $\mathbb{P}\{X \leq x_1\} \leq \mathbb{P}\{X \leq x_2\}$, sau $F(x_1) \leq F(x_2)$. Prin urmare, F este monoton crescătoare pe $\mathbb{R}$.
2. Orice funcţie monotonă pe un interval admite limite laterale finite in fiecare punct al intervalului. Dacă intervalul este deschis, atunci funcţia admite limite în extremităţile intervalului. Mai mult, mulţimea punctelor de discontinuitate ale unei astfel de funcţii este cel mult numărabilă. Funcţia F este monoton crescătoare pe $\mathbb{R}$, deci admite limite

laterale finite în toate numerele reale. Mai precis, pentru oricare $x \in \mathbb{R}$, limita la stânga a funcţiei F în punctul x este $F(x-0) = \sup_{t<x} F(t)$, iar limita la dreapta a funcţiei F în punctul x este $F(x+0) = \inf_{t>x} F(t)$. În plus, au loc inegalităţile:

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

Fixăm $x \in \mathbb{R}$, arbitrar. Notăm $A = \{X \leq x\}$ şi $A_n = \left\{X \leq x + \frac{1}{n}\right\} \in \mathcal{F},\ \forall\ n \in \mathbb{N}^*$. Şirul $(A_n)_{n\in\mathbb{N}^*}$ este monoton descrescător ($A_{n+1} \subset A_n,\ \forall\ n \in \mathbb{N}^*$) şi $\lim\limits_{n\to\infty} A_n = \cap_{n=1}^{\infty} A_n = A$ (notăm $A_n \downarrow A$). Conform Propoziţiei 1.5.1, avem $\lim\limits_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A)$. Dar $\mathbb{P}(A) = F(x)$ şi $\lim\limits_{n\to\infty} \mathbb{P}(A_n) = \lim\limits_{n\to\infty} F\left(x + \dfrac{1}{n}\right) = F(x+0)$. Obţinem $F(x+0) = F(x)$, deci F este continuă la dreapta în punctul x.

3. Conform monotoniei, F admite limite la ∞ şi $-\infty$. Deoarece $F(\mathbb{R}) \subset [0,1]$, aceste limite sunt finite. Mai precis,

$$\lim_{x\to-\infty} F(x) = \inf_{x\in\mathbb{R}} F(x) \geq 0$$

şi

$$\lim_{x\to\infty} F(x) = \sup_{x\in\mathbb{R}} F(x) \leq 1.$$

Definim şirurile de evenimente

$$B_n = \{X \leq -n\} \text{ şi } C_n = \{X \leq n\},\ n \in \mathbb{N}^*.$$

Avem $B_n \downarrow \cap_{n=1}^{\infty} B_n = \emptyset$, respectiv $C_n \uparrow \cup_{n=1}^{\infty} C_n = \Omega$. 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

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

$$= \mathbb{P}\left(\lim_{n\to\infty}\left\{X \leq x - \frac{1}{n}\right\}\right) = \mathbb{P}\left(\bigcup_{n=1}^{\infty}\left\{X \leq x - \frac{1}{n}\right\}\right) = \mathbb{P}\{X < x\}.$$

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

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

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

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

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

Din Propoziţia 2.3.1, punctul 5, rezultă următoarea proprietate: dacă F este continuă în $x_0 \in \mathbb{R}$, atunci $\mathbb{P}\{X = x_0\} = 0$.

Menţionăm următorul rezultat fundamental: pentru orice funcţie $F : \mathbb{R} \to [0,1]$ care satisface proprietăţile 1-3 din Propoziţia 2.3.1, există un spaţiu de probabilitate $(\Omega, \mathcal{F}, \mathbb{P})$ şi o variabilă aleatoare $X : \Omega \to \mathbb{R}$ astfel încât F să reprezinte funcţia de repartiţie a variabilei aleatoare X.

Două variabile aleatoare X şi Y care admit o funcţie de repartiţie comună vor fi denumite *identic distribuite*. Notăm $X \stackrel{d}{=} Y$.

Definim funcţia de repartiţie a unui vector aleator, care prezintă 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 $F : \mathbb{R}^d \to [0,1]$,*

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

2.4 Variabile aleatoare independente

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

Definiţia 2.4.1. *Variabilele aleatoare X_i, $i \in I$, definite pe un spaţiu 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

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

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

Propoziţia 2.4.1. *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 funcţia $F * G$ (convoluţia funcţiilor de repartiţie F şi G), definită prin:*

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

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*. Vom ilustra prin exemple semnificative cele două tipuri de variabile aleatoare.

2.5 Variabile aleatoare discrete. Exemple clasice

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

Definiţia 2.5.1. *Fie $(\Omega, \mathcal{F}, \mathbb{P})$ un spaţiu 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 $A_i = \{X = x_i\} \in \mathcal{F},\ i \in I$. Vom presupune $\mathbb{P}(A_i) > 0,\ \forall\ i \in I$. Atunci familia de evenimente $\{A_i, i \in I\}$ reprezintă un sistem complet de evenimente al spaţiului de probabilitatea $(\Omega, \mathcal{F}, \mathbb{P})$.

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:

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

Notăm $p_i = \mathbb{P}(A_i),\ i \in I$. Avem $p_i > 0,\ \forall\ i \in I$ şi

$$\sum_{i \in I} p_i = \sum_{i \in I} \mathbb{P}(A_i) = \mathbb{P}(\cup_{i \in I} A_i) = \mathbb{P}(\Omega) = 1.$$

Distribuţia variabilei aleatoare X se reprezintă convenţional prin:

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

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

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

Variabila aleatoare $X + Y$ are distribuţia

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

iar variabila aleatoare XY are distribuţia

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

Exemplul 2.5.1. *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 $X \sim B(1,p)$. Semnificaţia unei variabile aleatoare Bernoulli de parametru $p \in (0,1)$ este următoarea: în cadrul unei experienţe, urmărim producerea unui anumit eveniment A care are probabilitatea de apariţie $\mathbb{P}(A) = p$; variabila aleatoare Bernoulli indică prin valoarea 1 realizarea evenimentului A, iar prin valoarea 0 nerealizarea evenimentului respectiv. Constatăm că X este funcţia caracteristică c_A a evenimentului A.

Exemplul 2.5.2. *Variabile aleatoare binomiale.*

O variabilă aleatoare binomială X de parametri $n \in \mathbb{N}^*$ şi $p \in (0,1)$ ia toate valorile $k \in \{0, 1, \cdots, n\}$, cu probabilităţile $p_k = \mathbb{P}\{X = k\} = C_n^k p^k q^{n-k}$, unde $q = 1 - p \in (0,1)$. Variabila aleatoare X are deci distribuţia:

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

Avem

$$\sum_{k=0}^{n} p_k = \sum_{k=0}^{n} C_n^k p^k q^{n-k} = (q+p)^n = 1.$$

Notăm $X \sim B(n,p)$. O variabilă aleatoare binomială indică numărul de realizări ale unui eveniment A, având probabilitatea $p = \mathbb{P}(A)$, în n experienţe identice şi independente. Astfel, variabilele binomiale modelează schema binomială (Bernoulli).

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

Propoziţia 2.5.1. *Fie n variabile aleatoare independente şi identic distribuite $X_1, \cdots, X_n$, cu $X_i \sim B(1,p)$, pentru $i = 1, \cdots, n$. Dacă $X = \sum_{i=1}^n X_i$, atunci $X \sim B(n,p)$.*

Demonstraţie. Deoarece $X_i \in \{0,1\}$, $i = 1, \cdots, n$, avem $X \in \{0, 1, \cdots, n\}$. Notăm $p_k = \mathbb{P}\{X = k\}$, $k = 0, 1, \cdots, n$. Atunci:

$$p_0 = \mathbb{P}\{X_1 = 0, \cdots, X_n = 0\} = \mathbb{P}\left(\cap_{i=0}^n \{X_i = 0\}\right) \overset{indep.}{=} \prod_{i=1}^{n} \mathbb{P}\{X_i = 0\} = q^n;$$

$$p_n = \mathbb{P}\{X_1 = 1, \cdots, X_n = 1\} = \mathbb{P}\left(\cap_{i=0}^n \{X_i = 1\}\right) \overset{indep.}{=} \prod_{i=1}^{n} \mathbb{P}\{X_i = 1\} = p^n;$$

pentru $1 < k < n$,

$$p_k = \sum_{I \subset \{1,\cdots,n\},\ |I|=k} \mathbb{P}\{X_i = 1,\ \forall\, i \in I; X_j = 0,\ \forall\, j \in \{1, \cdots, n\} \setminus I\} =$$

$$= \sum_{I \subset \{1,\cdots,n\},\ |I|=k} \mathbb{P}\left[\left(\cap_{i \in I} \{X_i = 1\}\right) \cap \left(\cap_{j \in \{1,\cdots,n\} \setminus I} \{X_j = 0\}\right)\right] =$$

$$\overset{indep}{=} \sum_{I \subset \{1,\cdots,n\},\ |I|=k} \left(\prod_{i \in I} \mathbb{P}\{X_i = 1\}\right) \cdot \left(\prod_{j \in I^c} \mathbb{P}\{X_j = 0\}\right) = C_n^k p^k q^{n-k}.$$

Rezultă $X \sim B(n,p)$. $\square$

Exemplul 2.5.3. *Variabile aleatoare Poisson.*

O variabilă aleatoare X este de tip Poisson de parametru λ dacă ia toate valorile naturale, cu probabilităţile:

$$p_n = \mathbb{P}\{X = n\} = \frac{\lambda^n}{n!} e^{-\lambda},\ n \in \mathbb{N}.$$

Astfel, X are distribuţia:

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

Notăm $X \sim P(\lambda)$. Avem

$$\sum_{n=0}^{\infty} p_n = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} e^{-\lambda} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} = e^{-\lambda} e^{\lambda} = 1.$$

Exemplul 2.5.4. *Variabile aleatoare geometrice.*

O variabilă aleatoare geometrică X de parametru $p \in (0,1)$ ia valoarea $n \in \mathbb{N}^*$ cu probabilitatea $p_n = \mathbb{P}\{X = n\} = pq^{n-1}$, unde $q = 1 - p \in (0,1)$. Astfel, X are distribuţia:

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

Notăm $X \sim G(p)$. Avem

$$\sum_{n=1}^{\infty} p_n = \sum_{n=1}^{\infty} pq^{n-1} = p \sum_{n=1}^{\infty} q^{n-1} = p \cdot \frac{1}{1-q} = 1.$$

Variabila aleatoare geometrică de parametru p indică numărul de experienţe, identice şi independente, efectuate până la realizarea (apariţia), pentru prima dată, a unui eveniment A a cărui probabilitate de realizare (în cadrul fiecărei experienţe) este egală cu p. Variabilele aleatoare geometrice modelează schema geometrică.

2.6 Variabile aleatoare continue. Exemple clasice

O clasă importantă de variabile aleatoare o constituie variabilele de *tip continuu*. O variabilă aleatoare continuă admite *funcţie de densitate* (*densitate de probabilitate*).

Definiţia 2.6.1. *Fie $(\Omega, \mathcal{F}, \mathbb{P})$ un spaţiu de probabilitate şi $X : \Omega \to \mathbb{R}$ o variabilă aleatoare. X se numeşte variabilă aleatoare continuă dacă există o funcţie $f : \mathbb{R} \to [0, \infty)$, continuă pe $\mathbb{R}$, eventual cu excepţia a unui număr finit de puncte, astfel încât funcţia de repartiţie F a variabilei aleatoare X să fie determinată prin:*

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

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

Proprietăţile funcţiei de densitate a unei variabile aleatoare de tip continuu sunt evidenţiate în propoziţia de mai jos.

Propoziţia 2.6.1. *Fie* $f : \mathbb{R} \to [0, \infty)$ *funcţia de densitate asociată unei variabile aleatoare de tip continuu* $X : \Omega \to \mathbb{R}$*, iar* $F : \mathbb{R} \to [0, 1]$ *funcţia sa de repartiţie, definită prin:*

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

Atunci:

1. F *este continuă pe* $\mathbb{R}$*;*
2. F *este derivabilă în toate punctele* $x \in \mathbb{R}$ *de continuitate ale lui* f*, cu* $F'(x) = f(x)$*; dacă, în particular,* f *este continuă pe* $\mathbb{R}$*, atunci* F *este de clasă* $\mathcal{C}^1$ *pe* $\mathbb{R}$*, cu* $F' = f$*.*
3. $\displaystyle\int_{-\infty}^{\infty} f(t)\, dt = 1$*;*
4. $\mathbb{P}\{X = x\} = 0, \ \forall\, x \in \mathbb{R}$*;*
5. $\mathbb{P}\{a \leq X \leq b\} = \displaystyle\int_{a}^{b} f(t)\, dt, \ \forall\, [a, b] \subset \mathbb{R}.$

Demonstraţie.

1. Vom analiza doar cazul când f este mărginită pe $\mathbb{R}$. Presupunem că există $M > 0$ astfel ca $f(t) \in [0, M], \ \forall\, t \in \mathbb{R}$. Fie $x \in \mathbb{R}$, arbitrar. Atunci, pentru $h \in \mathbb{R}^*$, avem

$$|F(x+h) - F(x)| = \left| \int_{-\infty}^{x+h} f(t)\, dt - \int_{-\infty}^{x} f(t)\, dt \right| = \left| \int_{x}^{x+h} f(t)\, dt \right| \leq \left| \int_{x}^{x+h} |f(t)|\, dt \right| \leq M|h|.$$

Aplicând criteriul majorării, obţinem $\lim_{h \to 0} |F(x+h) - F(x)| = 0$, deci $\lim_{h \to 0} F(x+h) = F(x)$. Rezultă că F este continuă în x.

2. Fie $x \in \mathbb{R}$ un punct de continuitate a lui f. Considerăm $\varepsilon > 0$, arbitrar. Din continuitatea lui f în x, rezultă că există $\delta > 0$ astfel ca $|f(t) - f(x)| < \varepsilon, \ \forall\, t \in (x - \delta, x + \delta)$. Atunci, pentru $h \in (-\delta, \delta) \setminus \{0\}$, avem

$$\left| \frac{F(x+h) - F(x)}{h} - f(x) \right| = \frac{\left| \int_x^{x+h} f(t)\, dt - h f(x) \right|}{|h|} = \frac{\left| \int_x^{x+h} [f(t) - f(x)]\, dt \right|}{|h|} \leq$$

$$\leq \frac{\left| \int_x^{x+h} |f(t) - f(x)|\, dt \right|}{|h|} \leq \frac{\varepsilon |h|}{|h|} = \varepsilon.$$

Rezultă $\lim_{h \to 0} \dfrac{F(x+h) - F(x)}{h} = f(x)$, deci F este derivabilă în x, cu $F'(x) = f(x)$.

3. $\displaystyle\int_{-\infty}^{\infty} f(t)\, dt = \lim_{x \to \infty} \int_{-\infty}^{x} f(t)\, dt = \lim_{x \to \infty} F(x) = 1.$

4. F este continuă pe $\mathbb{R}$, deci $\mathbb{P}\{X = x\} = F(x) - F(x - 0) = 0, \ \forall\, x \in \mathbb{R}.$

5. Fie $a, b \in \mathbb{R}$, cu $a < b$. Atunci

$$\mathbb{P}\{a \leq X \leq b\} = \mathbb{P}(\{X \leq b\} \setminus \{X < a\}) = \mathbb{P}\{X \leq b\} - \mathbb{P}\{X < a\} =$$

$$= \mathbb{P}\{X \leq b\} - \mathbb{P}\{X \leq a\} = F(b) - F(a) = \int_a^b f(t)\, dt. \quad \square$$

Teorema 2.6.1. *Fie X şi Y două variabile aleatoare continue şi independente, cu densităţiile f şi respectiv g , $f, g : \mathbb{R} \to [0, \infty)$. Atunci variabila aleatoare $X+Y$ are densitatea $f * g$, unde*

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

reprezintă convoluţia densităţilor f şi g.

Reţinem următorul rezultat fundamental.

Teorema 2.6.2. *Dacă o funcţie $f : \mathbb{R} \to [0, \infty)$, continuă pe $\mathbb{R}$, eventual cu excepţia unui număr finit de puncte, are proprietatea*

$$\int_{-\infty}^{\infty} f(t) dt = 1,$$

atunci există un spaţiu de probabilitate $(\Omega, \mathcal{F}, \mathbb{P})$ şi o variabilă aleatoare $X : \Omega \to \mathbb{R}$ astfel încât f să reprezinte funcţia de densitate a lui X.

Vom prezenta în continuare câteva exemple clasice de repartiţii continue.

Exemplul 2.6.1. *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:

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

Verificăm condiţia $\displaystyle\int_{-\infty}^{\infty} \varphi(t)\, dt = 1$, utilizând funcţia Γ a lui Euler, definită prin

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx, \ \alpha > 0.$$

Este cunoscut faptul că $\Gamma(1/2) = \displaystyle\int_0^{\infty} x^{-1/2} e^{-x} dx = \sqrt{\pi}$. Astfel, avem

$$\int_{-\infty}^{\infty} \varphi(t)\, dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{t^2}{2}} dt \overset{x=t^2/2}{=} \sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{-x} (\sqrt{2x})' dx =$$

$$= \frac{1}{\sqrt{\pi}} \int_0^{\infty} \frac{e^{-x}}{\sqrt{x}} dx = \frac{1}{\sqrt{\pi}} \Gamma(1/2) = 1.$$

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

Generalizare. $X \sim N(\mu, \sigma^2)$, 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 2.6.2. *Distribuţia exponenţială.*

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

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

Funcţia de densitate exponenţială de parametru a este

$$f(x) = F'(x) = ae^{-ax}, \ \forall\, x \geq 0.$$

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

Exemplul 2.6.3. *Distribuţia* Γ.

Variabila aleatoare X are o distribuţie Γ de parametrii $n, a > 0$, notat prin $X \sim \Gamma(n, a)$, dacă ia valori în intervalul $(0, \infty)$ şi admite funcţia de densitate

$$f(x) = \frac{a^n x^{n-1} e^{-ax}}{\Gamma(n)}, \ \forall\, x > 0.$$

Pentru $n = 1$ se obţine distribuţia exponenţială de parametru a, deci $\Gamma(1, a) = Exp(a)$. Mai general, pentru $n \in \mathbb{N}^*$, funcţia de repartiţie a unei variabile aleatoare $X \sim \Gamma(n, a)$ poate fi explicitată. Astfel, $\Gamma(n) = (n-1)!$, deci funcţia de densitate devine

$$f(x) = \frac{a^n x^{n-1} e^{-ax}}{(n-1)!}, \ \forall\, x > 0.$$

Se obţine funcţia de repartiţie:

$$F(x) = \int_0^x f(t)\, dt = 1 - e^{-ax} \sum_{k=0}^{n-1} \frac{(ax)^k}{k!}, \ x \geq 0.$$

În statistică, este importantă distribuţia χ^2_n, definită ca $\Gamma(n/2, 1/2)$, cu $n \in \mathbb{N}^*$.

Exemplul 2.6.4. *Distribuţia* β.

Variabila aleatoare X are o distribuţie β de parametri $a, b > 0$, notat prin $X \sim \beta(a, b)$, dacă ia valori în intervalul $(0, 1)$ şi admite funcţia de densitate

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

unde

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

este funcţia lui Euler de a doua speţă.

2.7 Media. Medii de ordin superior

2.7.1 Definiţie, proprietăţi

Pentru o variabilă aleatoare reală, **media** şi **dispersia** reprezintă cele mai elementare şi în acelaşi timp cele mai semnificative caracteristici numerice.

Media unei variabile aleatoare (dacă există) indică "valoarea aşteptată" a variabilei aleatoare.

Definiţia 2.7.1. *Fie* $(\Omega, \mathcal{F}, \mathbb{P})$ *un spaţiu 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 după cum urmează.

1. Pentru o variabilă aleatoare X de tip discret, care ia valorile x_i cu probabilităţile p_i, pentru $i \in I$, media devine
$$\mathbb{E}(X) = \sum_{i \in I} x_i p_i.$$
Dacă I este o mulţime numărabilă, atunci media există în ipoteza că seria numerică $\sum_{i \in I} x_i p_i$ este convergentă.

2. Pentru o variabilă aleatoare X de tip continuu, care admite funcţia de densitate (densitatea de probabilitate) $f : \mathbb{R} \to [0, \infty)$, media devine
$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x) \, dx,$$
în ipoteza că integrala improprie este convergentă.

Următoarea propoziţie stabileşte proprietăţile de bază ale mediei variabilelor aleatoare.

Propoziţia 2.7.1. *Fie* X *şi* Y *două variabile aleatoare cu medie finită. Au loc proprietăţile:*

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 variabilelor aleatoare discrete. Rezultatele pentru variabilele aleatoare de tip continuu se pot deduce prin aproximarea acestora cu termenii unui şir de variabile aleatoare de tip discret şi trecerea la limită în şirul mediilor acestora.

Presupunem că variabilele aleatoare discrete 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. Pentru a demonstra relaţia $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$, considerăm evenimentele

$$A_{ij} = \{X = x_i,\ Y = y_j\},\quad (i,j) \in I \times J.$$

Notăm $p_{ij} = \mathbb{P}(A_{ij}),\ (i,j) \in I \times J$. 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

$$\mathbb{E}(X+Y) = \sum_{(i,j)\in I\times J} (x_i + y_j)p_{ij} = \sum_{i\in I}\sum_{j\in J} x_i p_{ij} + \sum_{j\in J}\sum_{i\in I} y_j p_{ij} =$$

$$= \sum_{i\in I}\left(x_i \sum_{j\in J} p_{ij}\right) + \sum_{j\in J}\left(y_j \sum_{i\in I} p_{ij}\right) = \sum_{i\in I} x_i p_i + \sum_{j\in J} y_j p'_j = \mathbb{E}(X) + \mathbb{E}(Y).$$

3. Variabila aleatoare XY are distribuţia

$$XY : \begin{pmatrix} x_i y_j \\ p_{ij} \end{pmatrix}_{(i,j)\in I\times J}.$$

Dacă X şi Y sunt independente, atunci $p_{ij} = p_i p'_j,\ \forall\ (i,j) \in I \times J$. Rezultă

$$\mathbb{E}(XY) = \sum_{(i,j)\in I\times J} (x_i y_j)p_{ij} = \sum_{(i,j)\in I\times J} (x_i p_i)(y_j p'_j) =$$

$$= \left(\sum_{i\in I} x_i p_i\right)\left(\sum_{j\in J} y_j p'_j\right) = \mathbb{E}(X)\mathbb{E}(Y).$$

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, o funcţie 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 2.7.2. *Fie $(\Omega, \mathcal{F}, \mathbb{P})$ un spaţiu de probabilitate şi $X : \Omega \to \mathbb{R}$ o variabilă aleatoare, având funcţia de repartiţie F.*

1. *Media de ordin $k \in \mathbb{N}^*$ a variabilei aleatoare X se defineşte prin*

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

în ipoteza că integrala este convergentă.

2. *Media absolută de ordin $k \in \mathbb{N}^*$ a variabilei aleatoare X se defineşte prin*

$$\mathbb{E}\left(|X|^k\right) = \int_{-\infty}^{\infty} |x|^k dF(x),$$

în ipoteza că integrala este convergentă.

Notăm $L^k(\Omega)$ mulţimea variabilelor aleatoare definite pe spaţiul de probabilitate $(\Omega, \mathcal{F}, \mathbb{P})$ cu medie absolută finită de ordinul k:

$$L^k(\Omega) = \{X : \Omega \to \mathbb{R} - \textit{variabilă aleatoare}: \ \mathbb{E}\left(|X|^k\right) < \infty\}.$$

Definim $L^\infty(\Omega) = \cap_{k=1}^{\infty} L^k(\Omega)$ mulţimea variabilelor aleatoare cu medie absolută finită de orice ordin.

Pentru oricare $k \in \mathbb{N}^*$, $L^k(\Omega)$ este un spaţiu vectorial real normat. Norma în $L^k(\Omega)$ se defineşte prin:

$$\|X\|_k = \left[\mathbb{E}\left(|X|^k\right)\right]^{1/k}, \ X \in L^k(\Omega).$$

Vom demonstra ulterior 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 $i = 1, \cdots, k$.

Vom explicita în continuare Definiţia 2.7.2 pentru variabile aleatoare discrete şi respectiv pentru variabile aleatoare de tip continuu.

I. **Variabile aleatoare discrete.**
Fie X o variabilă aleatoare discretă, cu distribuţia $X : \begin{pmatrix} x_i \\ p_i \end{pmatrix}_{i \in I}$. Atunci:

(a) $\mathbb{E}(X^k) = \sum_{i \in I} x_i^k p_i$;
dacă I este numărabilă, se presupune că seria $\sum_{i \in I} x_i^k p_i$ este convergentă;

(b) $\mathbb{E}(|X|^k) = \sum_{i \in I} |x_i|^k p_i$;
dacă I este numărabilă, se presupune că seria $\sum_{i \in I} |x_i|^k p_i$ este convergentă;

II. **Variabile aleatoare continue.**
Fie X o variabilă aleatoare de tip continuu, cu funcţia de densitate $f : \mathbb{R} \to [0, \infty)$. Atunci:

(a) $\mathbb{E}(X^k) = \int_{-\infty}^{\infty} x^k f(x)dx$, în ipoteza că integrala improprie este convergentă;

(b) $\mathbb{E}(|X|^k) = \int_{-\infty}^{\infty} |x|^k f(x)dx$, în ipoteza că integrala improprie este convergentă.

2.7.2 Mediile unor distribuţii clasice

Prezentăm în continuare mediile unor variabile aleatoare având distribuţii importante, discrete (Bernoulli, binomială, Poisson, geometrică) sau continue (normală, exponenţială, Gama).

1. **Distribuţia Bernoulli:** $X \sim B(1, p) \;\Rightarrow\; \mathbb{E}(X) = p$.

Demonstraţie. O variabila aleatoare $X \sim B(1, p)$, cu $p \in (0, 1)$ şi $q = 1 - p$, are distribuţia $X : \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}$. Atunci:

$$\mathbb{E}(X) = 0 \cdot q + 1 \cdot p = p. \;\square$$

2. **Distribuţia binomială:** $X \sim B(n, p) \;\Rightarrow\; \mathbb{E}(X) = np$.

Demonstraţie. O variabila aleatoare $X \sim B(n, p)$, cu $n \in \mathbb{N}^*$, $p \in (0, 1)$ şi $q = 1 - p$, are distribuţia $X : \begin{pmatrix} k \\ C_n^k p^k q^{n-k} \end{pmatrix}_{k \in \{0,1,\cdots,n\}}$. Atunci:

$$\mathbb{E}(X) = \sum_{k=0}^{n} k C_n^k p^k q^{n-k} = \sum_{k=1}^{n} n C_{n-1}^{k-1} p^k q^{n-k} \overset{i=k-1}{=} np \sum_{i=0}^{n-1} C_{n-1}^{i} p^i q^{(n-1)-i} = np(p+q)^{n-1} = np.$$

O metodă instructivă de calcul a mediei unei variabile aleatoare binomiale $X \sim B(n, p)$ se bazează pe faptul că X se poate reprezenta ca suma a n variabile aleatoare Bernoulli independente $X_i \sim B(1, p), i = 1, \cdots, n$. Astfel, avem

$$\mathbb{E}(X) = \mathbb{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{E}(X_i) = np. \;\square$$

3. **Distribuţia Poisson:** $X \sim P(\lambda) \;\Rightarrow\; \mathbb{E}(X) = \lambda$.

Demonstraţie. O variabila aleatoare $X \sim P(\lambda)$, cu $\lambda > 0$, are distribuţia următoare $X : \begin{pmatrix} n \\ \lambda^n e^{-\lambda}/n! \end{pmatrix}_{n \in \mathbb{N}}$. Avem:

$$\mathbb{E}(X) = \sum_{n=0}^{\infty} n \cdot \frac{\lambda^n e^{-\lambda}}{n!} = \lambda e^{-\lambda} \sum_{n=1}^{\infty} \frac{\lambda^{n-1}}{(n-1)!} \overset{i=n-1}{=} \lambda e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = \lambda e^{-\lambda} e^{\lambda} = \lambda. \;\square$$

4. **Distribuţia geometrică:** $X \sim G(p) \;\Rightarrow\; \mathbb{E}(X) = \dfrac{1}{p}$.

Demonstraţie. O variabila aleatoare $X \sim G(p)$, cu $p \in (0,1)$ şi $q = 1-p$, are distribuţia $X : \begin{pmatrix} n \\ pq^{n-1} \end{pmatrix}_{n\in\mathbb{N}^*}$. Avem:

$$\mathbb{E}(X) = \sum_{n=1}^{\infty} npq^{n-1} = p\sum_{n=1}^{\infty} nq^{n-1} =$$

$$= p\sum_{n=0}^{\infty} (q^n)'_q = p\left(\sum_{n=0}^{\infty} q^n\right)'_q = p\left(\frac{1}{1-q}\right)'_q = \frac{p}{(1-q)^2} = \frac{1}{p}. \;\square$$

5. **Distribuţia normală:** $X \sim N(\mu, \sigma^2) \;\Rightarrow\; \mathbb{E}(X) = \mu$.

Demonstraţie. O variabila aleatoare $X \sim N(\mu, \sigma^2)$ are funcţia de densitate

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

Avem:

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x\varphi(x)dx = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} xe^{-\frac{(x-\mu)^2}{2\sigma^2}}dx \overset{y=x-\mu}{=} \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}(y+\mu)e^{-\frac{y^2}{2\sigma^2}}dy =$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} ye^{-\frac{y^2}{2\sigma^2}}dy + \frac{\mu}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2\sigma^2}}dy \overset{impar./par.}{=} 0 + \frac{2\mu}{\sigma\sqrt{2\pi}}\int_{0}^{\infty} e^{-\frac{y^2}{2\sigma^2}}dy =$$

$$\overset{z=y^2/(2\sigma^2)}{=} \frac{2\mu}{\sigma\sqrt{2\pi}}\int_0^{\infty}\frac{\sigma\sqrt{2}}{2\sqrt{z}}e^{-z}dz = \frac{\mu}{\sqrt{\pi}}\int_0^{\infty} z^{1/2-1}e^{-z}dz = \frac{\mu}{\sqrt{\pi}}\Gamma(1/2) = \mu.$$

6. **Distribuţia exponenţială:** $X \sim Exp(a) \;\Rightarrow\; \mathbb{E}(X) = \dfrac{1}{a}$.

Demonstraţie. O variabila aleatoare pozitivă $X \sim Exp(a)$ are funcţia de densitate $f(x) = ae^{-ax},\ x \geq 0$. Obţinem

$$\mathbb{E}(X) = \int_0^{\infty} axe^{-ax}dx = -\int_0^{\infty} x\left(e^{-ax}\right)'dx = -\left. xe^{-ax}\right|_0^{\infty} + \int_0^{\infty} x'e^{-ax}dx = 0 - \left.\frac{e^{-ax}}{a}\right|_0^{\infty} = \frac{1}{a}.$$

7. **Distribuţia Gama:** $X \in \Gamma(n, a) \;\Rightarrow\; \mathbb{E}(X) = \dfrac{n}{a}$.

Demonstraţie. O variabila aleatoare pozitivă $X \in \Gamma(n,a)$, cu $n, a > 0$, are funcţia de densitate $f(x) = \dfrac{a^n}{\Gamma(n)}x^{n-1}e^{-ax},\ x \geq 0$. Avem:

$$\mathbb{E}(X) = \int_0^{\infty} xf(x)dx = \frac{a^n}{\Gamma(n)}\int_0^{\infty} x^n e^{-ax}dx \overset{y=ax}{=} \frac{a^n}{\Gamma(n)}\int_0^{\infty}\frac{y^n e^{-y}}{a^{n+1}}dy =$$

$$= \frac{1}{a\Gamma(n)}\int_0^{\infty} y^{(n+1)-1}e^{-y}dy = \frac{\Gamma(n+1)}{a\Gamma(n)} = \frac{n\Gamma(n)}{a\Gamma(n)} = \frac{n}{a}.$$

2.8 Dispersia

2.8.1 Definiţie, proprietăţi

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

Definiţia 2.8.1. *Fie $X \in L^2(\Omega)$, cu $\mathbb{E}(X) = \mu$. Mărimea:*

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

se numeşte dispersia lui X. Mărimea

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

se numeşte abaterea medie pătratică a variabilei aleatoare X.

Dispersia unei variabile aleatoare constante este egală cu 0. În mod explicit, pentru a variabilă aleatoare X, cu media $\mathbb{E}(X) = \mu$, dispersia se calculează astfel:

- dacă X este o variabilă aleatoare discretă cu valorile x_i, luate cu probabilităţile p_i, $i \in I$,
$$V(X) = \sum_{i\in I} (x_i - \mu)^2 \, p_i;$$
- dacă X este o variabilă aleatoare de tip continuu, cu funcţia de densitate f,
$$V(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx.$$

Dispersia are următoarele proprietăţi elementare.

Propoziţia 2.8.1. *Fie $X, Y \in L^2(\Omega)$, cu $\mathbb{E}(X) = \mu$ şi $\mathbb{E}(Y) = \eta$. Atunci:*

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

Demonstraţie.
1. Variabila aleatoare $(X-\mu)^2$ ia doar valori pozitive, deci $V(X) = \mathbb{E}\left[(X-\mu)^2\right] \geq 0$.
2. Fie $a \in \mathbb{R}$. Avem $\mathbb{E}(aX) = a\mu$. Atunci

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

3. Fie $a \in \mathbb{R}$. Avem $\mathbb{E}(X+a) = \mu + a$. Rezultă

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

4. Conform definiţiei:

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

5. Avem $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y) = \mu + \eta$. Deoarece variabilele aleatoare X şi Y sunt independente, media produsului XY este egală cu produsul mediilor: $\mathbb{E}(XY) = \mathbb{E}(X)\cdot\mathbb{E}(Y) = \mu\eta$. Rezultă

$$V(X+Y) = \mathbb{E}\left[((X+Y)-(\mu+\eta))^2\right] = \mathbb{E}\left(X^2 + 2XY + Y^2 - 2(\mu+\eta)(X+Y) + (\mu+\eta)^2\right) =$$
$$= \mathbb{E}\left(X^2\right) + 2\mathbb{E}(XY) + \mathbb{E}\left(Y^2\right) - 2(\mu+\eta)\mathbb{E}(X+Y) + (\mu+\eta)^2 =$$
$$= \mathbb{E}\left(X^2\right) + 2\mu\eta + \mathbb{E}\left(Y^2\right) - (\mu+\eta)^2 = \left[\mathbb{E}\left(X^2\right) - \mu^2\right] + \left[\mathbb{E}\left(Y^2\right) - \eta^2\right] \overset{4.}{=} V(X) + V(Y). \;\square$$

2.8.2 Dispersiile unor distribuţii clasice

Prezentăm dispersiile asociate unor distribuţii clasice, pentru care a fost evidenţiată anterior media. Vom utiliza, de regulă, formula $V(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X)$.

1. **Distribuţia Bernoulli:** $X \sim B(1,p) \;\Rightarrow\; V(X) = pq$.

Demonstraţie. O variabila aleatoare $X \sim B(1,p)$, având distribuţia $X : \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}$, cu $p \in (0,1),\; q = 1-p$, are media $\mathbb{E}(X) = p$. Calculăm $\mathbb{E}(X^2)$, unde $X^2 : \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}$. Avem $\mathbb{E}(X^2) = 0\cdot q + 1\cdot p = p$. Atunci $V(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X) = p - p^2 = pq$. $\square$

2. **Distribuţia binomială:** $X \sim B(n,p) \;\Rightarrow\; V(X) = npq$.

Demonstraţie. O variabila aleatoare $X \sim B(n,p)$, cu $n \in \mathbb{N}^*,\; p \in (0,1)$ şi $q = 1-p$, are distribuţia $X : \begin{pmatrix} k \\ C_n^k p^k q^{n-k} \end{pmatrix}_{k\in\{0,1,\cdots,n\}}$. Apelăm la reprezentarea variabilei aleatoare X ca suma a n variabile aleatoare Bernoulli independente $X_i \sim B(1,p), i = 1,\cdots,n$. Atunci

$$V(X) = V\left(\sum_{i=1}^{n} X_i\right) \overset{indep.}{=} \sum_{i=1}^{n} V(X_i) = npq. \;\square$$

3. **Distribuţia Poisson:** $X \sim P(\lambda) \;\Rightarrow\; V(X) = \lambda$.

Demonstraţie. O variabila aleatoare $X \sim P(\lambda)$, cu $\lambda > 0$, având distribuţia

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

are media $\mathbb{E}(X) = \lambda$. Calculăm $\mathbb{E}(X^2)$:

$$\mathbb{E}(X^2) = \sum_{n=0}^{\infty} n^2 \frac{\lambda^n e^{-\lambda}}{n!} = \sum_{n=2}^{\infty} n(n-1)\frac{\lambda^n e^{-\lambda}}{n!} + \sum_{n=1}^{\infty} n\frac{\lambda^n e^{-\lambda}}{n!} =$$

$$= \lambda^2 e^{-\lambda} \sum_{n=2}^{\infty} \frac{\lambda^{n-2}}{(n-2)!} + \mathbb{E}(X) \overset{i=n-2}{=} \lambda^2 e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} + \lambda = \lambda^2 e^{-\lambda} e^{\lambda} + \lambda = \lambda^2 + \lambda.$$

Atunci

$$V(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X) = (\lambda^2 + \lambda) - \lambda^2 = \lambda. \ \square$$

4. **Distribuţia geometrică:** $X \sim G(p) \ \Rightarrow \ V(X) = \dfrac{q}{p^2}$.

Demonstraţie. O variabila aleatoare $X \sim G(p)$, cu $p \in (0,1)$ şi $q = 1 - p$, având distribuţia $X : \begin{pmatrix} n \\ pq^{n-1} \end{pmatrix}_{n \in \mathbb{N}^*}$, are media $\mathbb{E}(X) = \dfrac{1}{p}$. Avem

$$\mathbb{E}(X^2) = \sum_{n=1}^{\infty} n^2 pq^{n-1} = \sum_{n=2}^{\infty} n(n-1)pq^{n-1} + \sum_{n=1}^{\infty} npq^{n-1} =$$

$$= pq \sum_{n=2}^{\infty} (q^n)'' + \mathbb{E}(X) = pq \left(\sum_{n=0}^{\infty} q^n \right)''_q + \frac{1}{p} =$$

$$= pq \left(\frac{1}{1-q} \right)''_q + \frac{1}{p} = \frac{2pq}{(1-q)^3} + \frac{1}{p} = \frac{2q+p}{p^2} = \frac{q+1}{p^2}.$$

Atunci $V(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X) = \dfrac{q+1}{p^2} - \dfrac{1}{p^2} = \dfrac{q}{p^2}$. $\square$

5. **Distribuţia normală:** $X \sim N(\mu, \sigma^2) \ \Rightarrow \ V(X) = \sigma^2$.

Demonstraţie. O variabila aleatoare $X \sim N(\mu, \sigma^2)$ are funcţia de densitate $\varphi(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $\forall \ x \in \mathbb{R}$, unde $\mu = \mathbb{E}(X) \in \mathbb{R}$ şi $\sigma > 0$. Calculăm $\mathbb{E}(X^2)$:

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 \varphi(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \overset{y=x-\mu}{=} \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (y+\mu)^2 e^{-\frac{y^2}{2\sigma^2}} dy =$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \left(\int_{-\infty}^{\infty} y^2 e^{-\frac{y^2}{2\sigma^2}} dy + 2\mu \int_{-\infty}^{\infty} y e^{-\frac{y^2}{2\sigma^2}} dy + \int_{-\infty}^{\infty} \mu^2 e^{-\frac{y^2}{2\sigma^2}} dy \right) =$$

$$\overset{par./impar.}{=} \frac{1}{\sigma\sqrt{2\pi}} \left(2 \int_{0}^{\infty} y^2 e^{-\frac{y^2}{2\sigma^2}} dy + 0 + 2\mu^2 \int_{0}^{\infty} e^{-\frac{y^2}{2\sigma^2}} dy \right) =$$

$$\overset{z=y^2/(2\sigma^2)}{=} \frac{2}{\sigma\sqrt{2\pi}} \left(\int_{0}^{\infty} \frac{\sigma^3\sqrt{2}z}{\sqrt{z}} e^{-z} dz + \mu^2 \int_{0}^{\infty} \frac{\sigma}{\sqrt{2z}} e^{-z} dz \right) =$$

$$= \frac{2\sigma^2}{\sqrt{\pi}} \int_{0}^{\infty} z^{3/2-1} e^{-z} dz + \frac{\mu^2}{\sqrt{\pi}} \int_{0}^{\infty} z^{1/2-1} e^{-z} dz = \frac{2\sigma^2}{\sqrt{\pi}} \cdot \Gamma(3/2) + \frac{\mu^2}{\sqrt{\pi}} \cdot \Gamma(1/2).$$

Avem $\Gamma(1/2) = \sqrt{\pi}$ şi $\Gamma(3/2) = 1/2\Gamma(1/2) = \sqrt{\pi}/2$. Rezultă

$$\mathbb{E}(X^2) = \frac{2\sigma^2}{\sqrt{\pi}} \cdot \frac{\sqrt{\pi}}{2} + \frac{\mu^2}{\sqrt{\pi}} \cdot \sqrt{\pi} = \sigma^2 + \mu^2.$$

Atunci

$$V(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X) = (\sigma^2 + \mu^2) - \mu^2 = \sigma^2.$$

În particular, dacă X are o distribuţie normală standard, deci $X \sim N(0,1)$, atunci avem $\mathbb{E}(X) = 0$ şi $V(X) = 1$. $\square$

6. **Distribuţia exponenţială:** $X \sim Exp(a) \;\Rightarrow\; V(X) = \dfrac{1}{a^2}$.

Demonstraţie. O variabila aleatoare pozitivă $X \sim Exp(a)$, cu $a > 0$, având funcţia de densitate $f(x) = ae^{-ax},\ x \geq 0$, are media $\mathbb{E}(X) = \dfrac{1}{a}$. Obţinem

$$\mathbb{E}(X^2) = \int_0^{\infty} ax^2 e^{-ax} dx = -\int_0^{\infty} x^2 \left(e^{-ax}\right)' dx = -\left. x^2 e^{-ax}\right|_0^{\infty} + \int_0^{\infty} (x^2)' e^{-ax} dx = \frac{2}{a}\mathbb{E}(X) = \frac{2}{a^2}.$$

Rezultă

$$V(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X) = \frac{2}{a^2} - \frac{1}{a^2} = \frac{1}{a^2}. \;\square$$

7. **Distribuţia Gama:** $X \in \Gamma(n,a) \;\Rightarrow\; V(X) = \dfrac{n}{a^2}$.

Demonstraţie. O variabila aleatoare pozitivă $X \in \Gamma(n,a)$, cu $n, a > 0$, are funcţia de densitate $f(x) = \dfrac{a^n}{\Gamma(n)} x^{n-1} e^{-ax},\ x \geq 0$ şi media $\mathbb{E}(X) = \dfrac{n}{a}$. Avem:

$$\mathbb{E}(X^2) = \int_0^{\infty} x^2 f(x) dx = \frac{a^n}{\Gamma(n)} \int_0^{\infty} x^{n+1} e^{-ax} dx \overset{y=ax}{=} \frac{a^n}{\Gamma(n)} \int_0^{\infty} \frac{y^{n+1} e^{-y}}{a^{n+2}} dy =$$

$$= \frac{1}{a^2 \Gamma(n)} \int_0^{\infty} y^{(n+2)-1} e^{-y} dy = \frac{\Gamma(n+2)}{a^2 \Gamma(n)} = \frac{n(n+1)\Gamma(n)}{a^2 \Gamma(n)} = \frac{n(n+1)}{a^2}.$$

Atunci

$$V(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X) = \frac{n(n+1)}{a^2} - \left(\frac{n}{a}\right)^2 = \frac{n}{a^2}. \;\square$$

2.9 Corelaţia

Corelaţia este un indicator al gradului de dependenţă dintre două variabile aleatoare.

Definiţia 2.9.1. *Fie $X, Y \in L^2(\Omega)$. Corelaţia variabilelor aleatoare X şi Y se defineşte prin:*

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

Presupunem $\sigma(X), \sigma(Y) > 0$. 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 alte proprietăţi ale corelaţiei.

Propoziţia 2.9.1. *Fie* $X, Y \in L^2(\Omega)$, *cu mediile* $\mathbb{E}(X) = \mu$ *şi* $\mathbb{E}(Y) = \eta$. *Atunci:*

1. $\lambda(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}(XY) - \mu\eta$;

2. *dacă* X *şi* Y *sunt independente, atunci* $\lambda(X,Y) = 0$*;*

3. $\rho(X,Y) \in [-1,1]$*;*

4. $\lambda(X,X) = V(X)$; $\rho(X,X) = 1$*;*

5. $\lambda(X,-X) = -V(X)$; $\rho(X,X) = -1$.

Demonstraţie.
1. Conform definiţiei,

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

2. Dacă X şi Y sunt independente, atunci $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y) = \mu\eta$, deci $\lambda(X,Y) = 0$.
3. 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

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

Atunci $\lambda^2(X,Y) \leq V(X)V(Y) = \sigma^2(X)\sigma^2(Y)$, de unde $|\lambda(X,Y)| \leq \sigma(X)\sigma(Y)$. Rezultă

$$|\rho(X,Y)| = \frac{|\lambda(X,Y)|}{\sigma(X)\sigma(Y)} \leq 1,$$

deci $\rho(X,Y) \in [-1,1]$.
4. Avem $\lambda(X,X) = \mathbb{E}(X^2) - \mathbb{E}^2(X) = V(X)$. Atunci $\rho(X,X) = \dfrac{V(X)}{\sigma^2(X)} = 1$.
5. Constatăm că $\lambda(X,-X) = \mathbb{E}(-X^2) - \mathbb{E}(X)\mathbb{E}(-X) = -\mathbb{E}(X^2) + \mathbb{E}^2(X) = -V(X)$ şi $V(X) = V(-X) = \sigma^2(X)$, de unde $\rho(X,-X) = \dfrac{-V(X)}{\sigma^2(X)} = -1$. □

Coeficientul de corelaţie normalizează corelaţia a două variabile aleatoare.

Dacă $\rho(X,Y) > 0$, atunci variabilele aleatoare X şi Y sunt *direct corelate*. Dacă $\rho(X,Y)$ este apropiat de 1, atunci între variabilele aleatoare X şi Y există o puternică dependenţă de tip direct. Dacă $\rho(X,Y) < 0$, atunci variabilele aleatoare X şi Y sunt *invers corelate*. Dacă $\rho(X,Y)$ este apropiat de -1, atunci între variabilele aleatoare X şi Y există o puternică dependenţă de tip invers.

2.10 Funcţia caracteristică

2.10.1 Definiţie, proprietăţi

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 2.10.1. *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.

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

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

Prezentăm în continuare principalele proprietăţi ale funcţiei caracteristice.

Teorema 2.10.1. *Fie X o variabilă aleatoare, cu funcţia de repartiţie F. Funcţia caracteristică $\varphi_X : \mathbb{R} \to \mathbb{C}$ a variabilei aleatoare X are următoarele proprietăţi:*

1. $\varphi_X(0) = 1$;
2. $\varphi_X(-t) = \overline{\varphi_X(t)},\ \forall\ t \in \mathbb{R}$;
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 $X \in L^\infty(\Omega)$, atunci φ_X este dezvoltabilă în serie Taylor în jurul originii:*

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

cu $\varphi_X^{(n)}(0) = i^n \mathbb{E}\left(X^n\right),\ n \in \mathbb{N}$. Astfel, are loc formula:

$$\varphi_X(t) = \sum_{n=0}^{\infty} \frac{(it)^n}{n!} \mathbb{E}(X^n),\ t \in \mathbb{R}. \tag{2.2}$$

Demonstrație.
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. $|\varphi_X(t)| = \left|\int_{-\infty}^{\infty} e^{itx} dF(x)\right| \leq \int_{-\infty}^{\infty} \left|e^{itx}\right| dF(x) = \int_{-\infty}^{\infty} 1 dF(x) = 1,\ \forall\ t \in \mathbb{R}.$
4. Fie $a, b \in \mathbb{R}$. Atunci:

$$\varphi_{aX+b}(t) = \int_{-\infty}^{\infty} e^{it(ax+b)} dF(x) = e^{ibt} \int_{-\infty}^{\infty} e^{i(at)x} dF(x) = e^{ibt} \varphi_X(at).$$

5. Fie $t \in \mathbb{R}$ şi $h > 0$. Pentru $x \in \mathbb{R}$, avem:

$$e^{ihx} - 1 = \cos(hx) + i\sin(hx) - 1 = -2\sin^2\frac{hx}{2} + 2i\sin\frac{hx}{2}\cos\frac{hx}{2} = 2i\sin\frac{hx}{2}\left(\cos\frac{hx}{2} + i\sin\frac{hx}{2}\right),$$

de unde

$$\left|e^{ihx} - 1\right| = \left|2i\sin\frac{hx}{2}\left(\cos\frac{hx}{2} + i\sin\frac{hx}{2}\right)\right| = 2\left|\sin\frac{hx}{2}\right| \leq 2\left|\frac{hx}{2}\right| = h|x|.$$

Rezultă

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

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

Fie $\varepsilon > 0$. Deoarece $\int_{-\infty}^{\infty} dF(x) = 1$, există $a > 0$ astfel ca $\int_{(-\infty,-a]\cup[a,\infty)} dF(x) < \frac{\varepsilon}{4}$. Rezultă

$$2\int_{(-\infty,-a]\cup[a,\infty)} \left|\sin\frac{hx}{2}\right| dF(x) \leq 2\int_{(-\infty,-a]\cup[a,\infty)} 1\, dF(x) < \frac{\varepsilon}{2},\ \forall h > 0.$$

Apoi, pentru $0 < h < \frac{\varepsilon}{2a}$, avem

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

Din inegalităţile de mai sus, obţinem:

$$|\varphi_X(t+h) - \varphi_X(t)| \leq 2\int_{-\infty}^{\infty} \left|\sin\frac{hx}{2}\right| dF(x) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,\ \forall t \in \mathbb{R},\ \forall\ h \in \left(0, \frac{\varepsilon}{2a}\right).$$

Ca urmare, φ_X este uniform continuă pe $\mathbb{R}$, deci şi continuă pe $\mathbb{R}$.
6. Derivata formală de ordinul n a funcţiei caracteristice este

$$\varphi_X^{(n)}(t) = \frac{\mathrm{d}^n}{\mathrm{d}t^n}\left(\int_{-\infty}^{\infty} e^{itx} dF(x)\right) = \int_{-\infty}^{\infty} \frac{\partial^n}{\partial t^n}\left(e^{itx}\right) dF(x) = i^n \int_{-\infty}^{\infty} x^n e^{itx} dF(x),\ t \in \mathbb{R}.$$

Rămâne de arătat că integrala $\int_{-\infty}^{\infty} x^n e^{itx} dF(x)$ este convergentă. Conform ipotezei,

$$\int_{-\infty}^{\infty} \left|x^n e^{itx}\right| dF(x) = \int_{-\infty}^{\infty} |x|^n \left|e^{itx}\right| dF(x) = \int_{-\infty}^{\infty} |x|^n \, dF(x) = \mathbb{E}(|X|^n) < \infty.$$

Din criteriul de convergenţă absolută, rezultă că integrala complexă $\int_{-\infty}^{\infty} x^n e^{itx} dF(x)$ este convergentă, deci formula de derivare este aplicabilă.
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 complexă, 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 $t \in \mathbb{R}$. Reţinem că dezvoltatea obţinută poate fi exprimată alternativ prin

$$\varphi_X(t) = \sum_{n=0}^{\infty} \frac{(it)^n}{n!} \mathbb{E}(X^n), \ t \in \mathbb{R}. \ \square$$

Teorema 2.10.2. *(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 2.10.3. *(Teorema de unicitate) Fie X o variabilă aleatoare cu funcţia de repartiţie F şi funcţia caracteristică φ_X. Atunci*

1. $F\left(x_2\right) - F\left(x_1\right) = \mathbb{P}\left\{x_1 < X \le x_2\right\} = \dfrac{1}{2\pi} \lim\limits_{c\to\infty} \displaystyle\int_{-c}^{c} \frac{e^{-itx_1} - e^{-itx_2}}{it} \varphi_X(t)\, dt, \ \forall \ x_1 < x_2;$

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

3. *Dacă X este de tip continuu, cu funcţia de densitate f, atunci:*
   $f(x) = \dfrac{1}{2\pi} \displaystyle\int_{-\infty}^{\infty} \varphi_X(t) e^{-itx} dt, \ \forall \ x \in \mathbb{R}.$

Demonstraţie. Se aplică transformarea Fourier inversă. $\square$

2.10.2 Funcţiile caracteristice ale unor distribuţii clasice.

1. Distribuţia Bernoulli: $X \sim B(1,p) \;\Rightarrow\; \varphi_X(t) = q + pe^{it},\ t \in \mathbb{R},\ q = 1-p.$

   Demonstraţie. $\varphi_X(t) = e^{it\cdot 0}q + e^{it\cdot 1}p = q + pe^{it},\ t \in \mathbb{R}.$
   Consecinţă. Prin dezvoltare în serie Taylor, obţinem:

   $$\varphi_X(t) = q + pe^{it} = 1 + \sum_{n=1}^{\infty} \frac{(it)^n}{n!}p,\ t \in \mathbb{R},$$

   deci $\mathbb{E}(X^n) = p,\ \forall\, n \in \mathbb{N}^*.$

2. Distribuţia binomială: $X \sim B(n,p) \;\Rightarrow\; \varphi_X(t) = \left(q + pe^{it}\right)^n,\ t \in \mathbb{R}.$

   Demonstraţie. Există n variabile aleatoare Bernoulli independente, de parametru p, $X_1, \cdots, X_n$, astfel ca $X = \sum_{k=1}^{n} X_k$. Atunci

   $$\varphi_X(t) = \varphi_{\sum_{k=1}^n X_k}(t) = \prod_{k=1}^{n} \varphi_{X_k}(t) = \left(q + pe^{it}\right)^n,\ t \in \mathbb{R}.$$

   Observaţie. Prin dezvoltarea binomului, urmată de dezvotarea în serie Taylor a exponenţialei, obţinem:

   $$\varphi_X(t) = \sum_{k=0}^{n} C_n^k q^{n-k}\left(pe^{it}\right)^k = \sum_{k=0}^{n} C_n^k q^{n-k} p^k \sum_{m=0}^{\infty} \frac{(itk)^m}{m!} = \sum_{m=0}^{\infty}\left(\sum_{k=0}^{n} k^m C_n^k q^{n-k} p^k\right)\frac{(it)^m}{m!},$$

   de unde deducem $\mathbb{E}(X^m) = \sum_{k=0}^{n} k^m C_n^k q^{n-k} p^k = \sum_{k=0}^{n} k^m \mathbb{P}\{X = k\},\ m \in \mathbb{N}^*.$
   Rezultatul poate fi obţinut direct din definiţia mediei de ordin superior a variabilei aleatoare binomiale X.

3. Distribuţia Poisson: $X \sim P(\lambda) \;\Rightarrow\; \varphi_X(t) = e^{\lambda\left(e^{it}-1\right)},\ t \in \mathbb{R}.$

   Demonstraţie. Pentru orice $t \in \mathbb{R}$, avem:

   $$\varphi_X(t) = \sum_{n=0}^{\infty} e^{itn}\mathbb{P}\{X = n\} = e^{-\lambda}\sum_{n=0}^{\infty} \frac{\lambda^n}{n!}e^{itn} = e^{-\lambda}\sum_{n=0}^{\infty} \frac{\left(\lambda e^{it}\right)^n}{n!} = e^{-\lambda}e^{\lambda e^{it}} = e^{\lambda\left(e^{it}-1\right)}.$$

4. Distribuţia geometrică: $X \sim G(p) \;\Rightarrow\; \varphi_X(t) = \dfrac{pe^{it}}{1-qe^{it}},\ t \in \mathbb{R},\ q = 1-p.$

   Demonstraţie. Pentru orice $t \in \mathbb{R}$, avem:

   $$\varphi_X(t) = \sum_{n=1}^{\infty} e^{itn}\mathbb{P}\{X = n\} = pe^{it}\sum_{n=1}^{\infty} e^{it(n-1)}q^{n-1} \overset{k=n-1}{=} pe^{it}\sum_{k=0}^{\infty}\left(qe^{it}\right)^k = \frac{pe^{it}}{1-qe^{it}}.$$

Aplicaţie. Derivând funcţia caracteristică, găsim $\varphi'_X(t) = \dfrac{ipe^{it}}{(1-qe^{it})^2}$, $t \in \mathbb{R}$. Aplicând Teorema 2.10.1, punctul 7, obţinem $\mathbb{E}(X) = \dfrac{\varphi'_X(0)}{i} = \dfrac{p}{(1-q)^2} = \dfrac{1}{p}$. Astfel, regăsim rezultatul obţinut anterior prin calculul direct al mediei.

5. Distribuţia normală: $X \sim N(\mu, \sigma^2) \;\Rightarrow\; \varphi_X(t) = e^{i\mu t - \frac{t^2\sigma^2}{2}}$, $t \in \mathbb{R}$.
   Distribuţia normală standard: $X \sim N(0,1) \;\Rightarrow\; \varphi_X(t) = e^{-\frac{t^2}{2}}$, $t \in \mathbb{R}$.

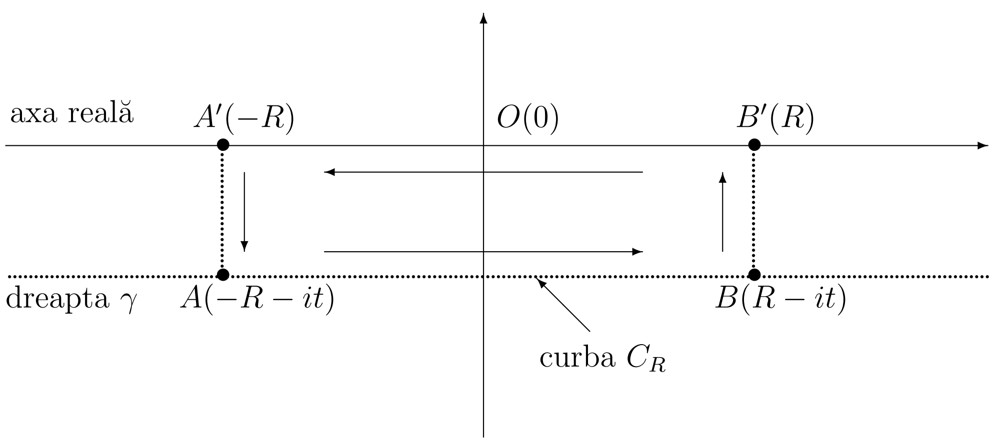
Demonstraţie. Vom trata doar cazul distribuţiei normale standard. Fie $X \sim N(0,1)$. Pentru $t \in \mathbb{R}^*$, avem

$$\varphi_X(t) = \mathbb{E}\left(e^{itX}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} e^{-\frac{x^2}{2}}\, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-it)^2}{2} - \frac{t^2}{2}}\, dx = \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} \int_{\gamma} e^{-\frac{z^2}{2}}\, dz,$$

unde γ este dreapta de ecuaţie $z = x - it$, $x \in \mathbb{R}$ din planul complex. Rămâne să demonstrăm relaţia

$$\int_{\gamma} e^{-\frac{z^2}{2}}\, dz = \sqrt{2\pi}. \tag{2.3}$$

Fie $R > 0$. Considerăm în planul complex punctele $A(-R-it)$, $B(R-it)$, $A'(-R)$ şi $B'(R)$. Definim curba închisă $C_R:\; A \to B \to B' \to A' \to A$, obţinută prin juxtapunerea segmentelor orientate $\overline{AB}$, $\overline{BB'}$, $\overline{B'A'}$ şi $\overline{A'A}$. Curba C_R este parcursă în sens pozitiv pentru $t > 0$, respectiv în sens negativ pentru $t < 0$.



Conform Teoremei lui Cauchy,

$$\int_{C_R} e^{-\frac{z^2}{2}}\, dz = 0.$$

Rezultă

$$\int_{\overline{AB}} e^{-\frac{z^2}{2}}\, dz + \int_{\overline{BB'}} e^{-\frac{z^2}{2}}\, dz + \int_{\overline{B'A'}} e^{-\frac{z^2}{2}}\, dz + \int_{\overline{A'A}} e^{-\frac{z^2}{2}}\, dz = 0,$$

sau

$$\int_{\overline{AB}} e^{-\frac{z^2}{2}}dz = -\int_{\overline{BB'}} e^{-\frac{z^2}{2}}dz + \int_{\overline{A'B'}} e^{-\frac{z^2}{2}}dz + \int_{\overline{AA'}} e^{-\frac{z^2}{2}}dz. \tag{2.4}$$

Au loc limitele de mai jos.

$$\lim_{R\to\infty}\int_{\overline{AB}} e^{-\frac{z^2}{2}}dz = \lim_{R\to\infty}\int_{-R}^{R} e^{-\frac{(x-it)^2}{2}}dx = \int_{-\infty}^{\infty} e^{-\frac{(x-it)^2}{2}}dx = \int_{\gamma} e^{-\frac{z^2}{2}}dz.$$

şi

$$\lim_{R\to\infty}\int_{\overline{A'B'}} e^{-\frac{z^2}{2}}dz = \lim_{R\to\infty}\int_{-R}^{R} e^{-\frac{x^2}{2}}dx = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}dx = \sqrt{2\pi}.$$

Segmentul orientat $\overline{BB'}$ admite parametrizarea $z = R + (y-t)i,\ y \in [0,t]$. Atunci

$$\left|\int_{\overline{BB'}} e^{-\frac{z^2}{2}}dz\right| = \left|\int_0^t e^{-\frac{[R+(y-t)i]^2}{2}} i dy\right| \le$$

$$\le \int_0^t \left| ie^{-\frac{R^2-(t-y)^2}{2}}[\cos R(t-y) + i\sin R(t-y)]\right| dy = e^{-\frac{R^2}{2}}\int_0^t e^{\frac{(t-y)^2}{2}}dy \le \frac{te^{\frac{t^2}{2}}}{e^{\frac{R^2}{2}}}.$$

Din $\displaystyle\lim_{R\to\infty}\frac{te^{\frac{t^2}{2}}}{e^{\frac{R^2}{2}}} = 0$ rezultă $\displaystyle\lim_{R\to\infty}\int_{\overline{BB'}} e^{-\frac{z^2}{2}}dz = 0$. Analog, $\displaystyle\lim_{R\to\infty}\int_{\overline{AA'}} e^{-\frac{z^2}{2}}dz = 0$.

Ca urmare, trecând la limită cu $R \to \infty$ în relaţia (2.4), obţinem (2.3). Astfel, am demonstrat $\varphi_X(t) = e^{-\frac{t^2}{2}},\ \forall\, t \in \mathbb{R}^*$. În plus, pentru $t = 0$, avem $\varphi_X(0) = 1 = e^{-\frac{0^2}{2}}$.
Aplicaţie. Din dezvoltatea

$$\varphi_X(t) = e^{-\frac{t^2}{2}} = \sum_{n=0}^{\infty}\frac{(-t^2/2)^n}{n!} = \sum_{n=0}^{\infty}\frac{(it)^{2n}}{n!2^n} = \sum_{n=0}^{\infty}\frac{(it)^{2n}}{(2n)!}\cdot\frac{(2n)!}{n!2^n}$$

rezultă

$$\mathbb{E}(X^k) = \begin{cases} 0, & k = 2n-1 \\ \frac{(2n)!}{n!2^n}, & k = 2n \end{cases},\ n \in \mathbb{N}^*.$$

6. Distribuţia exponenţială: $X \sim Exp(a) \ \Rightarrow\ \varphi_X(t) = \dfrac{a}{a-it},\ t \in \mathbb{R}$.

Demonstraţie.

$$\varphi_X(t) = \int_0^\infty ae^{itx}e^{-ax}dx = a\int_0^\infty e^{x(it-a)}dx = \frac{a}{it-a}e^{x(it-a)}\Big|_0^\infty = \frac{a}{a-it},\ t \in \mathbb{R}.$$

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

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

2.12 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 $\mathbb{P}\left\{|X| \leq \frac{1}{2}\right\}$.

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

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

   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:

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

Capitolul 3

Convergenţa şirurilor de variabile aleatoare

3.1 Introducere

Studiul convergenţei 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

Vom prezenta o serie de inegalităţi care joacă un rol important în teoria probabilităţilor.

Propoziţia 3.2.1. *Pentru $X \in L^n(\Omega)$, unde $n \in \mathbb{N}^*$, are loc inegalitatea:*

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

Dacă n este număr par, atunci avem egalitate.

Demonstraţie. Fie F funcţia de repartiţie a variabilei aleatoare $X \in L^n(\Omega)$. Pe baza unei proprietăţi cunoscute a integralei Stieltjes-Riemann în raport cu o funcţie crescătoare, obţinem:

$$|\mathbb{E}(X^n)| = \left| \int_{-\infty}^{\infty} x^n \, dF(x) \right| \leq \int_{-\infty}^{\infty} |x|^n \, dF(x) = \mathbb{E}(|X|^n).$$

Dacă n este număr par, atunci

$$|\mathbb{E}(X^n)| = \mathbb{E}(X^n) = \mathbb{E}(|X|^n). \quad \square$$

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

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

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

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

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

$$= a^n(\mathbb{P}\{X \leq -a\} + \mathbb{P}\{X \geq a\}) = a^n \mathbb{P}\{|X| \geq a\},$$

de unde obţinem concluzia. □

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*

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

Demonstraţie. Aplicăm inegalitatea lui Markov variabilei aleatoare $Z = X - \mu$, pentru $n = 2$. Astfel, avem:

$$\mathbb{P}\{|X - \mu| \geq a\} = \mathbb{P}\{|Z| \geq a\} \leq \frac{\mathbb{E}\left(|Z|^2\right)}{a^2} = \frac{\mathbb{E}\left[(X-\mu)^2\right]}{a^2} = \frac{V(X)}{a^2} = \frac{\sigma^2}{a^2}. \ \square$$

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*

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

În particular, pentru $Y = 1$, avem

$$\mathbb{E}(|X|) \leq \left[\mathbb{E}\left(|X|^p\right)\right]^{\frac{1}{p}}, \ \forall\ p > 1. \tag{3.1}$$

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

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

Inegalitatea Cauchy-Buniakovski se interpretează prin faptul că dispersia variabilei aleatoare $X \in L^2(\Omega)$, definită prin $V(X) = \mathbb{E}\left[(X - \mathbb{E}(X))^2\right] = \mathbb{E}(X^2) - \mathbb{E}^2(X)$, este pozitivă. Astfel, conform rezultatelor de mai sus, avem $|\mathbb{E}(X)| \leq \mathbb{E}(|X|) \leq \sqrt{\mathbb{E}\left(X^2\right)}$, de unde, prin ridicare la pătrat, obţinem: $\mathbb{E}^2(X) \leq \mathbb{E}\left(X^2\right)$, deci $V(X) \geq 0$. Menţionăm de asemenea că inegalitatea este expresia convexităţii funcţiei $f(x) = x^2,\ x \in \mathbb{R}$.

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 $p = \dfrac{s}{r} > 1$ şi $q = \dfrac{p}{p-1} > 1$. Avem $\dfrac{1}{p} + \dfrac{1}{q} = 1$. Atunci, conform inegalităţii (3.1), obţinem

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

Ridicând la puterea $\dfrac{1}{r}$ inegalitatea obţinută mai sus, obţinem concluzia. □

Consecinţă.

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

Astfel, dacă $X \in L^{k+1}(\Omega)$, atunci $\mathbb{E}(|X|^{k+1}) < \infty$. Din inegalitatea lui Lyapunov, rezultă

$$\mathbb{E}(|X|^k) \leq \left[\mathbb{E}(|X|^{k+1})\right]^{\frac{k}{k+1}} < \infty,$$

deci $X \in L^k(\Omega)$.

Enunţăm, fară demonstraţie, următorul rezultat important.

Propoziţia 3.2.2. *Fie $(\Omega, \mathcal{F}, \mathbb{P})$ un spaţiu de probabilitate şi k un număr natural nenul. Mulţimea*

$$L^k(\Omega) = \{X : \Omega \to \mathbb{R} : \ X - \textit{variabilă aleatoare}, \ \mathbb{E}(|X|^k) < \infty\}$$

este un spaţiu vectorial normat, înzestrat cu norma $\|\cdot\|_k : L^k(\Omega) \to [0, \infty)$ definită prin

$$\|X\|_k = \left[\mathbb{E}\left(|X|^k\right)\right]^{\frac{1}{k}}, \ X \in L^k(\Omega).$$

Notăm prin $(L^k(\Omega), \|\cdot\|_k)$ spaţiul vectorial real normat al variabilelor aleatoare cu medie absolută de ordinul k finită.

Inegalitatea lui Lyapunov se interpretează prin:

$$\|X\|_r \leq \|X\|_s, \ \forall\, X \in L^\infty(\Omega) := \cap_{k=1}^\infty L^k(\Omega), \ \forall\, s, r \in \mathbb{N}^*, \ r < s.$$

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 spaţiul 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 spaţiul 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 $n \geq 1$,

$$\{|X_n - X| > \varepsilon\} = \{X - X_n < -\varepsilon\} \cup \{X - X_n > \varepsilon\} \supset \{X - X_n < -\varepsilon\} = \{X < X_n - \varepsilon\},$$

avem

$$\{X \leq x_0 - \varepsilon\} \setminus \{|X_n - X| > \varepsilon\} \subset \{X \leq x_0 - \varepsilon\} \setminus \{X < X_n - \varepsilon\} =$$
$$= \{X \leq x_0 - \varepsilon\} \cap \{X < X_n - \varepsilon\}^c = \{X \leq x_0 - \varepsilon\} \cap \{X_n - \varepsilon \leq X\} \subset \{X_n \leq x_0\}.$$

Atunci, conform monotoniei probabilităţii, obţinem

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

Pe baza ipotezei, $\lim_{n\to\infty} \mathbb{P}\{|X_n - X| > \varepsilon\} = 0$. Astfel, prin trecere la limită în relaţia de mai sus, găsim:

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

Cu un raţionament similar, deducem

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

de unde

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

Prin trecere la limită, rezultă

$$\limsup_{n\to\infty} F_n(x_0) \leq F(x_0 + \varepsilon). \tag{3.3}$$

Relaţiile (3.2) şi (3.3) conduc la inegalităţile:

$$F(x_0 - \varepsilon) \leq \liminf_{n\to\infty} F_n(x_0) \leq \limsup_{n\to\infty} F_n(x_0) \leq F(x_0 + \varepsilon), \ \forall \varepsilon > 0. \tag{3.4}$$

Deoarece F este continuă în x_0, avem

$$\lim_{\varepsilon \searrow 0} F(x_0 - \varepsilon) = \lim_{\varepsilon \searrow 0} F(x_0 - \varepsilon) = F(x_0).$$

Atunci, din (3.4) şi criteriul cleşte, rezultă că şirul $(F_n(x_0))_{n\geq 1}$ este convergent, cu

$$\lim_{n\to\infty} F_n(x_0) = F(x_0).$$

Punctul x_0 de continuitate al funcţiei F fiind arbitrar, deducem că are loc convergenţa

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

2. Presupunem $X_n \overset{a.s.}{\to} X$. Avem

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

Atunci, pentru $\varepsilon > 0$, arbitrar, are loc incluziunea

$$\{X_n \to X\} \subset \cup_{n=1}^{\infty} \cap_{p=0}^{\infty} \{|X_{n+p} - X| < \varepsilon\}.$$

Conform ipotezei, rezultă

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

de unde

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

Şirul de evenimente $\left(\cap_{p=0}^{\infty}\{|X_{n+p} - X| < \varepsilon\}\right)_{n\geq 1}$ este monoton crescător, cu limita

$$\cup_{n=1}^{\infty} \cap_{p=0}^{\infty} \{|X_{n+p} - X| < \varepsilon\}.$$

Ca urmare,

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

Cum $\cap_{p=0}^{\infty}\{|X_{n+p} - X| < \varepsilon\} \subset \{|X_n - X| < \varepsilon\},\ \forall\, n \in \mathbb{N}^*$, avem

$$\mathbb{P}\left(\cap_{p=0}^{\infty}\{|X_{n+p} - X| < \varepsilon\}\right) \leq \mathbb{P}\left(\{|X_n - X| < \varepsilon\}\right) \leq 1,\ \forall\, n \in \mathbb{N}^*.$$

Prin trecece la limită cu $n \to \infty$ şi criteriul cleşte, rezultă

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

Trecând la evenimentele contrare, obţinem

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

Rezultă $X_n \overset{p}{\to} X$.

3. Implicaţia rezultă din inegalitatea lui Markov. Astfel, pentru oricare $\varepsilon > 0$, avem

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

Aplicând criteriul cleşte, obţinem $\lim\limits_{n\to\infty} \mathbb{P}\{|X_n - X| > \varepsilon\} = 0,\ \forall \varepsilon > 0$. Ca urmare, $X_n \overset{p}{\to} X$.

Observaţie. Convergenţa $X_n \overset{L^k}{\to} X$ corespunde convergenţei şirului $(X_n)_{n\geq 1}$ la X în spaţiul normat $(L^k(\Omega), \|\cdot\|_k)$, echivalentă cu condiţia:

$$\lim_{n\to\infty} \|X_n - X\|_k = 0.$$

Remarcăm de asemenea implicaţia:

$$X_n \overset{L^{k+1}}{\to} X \;\Rightarrow\; X_n \overset{L^k}{\to} X,\ \forall\, k \in \mathbb{N}^*,$$

dedusă din inegalitatea lui Lyapunov.

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 teoretică de realizare a evenimentului în cadrul fiecarei experienţe.* 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 este următoarea. Presupunem că un eveniment A se poate produce în cadrul unei experienţe cu probabilitatea teoretică $p = \mathbb{P}(A) \in (0, 1)$. Realizăm un şir de experienţe consecutive, identice, independente. Asociem şirului de experienţe şirul $(X_n)_{n\geq 1}$ de variabile aleatoare Bernoulli, independente şi identic distribuite (iid), cu distribuţia comună

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

unde valoarea 1 indică realizarea evenimentului A, iar valoarea 0 indică nerealizarea evenimentului A, în cadrul experienţelor $1, 2, \cdots$. Variabila aleatoare $S_n = X_1+X_2+\cdots+X_n$ indică numărul de apariţii (realizări) ale evenimentului A în primele n experienţe. Frecvenţa de realizare a evenimentului A în primele n experienţe este egală cu $\dfrac{S_n}{n}$. Fenomenul descris de legea numerelor mari este prin urmare

$$\frac{S_n}{n} \to p.$$

Trebuie însă dat un sens probabilistic precis convergenţei descrisă anterior. Vom observa în primul rând că limita p este media comună a variabilelor aleatoare $X_n, \; n \geq 1$. *Legea numerelor mari* evidenţiază această convergenţă în sensul unui tip de convergenţă slabă (*legea slabă a numerelor mari*) sau în sensul unui tip de convergenţă tare (*legea tare a numerelor mari*).

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), din spaţiul $L^2(\Omega)$, cu $\mathbb{E}(X_1) = \mu$ şi $V(X_1) = \sigma^2$. Notăm $S_n = X_1 + X_2 + \cdots + X_n, \; n \in \mathbb{N}^*$. 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. Avem $\mathbb{E}(S_n) = \displaystyle\sum_{k=1}^{n} \mathbb{E}(X_k) = n\mu$ şi, pe baza independenţei, $V(S_n) = \displaystyle\sum_{k=1}^{n} V(X_k) = n\sigma^2$. Fie $\varepsilon > 0$, arbitrar. Conform inegalităţii lui Chebyshev,

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

Din $\lim_{n\to\infty} \frac{\sigma^2}{n\varepsilon^2} = 0$, 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. □

Legea slabă a numerelor poate fi formulată pentru şiruri de variabile aleatoare independente cu medie comună şi dispersii uniform marginite (Chebyshev), respectiv pentru siruri de variabile de variabile aleatoare cu medie comună, având proprietatea $V(S_n)/n^2 \to 0$ (Markov). Prezentăm, fără demonstraţie, 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 $L^4(\Omega)$, cu media comună μ. Fie $S_n = X_1 + X_2 + \cdots + X_n,\ n \in \mathbb{N}^*$. Atunci

$$\frac{S_n}{n} \overset{a.s.}{\to} \mu.$$

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 standard* (denumită şi *legea gaussiană*). Rezultatul admite numeroase extinderi, inclusiv 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 a generalizat 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.

Prezentăm la început rezultatele datorate lui de Moivre şi Laplace, care se referă la şiruri de variabile aleatoare Bernoulli, independente şi identic distribuite.

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 $p \in (0,1)$, 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:

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

Demonstraţie. Vom utiliza aproximarea factorialului prin

$$n! = \sqrt{2n\pi}\left(\frac{n}{e}\right)^n \theta_n, \tag{3.5}$$

unde $\lim\limits_{n\to\infty} \theta_n = 1$. Abraham de Moivre a propus în 1733 aproximarea asimptotică a factorialului prin $n! \sim cn^{n+1/2}e^{-n}$, iar James Stirling a determinat ulterior constanta $c = \sqrt{2\pi}$. Relaţia (3.5) este cunoscută sub denumirea de *formula (aproximarea) lui Stirling*.

Fie $(X_n)_{n\geq 1}$ un şir de variabile aleatoare Bernoulli, independente şi identic distribuite, cu distribuţia comună

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

Atunci $S_n := \sum_{k=1}^{n} X_k \sim B(n,p)$, deci S_n are distribuţia

$$S_n : \begin{pmatrix} 0 & 1 & \cdots & k & \cdots & n \\ q^n & C_n^1 pq^{n-1} & \cdots & C_n^k p^k q^{n-k} & \cdots & p^n \end{pmatrix}, \ n \in \mathbb{N}^*.$$

Considerăm variabilele aleatoare Z_n definite prin

$$Z_n = \frac{S_n - np}{\sqrt{npq}}, \ n \geq 1.$$

Evidenţiem caracteristicile variabilelor aleatoare $Z_n, \ n \geq 1$.

1. Media egală cu 0: $\mathbb{E}(Z_n) = \dfrac{1}{\sqrt{npq}}\mathbb{E}(S_n - np)\dfrac{1}{\sqrt{npq}}[n\mathbb{E}(X_1) - np] = 0.$

2. Dispersia egală cu 1: $V(Z_n) = \left(\dfrac{1}{\sqrt{npq}}\right)^2 [V(S_n - np)] = \dfrac{1}{npq}V(S_n) = \dfrac{nV(X_1)}{npq} = 1.$

Spunem că variabila aleatoare Z_n s-a obţinut prin *normalizarea variabilei aleatoare* S_n. Fie

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

mulţimea valorilor variabilei aleatoare Z_n. Pentru $n \in \mathbb{N}^*$, avem

$$a_n := \min I_n = \frac{-np}{\sqrt{npq}} = -\sqrt{\frac{np}{q}}$$

şi

$$b_n := \max I_n = \frac{n - np}{\sqrt{npq}} = \sqrt{\frac{nq}{p}}$$

Constatăm că $\lim\limits_{n\to\infty} a_n = -\infty$, iar $\lim\limits_{n\to\infty} b_n = \infty$. Considerăm intervalul compact $[a,b]$, fixat. Există $N \in \mathbb{N}$ astfel ca $a_n \leq a < b \leq b_n$ şi $\dfrac{1}{\sqrt{npq}} < b - a$. Rezultă

$$A_n := I_n \cap [a,b] \neq \emptyset, \ \forall\, n \in \mathbb{N}, \ n \geq N.$$

Fie $n \geq N$. Pentru $x = \dfrac{k - np}{\sqrt{npq}} \in A_n$, notăm $p_n(x) = \mathbb{P}\{Z_n = x\}$. Avem

$$p_n(x) = \mathbb{P}\left\{ \frac{S_n - np}{\sqrt{npq}} = x \right\} = \mathbb{P}\{S_n = k\} = C_n^k p^k q^{n-k} = \frac{n!}{k!j!}p^k q^j,$$

unde $j = n - k$. Utilizând formula lui Stirling (3.5), obţinem

$$p_n(x) = \frac{\sqrt{2n\pi}\,(n/e)^n\,\theta_n}{(\sqrt{2k\pi}\,(k/e)^k\,\theta_k)(\sqrt{2j\pi}\,(j/e)^j\,\theta_j)}p^k q^j = \frac{\theta_n}{\theta_k\theta_j\sqrt{2\pi}}\sqrt{\frac{n}{kj}}\left(\frac{np}{k}\right)^k\left(\frac{nq}{j}\right)^j.$$

Înmulţind relaţia cu $\sqrt{2\pi npq}$, obţinem

$$p_n(x)\sqrt{2\pi npq} = \frac{\theta_n}{\theta_k\theta_j}\sqrt{\frac{n^2pq}{kj}}\left(\frac{np}{k}\right)^k\left(\frac{nq}{j}\right)^j. \tag{3.6}$$

Definim numerele naturale $k_n = [np + a\sqrt{npq}]$ şi $j_n = [nq - b\sqrt{npq}]$. Din $k = np + x\sqrt{npq}$ şi $j = n - k = nq - x\sqrt{npq}$, cu $a \le x \le b$, rezultă $k \ge k_n$ şi $j \ge j_n$.
Deoarece $\lim\limits_{n\to\infty} k_n = \lim\limits_{n\to\infty} j_n = \infty$, obţinem $\lim\limits_{n\to\infty} k = \lim\limits_{n\to\infty} j = \infty$, de unde

$$\lim_{n\to\infty}\frac{\theta_n}{\theta_k\theta_j} = 1, \tag{3.7}$$

iar convergenţa de mai sus este uniformă pentru $x \in [a, b]$.
Cum $x \in [a, b]$, deducem $\lim\limits_{n\to\infty}\dfrac{x}{\sqrt{n}} = 0$. Ca urmare, au loc limitele

$$\lim_{n\to\infty}\frac{np}{k} = \lim_{n\to\infty}\frac{np}{np + x\sqrt{npq}} = \lim_{n\to\infty}\frac{p}{p + x\sqrt{pq}/\sqrt{n}} = 1;$$
$$\lim_{n\to\infty}\frac{nq}{j} = \lim_{n\to\infty}\frac{nq}{n - np - x\sqrt{npq}} = \lim_{n\to\infty}\frac{q}{q - x\sqrt{pq}/\sqrt{n}} = 1;$$
$$\lim_{n\to\infty}\sqrt{\frac{n^2pq}{kj}} = \lim_{n\to\infty}\sqrt{\frac{np}{k}}\cdot\lim_{n\to\infty}\sqrt{\frac{nq}{j}} = 1, \tag{3.8}$$

iar convergenţele sunt uniforme pentru $x \in [a, b]$.
Aplicăm formula lui Taylor cu rest Lagrange funcţiei $f(t) = \ln(1 + t),\ t > -1$. Pentru $t > -1,\ t \ne 0$, există $\xi_t \in \mathbb{R}^*,\ |\xi_t| < |t|$, astfel ca

$$f(t) = f(0) + \frac{f'(0)}{1!}t + \frac{f''(0)}{2!}t^2 + \frac{f'''(\xi_t)}{3!}t^3,$$

de unde

$$\ln(1 + t) = t - \frac{t^2}{2} + \alpha(t)\frac{t^3}{3},\ \text{ cu } \lim_{t\to 0}\alpha(t) = 1. \tag{3.9}$$

Aplicând (3.9), obţinem

$$\ln\left(\frac{np}{k}\right)^k = k\ln\left(\frac{np}{k}\right) = (np + x\sqrt{npq})\ln\left(1 - \frac{x\sqrt{npq}}{np + x\sqrt{npq}}\right) =$$

$$= (np + x\sqrt{npq})\left(-\frac{x\sqrt{npq}}{np + x\sqrt{npq}} - \frac{x^2npq}{2(np + x\sqrt{npq})^2} - \alpha(n, x)\frac{x^3(npq)^{3/2}}{(np + x\sqrt{npq})^3}\right) =$$

$$= -x\sqrt{npq} - \frac{x^2npq}{2(np + x\sqrt{npq})} - \alpha_1(n, x)\frac{x^3(npq)^{3/2}}{3(np + x\sqrt{npq})^2},$$

cu $\lim\limits_{n\to\infty} \alpha_1(n,x) = 1$. Analog obţinem

$$\ln\left(\frac{nq}{j}\right)^j = x\sqrt{npq} - \frac{x^2 npq}{2(nq - x\sqrt{npq})} + \alpha_2(n,x)\frac{x^3(npq)^{3/2}}{3(nq - x\sqrt{npq})^2},$$

cu $\lim\limits_{n\to\infty} \alpha_2(n,x) = 1$. Prin adunarea relaţiilor de mai sus, rezultă

$$\ln\left[\left(\frac{np}{k}\right)^k\left(\frac{nq}{j}\right)^j\right] = -\frac{x^2}{2}\left(\frac{npq}{np + x\sqrt{npq}} + \frac{npq}{nq - x\sqrt{npq}}\right) +$$

$$+\frac{x^3}{3}\left[\alpha_2(n,x)\frac{(npq)^{3/2}}{(nq - x\sqrt{npq})^2} - \alpha_1(n,x)\frac{(npq)^{3/2}}{(np + x\sqrt{npq})^2}\right].$$

Atunci

$$\lim_{n\to\infty} \ln\left[\left(\frac{np}{k}\right)^k\left(\frac{nq}{j}\right)^j\right] = -\frac{x^2}{2}(q+p) = -\frac{x^2}{2},$$

de unde

$$\lim_{n\to\infty}\left[\left(\frac{np}{k}\right)^k\left(\frac{nq}{j}\right)^j\right] = e^{-x^2/2}. \tag{3.10}$$

În plus, convergenţa (3.10) este uniformă datorită condiţiei $x \in [a,b]$.
Din relaţiile (3.6), (3.7), (3.8) şi (3.10) rezultă convergenţa uniformă

$$p_n(x)\sqrt{2\pi npq} \overset{n\to\infty}{\longrightarrow} e^{-\frac{x^2}{2}},\ \forall\, x \in A_n.$$

Atunci

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

Teorema de Moivre-Laplace conduce la versiunea binomială a *Teoremei limită centrală.*

Teorema 3.5.2. *Fie $(X_n)_{n\geq 1}$ un şir de variabile aleatoare Bernoulli independente şi identic distribuite, de parametru $p \in (0,1)$. Notăm $S_n = \sum_{k=1}^n X_k,\ n \in \mathbb{N}^*$.*
Pentru oricare interval $[a,b] \subset \mathbb{R}$, are loc limita:

$$\lim_{n\to\infty}\mathbb{P}\left\{a \leq \frac{S_n - np}{\sqrt{npq}} \leq b\right\} = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}\,dx.$$

Demonstraţie. Fie $[a,b] \subset \mathbb{R}$. Fie $\varepsilon > 0$. Din Teorema de Moivre-Laplace, rezultă că există $n_\varepsilon \in \mathbb{N}^*$, astfel ca

$$\left|\sqrt{2\pi npq}\cdot\mathbb{P}\left\{\frac{S_n - np}{\sqrt{npq}} = x\right\} - e^{-\frac{x^2}{2}}\right| < \varepsilon,\ \forall\, x \in A_n,\ \forall\, n \geq n_\varepsilon,$$

unde $A_n = \left\{\dfrac{k - np}{\sqrt{npq}},\ k = 0, 1, \cdots, n\right\} \cap [a,b]$.

Fie $n \geq n_\varepsilon$. Notăm $m_n = |A_n|$. Reprezentăm mulţimea A_n prin $A_n = \{x_1, x_2, \cdots, x_{m_n}\}$, unde $x_1 = \min A_n \in \left[a, a + \frac{1}{\sqrt{npq}}\right)$, $x_{m_n} = \max A_n \in \left(b - \frac{1}{\sqrt{npq}}, b\right]$, iar

$$x_{i+1} = x_i + \frac{1}{\sqrt{npq}}, \ i = 1, \cdots, m_n - 1.$$

Sumând inegalităţile

$$\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi npq}} - \frac{\varepsilon}{\sqrt{2\pi npq}} < \mathbb{P}\left\{\frac{S_n - np}{\sqrt{npq}} = x\right\} < \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi npq}} + \frac{\varepsilon}{\sqrt{2\pi npq}}, \ x \in A_n,$$

obţinem

$$\frac{1}{\sqrt{2\pi}}\left(\sum_{x \in A_n} \frac{1}{\sqrt{npq}} e^{-\frac{x^2}{2}} - \frac{\varepsilon m_n}{\sqrt{npq}}\right) < P_n < \frac{1}{\sqrt{2\pi}}\left(\sum_{x \in A_n} \frac{1}{\sqrt{npq}} e^{-\frac{x^2}{2}} + \frac{\varepsilon m_n}{\sqrt{npq}}\right), \qquad (3.11)$$

unde

$$P_n := \mathbb{P}\left\{a \leq \frac{S_n - np}{\sqrt{npq}} \leq b\right\}.$$

Definim diviziunea $\Delta_n = (a, x_1, x_2, \cdots, x_{m_n}, b)$ a intervalului $[a, b]$ şi sistemul de puncte intermediare $\Lambda_n = (x_1, x_2, \cdots, x_{m_n}, b)$ asociat lui Δ_n. Fie $f : [a, b] \to \mathbb{R}$, $f(x) = e^{-x^2/2}$. Suma Riemann a funcţiei f, asociată diviziunii Δ_n şi sistemului de puncte intermediare Λ_n este

$$\sigma_{\Delta_n}(f, \Lambda_n) = f(x_1)(x_1 - a) + \sum_{i=2}^{m_n} f(x_i)(x_i - x_{i-1}) + f(b)\,(b - x_{m_n}) =$$

$$= \frac{1}{\sqrt{npq}} \sum_{x \in A_n} e^{-\frac{x^2}{2}} - e^{-\frac{x_1^2}{2}}\left(a + \frac{1}{\sqrt{npq}} - x_1\right) + e^{-\frac{b^2}{2}}\,(b - x_{m_n}).$$

Prin trecere la limită, rezultă

$$\int_a^b e^{-\frac{x^2}{2}} dx = \lim_{n\to\infty}\left(\frac{1}{\sqrt{npq}} \sum_{x \in A_n} e^{-\frac{x^2}{2}}\right), \qquad (3.12)$$

deoarece

$$0 \leq e^{-\frac{x_1^2}{2}}\left(a + \frac{1}{\sqrt{npq}} - x_1\right) < \frac{1}{\sqrt{npq}}$$

şi

$$0 \leq e^{-\frac{b^2}{2}}\,(b - x_{m_n}) < \frac{1}{\sqrt{npq}},$$

ceea ce implică

$$\lim_{n\to\infty}\left[-e^{-\frac{x_1^2}{2}}\left(a + \frac{1}{\sqrt{npq}} - x_1\right) + e^{-\frac{b^2}{2}}\,(b - x_{m_n})\right] = 0.$$

Analog se arată

$$\lim_{n\to\infty} \frac{m_n}{\sqrt{npq}} = \int_a^b 1\,dx = b - a. \tag{3.13}$$

Atunci, prin trecere la limită în (3.11), pe baza relaţiilor (3.12) şi (3.13), rezultă:

$$\frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx - \frac{\varepsilon(b-a)}{\sqrt{2\pi}} \le \liminf_{n\to\infty} P_n \le \limsup_{n\to\infty} P_n \le \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx + \frac{\varepsilon(b-a)}{\sqrt{2\pi}}.$$

Cum $\varepsilon > 0$ este arbitrar, deducem că şirul $(P_n)_{n\ge1}$ converge la $\frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx$, deci are loc limita

$$\lim_{n\to\infty} \mathbb{P}\left\{a \le \frac{S_n - np}{\sqrt{npq}} \le b\right\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx. \ \square$$

Prezentăm în continuare forma generală a *Teoremei limită centrală.*

Teorema 3.5.3. *(Teorema limită centrală) Fie $(X_n)_{n\ge1}$ 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.
$$\lim_{n\to\infty} \mathbb{P}\left\{a \le \frac{S_n - n\mu}{\sigma\sqrt{n}} \le b\right\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx,\ \forall\ a, b \in \mathbb{R},\ a < b;$$

2. *(versiunea Lindeberg-Lévy) Are loc următoarea convergenţa în distribuţie către o variabilă aleatoare gaussiană:*
$$\frac{\sqrt{n}}{\sigma}\left(\frac{S_n}{n} - \mu\right) \xrightarrow{d} Z,$$
unde $Z \sim N(0, 1)$.

Demonstraţie.
Fie $(\Omega, \mathcal{F}, \mathbb{P})$ un spaţiu de probabilitate şi $(X_n)_{n\ge1}$ un şir de variabile aleatoare independente şi identic distribuite din $L^2(\Omega)$, cu $\mathbb{E}(X_n) = \mu$ şi $V(X_n) = \sigma^2$, pentru $n \ge 1$. Definim şirul de variabile aleatoare, independente şi identic distribuite, $Z_n = \frac{X_n - \mu}{\sigma},\ n \ge 1$, cu media $\mathbb{E}(Z_n) = \frac{1}{\sigma}(\mathbb{E}(X_n) - \mu) = 0$ şi dispersia $V(Z_n) = \frac{1}{\sigma^2}V(X_n - \mu) = \frac{1}{\sigma^2}V(X_n) = 1$, pentru orice $n \in \mathbb{N}^*$. Fie $Z \sim N(0, 1)$ o variabilă aleatoare normală standard definită pe $(\Omega, \mathcal{F}, \mathbb{P})$. Atunci

$$\frac{\sqrt{n}}{\sigma}\left(\frac{S_n}{n} - \mu\right) = \frac{1}{\sqrt{n}} \sum_{k=1}^n \frac{X_k - \mu}{\sigma} = \frac{1}{\sqrt{n}} \sum_{k=1}^n Z_k.$$

Pentru versiunea Lindeberg-Lévi a teoremei, avem de demonstrat

$$\frac{1}{\sqrt{n}} \sum_{k=1}^n Z_k \xrightarrow{d} Z. \tag{3.14}$$

Datorită injectivităţii transformării Fourier, convergenţa în distribuţie (3.14) este echivalentă cu convergenţa punctuală pe $\mathbb{R}$:

$$\varphi_{\frac{1}{\sqrt{n}}\sum_{k=1}^{n} Z_k} \to \varphi_Z, \quad n \to \infty. \tag{3.15}$$

Variabila aleatoare normală standard Z are funcţia caracteriatică:

$$\varphi_Z(t) = e^{-\frac{t^2}{2}}, \ t \in \mathbb{R}. \tag{3.16}$$

Notăm $\varphi : \mathbb{R} \to \mathbb{C}$ funcţia caracteristică comună a variabilelor aleatoare $Z_n, \ n \in \mathbb{N}^*$. Conform proprietăţilor funcţiei caracteristice (Teorema 2.9.1 şi Teorema 2.9.2), avem

$$\varphi_{\frac{1}{\sqrt{n}}\sum_{k=1}^{n} Z_k}(t) = \varphi_{\sum_{k=1}^{n} Z_k}\left(t/\sqrt{n}\right) \overset{indep.}{=} \prod_{k=1}^{n} \varphi_{Z_k}\left(t/\sqrt{n}\right) = \varphi^n\left(t/\sqrt{n}\right), \ t \in \mathbb{R}.$$

Astfel, convergenţa (3.15) se explicitează prin:

$$\lim_{n\to\infty} \varphi^n\left(\frac{t}{\sqrt{n}}\right) = e^{-\frac{t^2}{2}}, \ \forall\, t \in \mathbb{R}. \tag{3.17}$$

Pentru $t = 0$, avem $\varphi^n\left(0/\sqrt{n}\right) = 1 = e^{-\frac{0^2}{2}}, \ n \geq 1$.
Fie $t \in \mathbb{R} \setminus \{0\}$. Arătăm:

$$\Delta_n := \left|\varphi^n\left(\frac{t}{\sqrt{n}}\right) - e^{-\frac{t^2}{2}}\right| \overset{n\to\infty}{\longrightarrow} 0. \tag{3.18}$$

Pentru $n \geq 1$, avem

$$\Delta_n = \left|\varphi^n\left(\frac{t}{\sqrt{n}}\right) - \left(e^{-\frac{t^2}{2n}}\right)^n\right| = \left|\varphi\left(\frac{t}{\sqrt{n}}\right) - e^{-\frac{t^2}{2n}}\right| \cdot \left|\sum_{k=0}^{n-1} \varphi^k\left(\frac{t}{\sqrt{n}}\right)\left(e^{-\frac{t^2}{2n}}\right)^{n-1-k}\right|.$$

Deoarece $|\varphi(u)| \leq 1, \ \forall\, u \in \mathbb{R}$, iar $e^{-u} \in (0,1], \ \forall\, u \geq 0$, obţinem

$$\left|\sum_{k=0}^{n-1} \varphi^k\left(\frac{t}{\sqrt{n}}\right)\left(e^{-\frac{t^2}{2n}}\right)^{n-1-k}\right| \leq \sum_{k=0}^{n-1}\left|\varphi^k\left(\frac{t}{\sqrt{n}}\right)\right| e^{-\frac{t^2(n-1-k)}{2n}} \leq n,$$

de unde

$$\Delta_n \leq n\left|\varphi\left(\frac{t}{\sqrt{n}}\right) - e^{-\frac{t^2}{2n}}\right| \leq n\left|\varphi\left(\frac{t}{\sqrt{n}}\right) - 1 + \frac{t^2}{2n}\right| + n\left|1 - \frac{t^2}{2n} - e^{-\frac{t^2}{2n}}\right|. \tag{3.19}$$

Din formula lui Taylor cu rest Lagrange, rezultă

$$\left|e^{-u} - 1 + u\right| \leq \frac{u^2}{2}, \ \forall\, u \geq 0. \tag{3.20}$$

De asemenea, se obţin inegalităţile complexe:

$$\left|e^{iu} - 1 - iu\right| \leq \frac{u^2}{2}, \ \forall\, u \in \mathbb{R}; \tag{3.21}$$

$$\left|e^{iu} - 1 - iu - \frac{(iu)^2}{2}\right| \leq \frac{|u|^3}{6}, \ \forall\, u \in \mathbb{R}. \tag{3.22}$$

Conform (3.20),

$$\left|1-\frac{t^2}{2n}-e^{-\frac{t^2}{2n}}\right|\leq\frac{1}{2}\left(\frac{t^2}{2n}\right)^2=\frac{t^4}{8n^2}.$$

Atunci

$$n\left|1-\frac{t^2}{2n}-e^{-\frac{t^2}{2n}}\right|\leq\frac{t^4}{8n},\ \forall\, n\in\mathbb{N}^*,$$

de unde

$$\lim_{n\to\infty}n\left|1-\frac{t^2}{2n}-e^{-\frac{t^2}{2n}}\right|=0. \tag{3.23}$$

Pe de altă parte, deoarece $\mathbb{E}(Z_1)=0$ şi $\mathbb{E}(Z_1^2)=V(Z_1)+\mathbb{E}^2(Z_1)=1$, are loc formula:

$$n\left|\varphi\left(\frac{t}{\sqrt{n}}\right)-1+\frac{t^2}{2n}\right|=n\left|\mathbb{E}\left(e^{itZ_1/\sqrt{n}}-1-\frac{it}{\sqrt{n}}Z_1-\frac{(it)^2}{2n}Z_1^2\right)\right|.$$

Pentru o variabilă aleatoare V, avem $|\mathbb{E}(V)|\leq\mathbb{E}(|V|)$. Astfel, obţinem:

$$n\left|\varphi\left(\frac{t}{\sqrt{n}}\right)-1+\frac{t^2}{2n}\right|\leq n\mathbb{E}\left(\left|e^{itZ_1/\sqrt{n}}-1-\frac{it}{\sqrt{n}}Z_1-\frac{(it)^2}{2n}Z_1^2\right|\right). \tag{3.24}$$

Fie $\varepsilon>0$. Definim $\delta=\dfrac{3\varepsilon}{|t|^3}$ şi evenimentul $A_n(\delta)=\{|Z_1|>\delta\sqrt{n}\}$. Considerăm variabila aleatoare $1_{A_n(\delta)}$ indicatoare a evenimentului $A_n(\delta)$, definită prin

$$1_{A_n(\delta)}(\omega)=\begin{cases}1, & \omega\in A_n(\delta)\\ 0, & \omega\notin A_n(\delta)\end{cases}.$$

Pentru $\omega\in A_n(\delta)$, , are loc majorarea

$$\left|e^{itZ_1(\omega)/\sqrt{n}}-1-\frac{it}{\sqrt{n}}Z_1(\omega)-\frac{(it)^2}{2n}Z_1^2(\omega)\right|\leq\left|e^{itZ_1(\omega)/\sqrt{n}}-1-\frac{it}{\sqrt{n}}Z_1(\omega)\right|+\left|-\frac{(it)^2}{2n}Z_1^2(\omega)\right|\leq$$

$$\stackrel{(3.21)}{\leq}\frac{t^2}{2n}Z_1^2(\omega)+\frac{t^2}{2n}Z_1^2(\omega)=\frac{t^2}{n}Z_1^2(\omega),$$

iar pentru $\omega\in A_n^c(\delta)$, are loc majorarea

$$\left|e^{itZ_1(\omega)/\sqrt{n}}-1-\frac{it}{\sqrt{n}}Z_1(\omega)-\frac{(it)^2}{2n}Z_1^2(\omega)\right|\stackrel{(3.22)}{\leq}\frac{|t|^3}{6n\sqrt{n}}|Z_1(\omega)|^3.$$

Rezultă

$$\left|e^{itZ_1/\sqrt{n}}-1-\frac{it}{\sqrt{n}}Z_1-\frac{(it)^2}{2n}Z_1^2\right|\leq\frac{t^2}{2n}Z_1^2\cdot 1_{A_n(\delta)}+\frac{|t|^3}{6n\sqrt{n}}|Z_1|^3\cdot 1_{A_n^c(\delta)}.$$

Astfel,

$$n\mathbb{E}\left(\left|e^{itZ_1/\sqrt{n}}-1-\frac{it}{\sqrt{n}}Z_1-\frac{(it)^2}{2n}Z_1^2\right|\right)\leq\frac{t^2}{2}\mathbb{E}\left(Z_1^2\cdot 1_{A_n(\delta)}\right)+\frac{|t|^3}{6\sqrt{n}}\mathbb{E}\left(|Z_1|^3\cdot 1_{A_n^c(\delta)}\right).$$

Fie $F:\mathbb{R}\to[0,1]$ funcţia de repartiţie a variabilei aleatoare Z_1. Deoarece $\mathbb{E}(Z_1^2)=1$, avem

$$\frac{t^2}{2}\mathbb{E}\left(Z_1^2\cdot 1_{A_n(\delta)}\right)=\frac{t^2}{2}\int_{|x|>\delta\sqrt{n}}x^2dF(x)\stackrel{n\to\infty}{\longrightarrow}0,$$

deci există $n_\varepsilon \in \mathbb{N}^*$ astfel ca $\dfrac{t^2}{2}\mathbb{E}\left(Z_1^2 \cdot 1_{A_n(\delta)}\right) < \dfrac{\varepsilon}{2},\ \forall n \geq n_\varepsilon$.
Apoi,

$$\frac{|t|^3}{6\sqrt{n}}\mathbb{E}\left(|Z_1|^3 \cdot 1_{A_n^c(\delta)}\right) = \frac{|t|^3}{6\sqrt{n}}\int_{|x|\leq\delta\sqrt{n}} |x|^3 dF(x) \leq$$

$$\leq \frac{|t|^3\delta}{6}\int_{|x|\leq\delta\sqrt{n}} |x|^2 dF(x) \leq \frac{|t|^3\delta}{6}\int_{-\infty}^{\infty} |x|^2 dF(x) = \frac{|t|^3\delta}{6}\mathbb{E}(Z_1^2) = \frac{|t|^3\delta}{6} = \frac{\varepsilon}{2}.$$

Prin urmare,

$$n\mathbb{E}\left(\left|e^{itZ_1/\sqrt{n}} - 1 - \frac{it}{\sqrt{n}}Z_1 - \frac{(it)^2}{2n}Z_1^2\right|\right) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,\ \forall n \geq n_\varepsilon.$$

Cum $\varepsilon > 0$ este arbitrar, rezultă $\lim\limits_{n\to 0} n\mathbb{E}\left(\left|e^{itZ_1/\sqrt{n}} - 1 - \dfrac{it}{\sqrt{n}}Z_1 - \dfrac{(it)^2}{2n}Z_1^2\right|\right) = 0$. Atunci, din (3.24), obţinem

$$\lim_{n\to\infty} n\left|\varphi\left(\frac{t}{\sqrt{n}}\right) - 1 + \frac{t^2}{2n}\right| = 0. \tag{3.25}$$

Din inegalitatea (3.19) şi limitele (3.23) şi (3.25), rezultă

$$\lim_{n\to\infty} \Delta_n = 0,$$

ceea ce încheie demonstraţia versiunii Lindeberg-Lévi a teoremei.

Demonstrăm prima afirmaţie din enunţul teoremei. Fie $F_n : \mathbb{R} \to [0,1]$ funcţia de repartiţie a variabilei aleatoare $\dfrac{S_n - n\mu}{\sigma\sqrt{n}}$, iar $\Phi : \mathbb{R} \to [0,1]$ funcţia de repartiţie normală standard. Convergenţa în distribuţie demonstrată înseamnă:

$$\lim_{n\to\infty} F_n(x) = \Phi(x),\ \forall x \in \mathbb{R}.$$

Fie $[a,b] \subset \mathbb{R}$. Avem

$$\lim_{n\to\infty} \mathbb{P}\left\{a < \frac{S_n - n\mu}{\sigma\sqrt{n}} \leq b\right\} = \lim_{n\to\infty} [F_n(b) - F_n(a)] = \Phi(b) - \Phi(a) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}} dx.$$

Astfel, Teorema limită centrală este complet demonstrată. □

Comentariu. În cazul binomial, $\sigma = \sqrt{pq}$, deci Teorema 3.5.2 reprezintă un caz particular al Teoremei 3.5.3.

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

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

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

Au loc implicaţiile

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

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 $X \sim N(0, 1)$.

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 societăţii umane are adeseori la bază *experimentul*. 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. Verosimilitatea prognozei (încrederea pe care o putem avea rezultatele deduse din analiza datelor experimentale), este analizată cu instrumente matematice .

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. Scopul principal al unui studiu statistic îl reprezintă ***estimarea*** unei caracteristici a populaţiei statistice, pe baza datelor înregistrate în urma sondajului efectuat pentru un eşantion reprezentativ al populaţiei respective. Funcţia matematică care realizează estimarea se numeşte ***estimator***.

4.2 Noţiuni de statistică descriptivă

4.2.1 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 $f : \mathbb{R} \to [0, \infty)$ 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 volumul eşantionului tinde (ipotetic) la infinit.
- **Teoria probabilităţilor** operează cu noţiunile de *spaţiu de probabilitate* şi *variabilă aleatoare*. O variabilă aleatoare este o funcţie X, definită pe un spaţiu de 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:

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

  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:

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

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

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

Un caz particular îl constituie *variabilele aleatoare finite.* O astfel de variabilă aleatoare X ia 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:

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

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

$$\mu = \mathbb{E}(X) = x_1p_1 + x_2p_2 + \cdots + x_kp_k = \sum_{i=1}^{k} x_ip_i$$

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

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

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

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

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 μ şi dispersie σ^2 se "concentrează" cu o probabilitate de 95% în intervalul $[\mu - 2\sigma, \mu + 2\sigma]$, centrat în valoarea medie μ. Mai precis avem:

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

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

2. **Repartiţia** χ^2 cu $m \in \mathbb{N}^*$ **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:

$$\mathbb{E}(X) = m$$

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

Repartiţia χ^2 are un rol important în statistica matematică în verificarea *ipotezelor statistice.*

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

În 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 de valori.

Considerăm un eşantion fixat de volum n. Aceasta înseamnă că numărul de unităţi statistice pentru care se înregistrează valoarea caracteristicii este egal cu n.

Presupunem că valorile înregistrate ale caracteristicii sunt numerele reale

$$x_1 < x_2 < \cdots < x_k$$

unde k este un număr natural mai mic decât 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 $\displaystyle\sum_{i=1}^{k} n_i = n$.

- **Frecvenţa relativă a valorii** x_i este numărul $f_i = \frac{n_i}{n}$.

  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:*

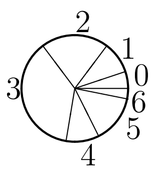
numărul de aruncări reuşite	0	1	2	3	4	5	6	
numărul de participanţi	2	4	8	15	4	6	1	total 40

Volumul eşantionului 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.

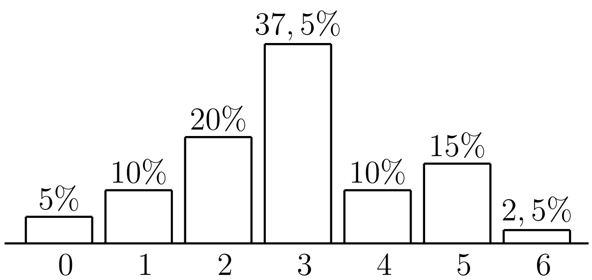
A) Tabelul frecvenţelor

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 %
	$\sum_{i=1}^{7} n_i = 40$	$\sum_{i=1}^{7} f_i = 1$	$\sum_{i=1}^{7} p_i = 100$

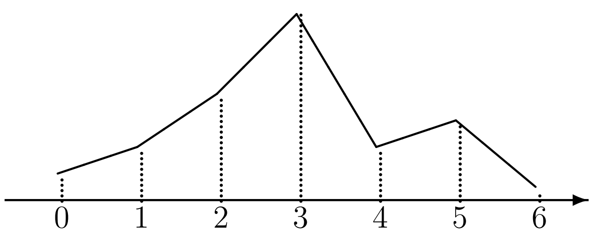
B) Diagrama de structură



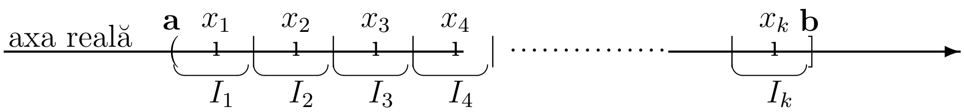
C) Histograma



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)	relative (f_i)	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 %
		$\sum_{i=1}^{6} n_i = 148$	$\sum_{i=1}^{6} f_i = 1$	$\sum_{i=1}^{6} p_i = 100$

4.2.3 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 ”aşteptată” 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 X, denumită variabilă aleatoare empirică, care ia valorile $x_1 < x_2 < \cdots < x_k$ cu frecvenţele absolute $n_1, n_2, \cdots, n_k$ şi frecvenţele relative $f_1, f_2, \cdots, f_k$.

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

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

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

- **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:

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

unde $\overline{x} = \mathbb{E}(X)$ 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.

4.2.4 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:

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

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.*

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

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

Rezultă că timpul mediu înregistrat al probei este:

$$\mathbb{E}(Y) = \overline{x} = 49,29' \cong 49'18''.$$

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.*

4.2.6 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.

4.2.7 Corelaţia şi coeficientul de corelaţie

Fie $\overline{x} = \mathbb{E}(X)$, $\overline{y} = \mathbb{E}(Y)$ 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 defineşte prin:

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

Coeficientul de corelaţie (Pearson) $r = r_{X,Y}$ **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.*

4.2.8 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.*

4.3 Statistica eşantioanelor Bernoulli

Cel mai simplu studiu statistic se referă la analiza cazului binar: fiecare individ (element) al populaţiei statistice are valoarea 1 sau 0 (interpretare posibilă: *la un chestionar, fiecare unitate statistică optează pentru 0 sau 1*).

Ne interesează proporţia indivizilor cu valoarea 1 în cadrul întregii populaţii statistice. Examinarea (chestionarea) prin sondaj a fiecărui individ al eşantionului selecţionat ne va permite să determinăm numărul indivizilor care "valorează 1" dintr-un eşantion de volum n. Pe baza datelor obţinute, dorim să apreciem care este probabilitatea ca un individ ales arbitrar din populaţia statistică să "valoreze 1".

Cazul unei populaţii de volum mic în raport cu eşantionul considerat

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.
Fie a numărul de indivizi ai populaţiei statistice cu caracteristica de valoare 1 şi b numărul de indivizi ai populaţiei statistice cu caracteristica de valoare 0. Astfel $N = a+b$ reprezintă volumul populaţiei, iar

$$\theta = \frac{a}{a+b}$$

reprezită probabilitatea ca un individ X, arbitrar, al populaţiei statistice să valoreze 1. Avem deci

$$\theta = \mathbb{P}\{X = 1\} = \frac{a}{a+b}.$$

Să presupunem la un sondaj de volum $n < N$, se constată că $0 \le i \le n$ indivizi au valoarea caracteristicii egală cu 1.

Aproximarea $\theta \approx \frac{i}{n}$ este plauzibilă, dar trebuie privită cu precauţie, deoarece probabilitatea ca exact i indivizi din n să valoreze 1 este dată de

$$p_{i/n} = \frac{C_a^i C_b^{n-i}}{C_N^n},$$

conform **schemei hipergeometrice**.

Cazul unei populaţii de volum mare în raport cu eşantionul considerat
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 probabilitatea (şansa) ca unitatea statistică X să aibă valoarea 1. Evident, $\mathbb{P}\{X = 0\} = 1 - \theta$ reprezintă probabilitatea (şansa) ca X să ia valoarea 0.

Obiectul studiului statistic întreprins constă în **estimarea** parametrului θ, din observarea unui eşantion de volum n.

Eşantionul poate fi reprezentat prin vectorul aleator $(X_1, X_2, \cdots, X_n)$, cu componentele variabile aleatoare iid, de tip $B(1, \theta)$.

Distribuţia eşantionului, care ia valori în mulţimea $\{0, 1\}^n$, este determinată de distribuţia (comună) a variabilelor aleatoare $X_1, X_2, \cdots, X_n$.

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:

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

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 $\overline{X}_n$, ca estimator a lui θ.

4.3.1 Proprietăţile estimatorului $\overline{X}_n$

1. $\overline{X}_n$ *este un* ***estimator fără abatere*** *a lui* θ*, adică:*

$$\mathbb{E}_\theta(\overline{X}_n) = \theta, \ \forall\, \theta \in (0,1).$$

Demonstraţie. Pentru orice $\theta \in (0,1)$ avem:

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

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

Demonstraţie. Fie $g : \{0,1,2,\cdots,n\} \to \mathbb{R}$ o funcţie cu proprietatea că $\widetilde{X} = g(S_n)$ este un estimator fără abatere a parametrului θ, adică:

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

Deoarece $S_n \sim B(n,\theta)$, avem

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

iar pe de altă parte

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

Astfel, $\sum_{k=0}^{n} g(k) C_n^k \theta^k (1-\theta)^{n-k} = \theta = \sum_{k=0}^{n} \frac{k}{n} C_n^k \theta^k (1-\theta)^{n-k}$, $\forall\, \theta \in (0,1)$, de unde 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 $f(X) = \sum_{k=0}^{n} \left(g(k) - \frac{k}{n}\right) C_n^k X^k$ are o infinitate de rădăcini, deci este polinomul identic nul.

Prin urmare $g(k) = k/n, \ k = 0,1,\cdots,n$, deci

$$\widetilde{X} = g(S_n) = \frac{S_n}{n} = \overline{X}_n.$$

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

Demonstraţie.
Să presupunem că, în urma sodajului, au fost observate valorile $(x_1, x_2, \cdots, x_n) \in \{0,1\}^n$. Notăm $s_n = \sum_{k=1}^{n} x_k$. Astfel, s_n indică numărul de componente egale cu 1 ale vectorului $(x_1, x_2, \cdots, x_n)$.

Datorită independenţei variabilelor aleatoare $X_1, X_2, \cdots, X_n$, probabilitatea de a observa secvenţă de valori $(x_1, x_2, \cdots, x_n)$ este

$$\mathbb{P}\{X_1 = x_1,\ X_2 = x_2,\ \cdots,\ X_n = x_n\} = \prod_{i=1}^{n} \mathbb{P}\{X_i = x_i\} = \theta^{s_n}(1-\theta)^{n-s_n}.$$

Atunci, cea mai *verosimilă* valoare a parametrului θ este cea care maximizează funcţia

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

Evident, $0 \leq s_n \leq n$. Avem

$$L'(x) = s_n x^{s_n-1}(1-x)^{n-s_n} - (n-s_n)x^{s_n}(1-x)^{n-s_n-1} =$$

$$= x^{s_n-1}(1-x)^{n-s_n-1}\left[s_n(1-x) - (n-s_n)x\right] = x^{s_n-1}(1-x)^{n-s_n-1}(s_n - nx).$$

1) Dacă $s_n \in (1,n)$, atunci $\dfrac{s_n}{n} \in (0,1)$. Deducem că L' admite rădăcina unică $\overline{x} = \dfrac{s_n}{n}$, este strict pozitivă pe intervalul $(0,\overline{x})$ şi strict negativă pe intervalul $(\overline{x},1)$. Ca urmare, L îşi atinge maximul în punctul $\overline{x} = \dfrac{s_n}{n}$.
2) Daca $s_n = 0$, atunci funcţia L este strict descrescătoare pe $(0,1)$ şi vom alege, în mod firesc, $\overline{x} = 0 = \dfrac{s_n}{n}$.
3) Dacă $s_n = n$, atunci funcţia L este strict crescătoare pe $(0,1)$ şi alegem $\overline{x} = 1 = \dfrac{s_n}{n}$.
Rezultă că

$$\overline{X}_n = \frac{S_n}{n}$$

reprezintă estimatorul de maximă verosimilitate a lui θ.

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

Pe baza inegalităţii lui Cebîşev, se demonstrează convergenţa în probabilitate stabilită de **Legea slabă numerelor mari**:

$$\overline{X}_n = \frac{S_n}{n} \xrightarrow{p} \theta.$$

Ca urmare, are loc şi convergenţa in distribuţie

$$\overline{X}_n = \frac{S_n}{n} \xrightarrow{d} \theta$$

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

5. *Când volumul eşantionului tinde la infinit,* ***riscul pătratic mediu al estimatorului*** *$\overline{X}_n$ a lui θ* ***converge la 0****.*

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

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

$$= \frac{nV(X_1)}{n^2} = \frac{\theta(1-\theta)}{n} \leq \frac{1}{4n},$$

de unde

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

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

Demonstraţie. Conform **Teoremei limită centrală**,

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

unde $Z \sim N(0,1)$.

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

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

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

4.3.2 Estimare bayesiană

Media empirică $\overline{X}_n$ 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

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

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

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

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

4.4 Statistici, estimatori

Am urmărit anterior exemplul concret al estimării parametrului unei distribuţii Bernoulli. Vom defini acum, în cadrul general, noţiunile de **statistică** şi **estimator**.

În cele ce urmează, X desemnează o caracteristică statistică cantitativă (măsurabilă) a unei populaţii statistice P. Caracteristica X se interpretează din punct de vedere teoretic ca o variabilă aleatoare. Concret, vom presupune că $X : \Omega \to \mathbb{R}$ este o variabilă aleatoare definită pe un spaţiu de probabilitate $(\Omega, \mathcal{F}, \mathbb{P})$, de regulă neexplicitat.

Considerăm ***variabilele aleatoare de selecţie*** $X_1, X_2, \ldots, X_n$, reprezentând valorile măsurate ale caracteristicii pentru indivizii i_1, $i_2, \ldots, i_n$ selecţionaţi prin sondaj din populaţia statistică P.

Teoretic, considerăm un şir de variabile aleatoare independente şi identic distribuite (i.i.d), având distribuţia variabilei aleatoare X. Astfel, X_1, $X_2, \ldots$ reprezintă ”replici” independente ale variabilei aleatoare X.

În rezumat, formalizarea teoretică este realizată în felul următor. Fie

$$\Omega^\infty = \{\overline{\omega} = (\omega_i)_{i\in\mathbb{N}^*} : \ \omega_i \in \Omega, \ \forall\, i \in \mathbb{N}^*\}$$

mulţimea şirurilor cu termenii în mulţimea Ω. Pentru fiecare $n \in \mathbb{N}^*$, variabila aleatoare de selecţie X_n se defineşte prin:

$$X_n : \Omega^\infty \to \mathbb{R}, \ X_n(\overline{\omega}) = X(\omega_n), \ \overline{\omega} \in \Omega^\infty,$$

unde ω_n este termenul de rang n al şirului $\overline{\omega} \in \Omega^\infty$.

Se construieşte spaţiul de probabilitate $(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty)$ (nu dăm detalii), astfel încât, pentru oricare număr natural nenul n, să avem

$$\mathbb{P}^\infty\{X_n \in A\} = \mathbb{P}\{X \in A\}, \ \forall\, A \in \mathcal{B}_{\mathbb{R}},$$

iar

$$\mathbb{P}^\infty\{X_i \in A_i, \ i = \overline{1,n}\} = \prod_{i=1}^{n} \mathbb{P}^\infty\{X_i \in A_i\} = \prod_{i=1}^{n} \mathbb{P}\{X \in A_i\}, \ \forall\, A_i \in \mathcal{B}_{\mathbb{R}}, \ \overline{1,n}.$$

Enunţăm definiţiile conceptelor de ***estimator*** şi ***statistică***.

Definiţia 4.4.1. *Considerăm o variabilă aleatoare X definită pe spaţiul de probabilitate $(\Omega, \mathcal{F}, \mathbb{P})$ şi un şir de variabile aleatoare de selecţie $(X_n)_{n\geq 1}$. Fie $A \in \mathbb{R}$ o caracteristică numerică a distribuţiei lui X, iar $(g_n)_{n\geq 1}$ un şir de funcţii, $g_n : \mathbb{R}^n \to \mathbb{R}, \ n \in \mathbb{N}^*$.*

1. *Şirul de variabile aleatoare $(g_n(X_1, X_2, \cdots, X_n))_{n\geq 1}$, unde*

   $$g_n(X_1, X_2, \cdots, X_n) : \Omega^\infty \to \mathbb{R}, \ n \in \mathbb{N}^*,$$

   *se numeşte **statistică** (asociată şirurilor $(g_n)_{n\geq 1}$ şi $(X_n)_{n\geq 1}$).*

2. *Statistica $(g_n(X_1, X_2, \cdots, X_n))_{n\geq 1}$ se numeşte:*

   *(a) **corectă relativ la** A dacă*

   (i) $\lim\limits_{n\to\infty} \mathbb{E}[g_n(X_1, X_2, \cdots, X_n)] = A$;

   (ii) $\lim\limits_{n\to\infty} V[g_n(X_1, X_2, \cdots, X_n)] = 0$.

   *(b) **fără abatere (absolut corectă/nedeplasată) relativ la** A dacă*

   (i) $\mathbb{E}[g_n(X_1, X_2, \cdots, X_n)] = A, \ \forall\, n \in \mathbb{N}^*$;

   (ii) $\lim\limits_{n\to\infty} V[g_n(X_1, X_2, \cdots, X_n)] = 0$.

(c) ***estimator pentru*** *A dacă*

$$\lim_{n\to\infty} \mathbb{P}^\infty\{|g_n(X_1, X_2, \cdots, X_n) - A| > \varepsilon\} = 0,\ \forall\, \varepsilon > 0.$$

Evident, o statistică fără abatere relativ la A este o statistică corectă relativ la A. Are loc următorul rezultat fundamental.

Teorema 4.4.1. *O statistică corectă relativ la o caracteristică numerică a unei distribuţii este un estimator pentru acea caracteristică.*

Demonstraţie.
Fie $(g_n(X_1, X_2, \cdots, X_n))_{n\geq 1}$ o statistică corectă relativ la A. Atunci sunt satisfăcute condiţiile:

(i) $\lim\limits_{n\to\infty} \mathbb{E}[g_n(X_1, X_2, \cdots, X_n)] = A$;

(ii) $\lim\limits_{n\to\infty} V[g_n(X_1, X_2, \cdots, X_n)] = 0$.

Fie $\varepsilon > 0$, arbitrar.
Din (i) rezultă că există $N \in \mathbb{N}^*$ astfel încât

$$|\mathbb{E}[g_n(X_1, X_2, \cdots, X_n)] - A| < \frac{\varepsilon}{2},\ \forall\, n \geq N,$$

de unde

$$\varepsilon - |\mathbb{E}[g_n(X_1, X_2, \cdots, X_n)] - A| > \frac{\varepsilon}{2},\ \forall\, n \geq N.$$

Considerăm în continuare $n \in \mathbb{N}^*,\ n \geq N$. Din inegalitatea

$$|g_n(X_1, \cdots, X_n) - A| \leq |g_n(X_1, \cdots, X_n) - \mathbb{E}[g_n(X_1, \cdots, X_n)]| + |\mathbb{E}[g_n(X_1, \cdots, X_n)] - A|,$$

obţinem

$$\begin{aligned} &\{|g_n(X_1, \cdots, X_n) - A| > \varepsilon\} \subset \\ &\subset \{|g_n(X_1, \cdots, X_n) - \mathbb{E}[g_n(X_1, \cdots, X_n)]| + |\mathbb{E}[g_n(X_1, \cdots, X_n)] - A| > \varepsilon\} = \\ &= \{|g_n(X_1, \cdots, X_n) - \mathbb{E}[g_n(X_1, \cdots, X_n)]| > \varepsilon - |\mathbb{E}[g_n(X_1, \cdots, X_n)] - A|\} \subset \\ &\subset \left\{|g_n(X_1, \cdots, X_n) - \mathbb{E}[g_n(X_1, \cdots, X_n)]| > \frac{\varepsilon}{2}\right\}. \end{aligned}$$

Probabilitatea este o funcţie monoton crescătoare, deci

$$\mathbb{P}^\infty\{|g_n(X_1, \cdots, X_n) - A| > \varepsilon\} \leq \mathbb{P}^\infty\left\{|g_n(X_1, \cdots, X_n) - \mathbb{E}[g_n(X_1, \cdots, X_n)]| > \frac{\varepsilon}{2}\right\},\ \forall\, n \geq \mathbb{N}.$$

Conform inegalităţii lui Cebîşev, avem

$$\mathbb{P}^\infty\left\{|g_n(X_1, \cdots, X_n) - \mathbb{E}[g_n(X_1, \cdots, X_n)]| > \frac{\varepsilon}{2}\right\} \leq \frac{V[g_n(X_1, X_2, \cdots, X_n)]}{(\varepsilon/2)^2},$$

de unde

$$\mathbb{P}^\infty\{|g_n(X_1, \cdots, X_n) - A| > \varepsilon\} \leq \frac{4V[g_n(X_1, X_2, \cdots, X_n)]}{\varepsilon^2},\ \forall\, n \geq N.$$

Din condiţia (ii) şi criteriul majorării, obţinem

$$\lim_{n\to\infty} \mathbb{P}^{\infty}\{|g_n(X_1, \cdots, X_n) - A| > \varepsilon\} = 0.$$

Rezultă că statistica $(g_n(X_1, X_2, \cdots, X_n))_{n\geq 1}$ este un estimator pentru A. □

Observaţia 4.4.1. *O statistică fără abatere relativ la o caracteristică numerică a unei distribuţii este de asemenea un estimator pentru acea caracteristică. În acest caz, estimatorul se va numi fără abatere.*

Definiţia 4.4.2. *Pentru şirul de variabile aleatoare de selecţie $(X_n)_{n\geq 1}$ definim:*

1. *media empirică (media de selecţie) prin*

$$\overline{X}_n = \frac{1}{n}(X_1 + X_2 + \cdots + X_n), \quad n \in \mathbb{N}^*.$$

2. *dispersia empirică (dispersia de selecţie) prin*

$$\overline{V}_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2, \quad n \in \mathbb{N}^*.$$

3. *dispersia empirică modificată (dispersia de selecţie modificată) prin*

$$\overline{S}_n = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2, \quad n \in \mathbb{N}^*,\ n \geq 2.$$

Teorema 4.4.2. *Fie $X \in \mathcal{L}^2(\Omega)$. Statistica $(g_n(X_1, X_2, \cdots, X_n))_{n\geq 1}$ definită prin*

$$g_n(X_1, X_2, \cdots, X_n) = \overline{X}_n,\ \forall\, n \in \mathbb{N}^*,$$

este fără abarere (absolut corectă/nedeplasată) relativ la $A = \mathbb{E}(X)$.

Demonstraţie.
Trebuie să demonstrăm relaţiile:

(i) $\mathbb{E}[g_n(X_1, X_2, \cdots, X_n)] = A,\ \forall\, n \in \mathbb{N}^*$;

(ii) $\lim\limits_{n\to\infty} V[g_n(X_1, X_2, \cdots, X_n)] = 0$.

Avem

$$\mathbb{E}[g_n(X_1, X_2, \cdots, X_n)] = \mathbb{E}(\overline{X}_n) = \frac{1}{n}\mathbb{E}\left(\sum_{i=1}^{n} X_i\right) = \frac{\sum_{i=1}^{n}\mathbb{E}(X_i)}{n} = \frac{nA}{n} = A,\ \forall\, n \in \mathbb{N}^*.$$

Apoi

$$V[g_n(X_1, X_2, \cdots, X_n)] = V(\overline{X}_n) = \frac{1}{n^2}V\left(\sum_{i=1}^{n} X_i\right) \overset{indep.}{=} \frac{\sum_{i=1}^{n} V(X_i)}{n^2} = \frac{nV(X)}{n^2} = \frac{V(X)}{n} \overset{n\to\infty}{\longrightarrow} 0.$$

Deci $(\overline{X}_n)_{n\geq 1}$ este o statistică fără abatere relativ la medie. □

Prim urmare, media empirică este şi un estimator fără abatere pentru medie, fapt demonstrat anterior pentru cazul particular al eşantioanelor Bernoulli.

Teorema 4.4.3. *Fie* $X \in \mathcal{L}^4(\Omega)$.

1. *Statistica* $(g_n(X_1, X_2, \cdots, X_n))_{n\geq 1}$ *definită prin*

$$g_n(X_1, X_2, \cdots, X_n) = \overline{V}_n, \ \forall \ n \in \mathbb{N}^*,$$

este corectă relativ la $A = V(X)$.

2. *Statistica* $(g_n(X_1, X_2, \cdots, X_n))_{n\geq 1}$ *definită prin*

$$g_n(X_1, X_2, \cdots, X_n) = \overline{S}_n, \ \forall \ n \in \mathbb{N}^*,$$

este absolut corectă (fără abatere) relativ la $A = V(X)$.

Demonstraţie.
Demonstrăm la început relaţiile

$$\lim_{n\to\infty} \mathbb{E}\left(\overline{V}_n\right) = V(X), \tag{4.1}$$

$$\mathbb{E}\left(\overline{S}_n\right) = V(X), \ \forall n \in \mathbb{N}^*. \tag{4.2}$$

Avem

$$\mathbb{E}(\overline{V}_n) = \mathbb{E}\left[\frac{\sum_{i=1}^n (X_i - \overline{X}_n)^2}{n}\right] = \frac{\sum_{i=1}^n \mathbb{E}\left[(X_i - \overline{X}_n)^2\right]}{n}. \tag{4.3}$$

Pentru $i \in \{1, 2, \cdots, n\}$, obţinem

$$\mathbb{E}\left[(X_i - \overline{X}_n)^2\right] = \frac{1}{n^2}\mathbb{E}\left[\left(nX_i - \sum_{k=1}^n X_k\right)^2\right] = \frac{1}{n^2}\mathbb{E}\left[\left((n-1)X_i - \sum_{k\neq i} X_k\right)^2\right] =$$

$$= \frac{\mathbb{E}\left[(n-1)^2X_i^2 - 2(n-1)\sum_{k\neq i} X_iX_k + \sum_{k\neq i} X_k^2 + \sum_{k,j\neq i,\ k\neq j} X_kX_j\right]}{n^2} =$$

$$= \frac{(n-1)^2\mathbb{E}(X_i^2) - 2(n-1)\sum_{k\neq i}\mathbb{E}(X_iX_k) + \sum_{k\neq i}\mathbb{E}(X_k^2) + \sum_{k,j\neq i,\ k\neq j}\mathbb{E}(X_kX_j)}{n^2}.$$

Datorită faptului că variabilelor aleatoare $X_1, X_2, \cdots, X_n$ sunt independente, cu distribuţia variabilei aleatoare X, rezultă

$$\mathbb{E}(X_i^2) = \mathbb{E}(X_k^2) = \mathbb{E}(X^2),\ \mathbb{E}(X_iX_k) = \mathbb{E}(X_i)\mathbb{E}(X_k) = \mathbb{E}^2(X) \text{ şi } \mathbb{E}(X_kX_j) = \mathbb{E}(X_k)\mathbb{E}(X_j) = \mathbb{E}^2(X).$$

Atunci

$$\mathbb{E}\left[(X_i - \overline{X}_n)^2\right] = \frac{(n-1)^2\mathbb{E}(X^2) - 2(n-1)^2\mathbb{E}^2(X) + (n-1)E(X^2) + (n-1)(n-2)\mathbb{E}^2(X)}{n^2} =$$

$$= \frac{n(n-1)\mathbb{E}(X^2) - n(n-1)\mathbb{E}^2(X)}{n^2} = \frac{n-1}{n}V(X),\ \forall i \in \{1, 2, \cdots, n\}.$$

Înlocuind în (4.3) obţinem:

$$\mathbb{E}(\overline{V}_n) = \frac{n-1}{n}V(X) \overset{n\to\infty}{\longrightarrow} V(X), \tag{4.4}$$

deci relaţia (4.1) este demonstrată. Relaţia (4.2) rezultă astfel:

$$\mathbb{E}(\overline{S}_n) = \mathbb{E}\left(\frac{n}{n-1}\overline{V}_n\right) = \frac{n}{n-1}\mathbb{E}(\overline{V}_n) \overset{(4)}{=} \frac{n}{n-1}\cdot\frac{n-1}{n}\cdot V(X) = V(X),\ \forall n \in \mathbb{N}^*.$$

Demonstrăm acum relaţiile

$$\lim_{n\to\infty} V(\overline{S}_n) = 0, \tag{4.5}$$

$$\lim_{n\to\infty} V(\overline{V}_n) = 0. \tag{4.6}$$

Pentru $n \geq 2$, avem

$$V(\overline{S}_n) = \mathbb{E}\left(\overline{S}_n^2\right) - \mathbb{E}^2\left(\overline{S}_n\right) \overset{(4.2)}{=} \mathbb{E}\left(\overline{S}_n^2\right) - V^2(X). \tag{4.7}$$

Explicităm $\overline{S}_n$ şi $\overline{S}_n^2$. Avem:

$$\overline{S}_n = \frac{\sum_{i=1}^n (X_i - \overline{X}_n)^2}{n-1} = \frac{\sum_{i=1}^n (X_i^2 - 2X_i\overline{X}_n + \overline{X}_n^2)}{n-1} =$$

$$= \frac{\sum_{i=1}^n X_i^2 - 2n\overline{X}_n^2 + n\overline{X}_n^2}{n-1} = \frac{\sum_{i=1}^n X_i^2 - n\overline{X}_n^2}{n-1}$$

şi

$$\overline{S}_n^2 = \frac{\left(\sum_{i=1}^n X_i^2\right)^2 - 2n\overline{X}_n^2\left(\sum_{i=1}^n X_i^2\right) + n^2\overline{X}_n^4}{(n-1)^2}.$$

Avem:

$$\mathbb{E}\left[\left(\sum_{i=1}^n X_i^2\right)^2\right] = \mathbb{E}\left[\sum_{i=1}^n X_i^4 + \sum_{i\neq j} X_i^2X_j^2\right] =$$

$$= \sum_{i=1}^n \mathbb{E}(X_i^4) + \sum_{i\neq j}\mathbb{E}(X_i^2X_j^2) = n\mathbb{E}(X^4) + n(n-1)\mathbb{E}^2(X^2).$$

Procedând analog pentru ceilalţi termeni, obţinem în final

$$\mathbb{E}\left(\overline{S}_n^2\right) = \frac{1}{n}\mathbb{E}(X^4) + \frac{n^2-2n+3}{n(n-1)}V^2(X).$$

Înlocuind $\mathbb{E}\left(\overline{S}_n^2\right)$ în relaţia (4.7) găsim:

$$V(\overline{S}_n) = \frac{1}{n}\left[\mathbb{E}(X^4) - \frac{n-3}{n-1}V^2(X)\right] \overset{n\to\infty}{\longrightarrow} 0.$$

În sfârşit,

$$V(\overline{V}_n) = V\left(\frac{n-1}{n}\overline{S}_n\right) = \frac{(n-1)^2}{n^2}V(\overline{S}_n) \overset{n\to\infty}{\longrightarrow} 0,$$

ceea ce încheie demonstraţia. □

4.5 Intervale de încredere

Considerăm o caracteristică cantitativă a unei populaţii P, modelată prin variabila aleatoare X. Fie A un parametru (o caracteristică numerică) a variabilei aleatoare X.

Definiţia 4.5.1. *Presupunem că $g' = (g'_n(X_1, X_2, \ldots, X_n))_{n\in\mathbb{N}}$ şi $g'' = (g''_n(X_1, X_2, \ldots, X_n))_{n\in\mathbb{N}}$ reprezintă două statistici pentru parametrul A, cu proprietatea*

$$g'_n(x_1, \ldots, x_n) < g''_n(x_1, \ldots, x_n),\ \forall\ x_1, \ldots, x_n \in \mathbb{R},\ \forall\ n \in \mathbb{N}^*.$$

Spunem că $[g', g'']$ este un interval de încredere de nivel $\alpha \in (0, 1)$ pentru parametrul A dacă

$$\mathbb{P}\{g' \leq A \leq g''\} \geq \alpha.$$

Explicit,

$$\mathbb{P}\{g'_n(X_1, \ldots, X_n) \leq A \leq g''_n(X_1, \ldots, X_n)\} \geq \alpha,\ \forall\ n \in \mathbb{N}^*.$$

Numărul $\varepsilon = 1 - \alpha \in (0, 1)$ se numeşte prag de încredere pentru intervalul $[g', g'']$.

Observaţia 4.5.1. *De regulă, se consideră $0,9 < \alpha < 1$, deci $0 < \varepsilon < \dfrac{1}{10}$.*

Definiţia poate fi înteleasă astfel. Pentru un număr natural arbitrar $n > 1$, considerăm N selecţii $(x_1, \ldots, x_n)$ de volum n din populaţia P, privitoare la caracteristica măsurată X. Notăm K_N numărul de selecţii pentru care $g'_n(x_1, \ldots, x_n) \leq A \leq g''_n(x_1, \ldots, x_n)$. Atunci $\dfrac{K_N}{N} > \alpha$, pentru N "mare".

Un interval de încredere pentru un paramentru A poate fi construit pe modelul

$$[g_n(X_1, \ldots, X_n) - a_n, g_n(X_1, \ldots, X_n) + a_n],$$

unde $g = (g_n(X_1, \ldots, X_n))_{n\in\mathbb{N}^*}$ este o statistică corectă relativ la A, iar $(a_n)_{n\in\mathbb{N}}$ este un şir pozitiv convergent la 0.

4.5.1 Exemplu: intervale de încredere pentru medie

Vom exemplifica noţiunea de interval de încredere. Presupunem că o caracteristică X a populaţiei P are o distribuţie normală de dispersie cunoscută σ^2 şi medie necunosctă μ. Ne propunem să definim intervale de încredere pentru μ. Ne vom servi de următorul rezultat.

Propoziţia 4.5.1. *Dacă $X_1, X_2, \ldots, X_n$ sunt variabile aleatoare normale independente, de medie comună μ şi dispersie comună σ^2, atunci variabila aleatoare*

$$\overline{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

are o distribuţie normală de medie μ şi dispersie σ^2/n, iar variabila aleatoare $\overline{Z}_n = \dfrac{\overline{X}_n - \mu}{\sigma/\sqrt{n}}$ are o distribuţie normală standard.

Demonstraţie. Variabilele aleatoare X_k au funcţia caracteristică comună

$$\varphi(t) = e^{i\mu t - \frac{t^2\sigma^2}{2}},\ t \in \mathbb{R}.$$

Atunci variabila aleatoare $\overline{X}_n$ are funcţia caracteristică

$$\varphi_{\overline{X}_n}(t) = \varphi_{(X_1+\cdots+X_n)/n}(t) = \varphi_{X_1+\cdots+X_n}(t/n) = \prod_{k=1}^{n} \varphi_{X_k}(t/n) = \varphi^n(t/n) =$$

$$= \left(e^{i\mu\frac{t}{n} - \frac{t^2\sigma^2}{2n^2}}\right)^n = e^{i\mu t - \frac{t^2\sigma^2}{2n}},\ t \in \mathbb{R}.$$

Din teorema de inversiune rezultă că variabila aleatoare $\overline{X}_n$ are distribuţia normală de medie μ şi dispersie σ^2/n.

Apoi

$$\varphi_{\overline{Z}_n}(t) = \mathbb{E}\left(e^{it\cdot\frac{\overline{X}_n-\mu}{\sigma/\sqrt{n}}}\right) = e^{-\frac{it\mu\sqrt{n}}{\sigma}}\mathbb{E}\left(e^{i\left(\frac{t\sqrt{n}}{\sigma}\right)\overline{X}_n}\right) = e^{-\frac{it\mu\sqrt{n}}{\sigma}}\varphi_{\overline{X}_n}\left(\frac{t\sqrt{n}}{\sigma}\right) =$$

$$= e^{-\frac{it\mu\sqrt{n}}{\sigma}} \cdot e^{i\mu\left(\frac{t\sqrt{n}}{\sigma}\right) - \frac{\left(\frac{t\sqrt{n}}{\sigma}\right)^2\sigma^2}{2n}} = e^{-\frac{t^2}{2}},\ t \in \mathbb{R}.$$

Din teorema de inversiune rezultă că variabila aleatoare $\overline{Z}_n$ are distribuţia normală standard. □

Pentru variabilele aleatoare de selecţie $X_1, X_2, \ldots, X_n$, definim variabila aleatoare

$$\overline{Z}_n = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}}.$$

Din Propoziţia 4.5.1 deducem că $\overline{Z}_n \sim N(0,1)$. Fie $\alpha \in (0,1)$, fixat. Ne propunem să determinăm un interval de încredere de nivel α pentru parametrul necunoscut μ. Notăm

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}}\, dt,\ x > 0.$$

Determinăm (din tabele sau cu ajutorul unui soft statistic) *cuantila* $a > 0$ astfel ca

$$2\Phi(a) = \alpha.$$

Avem

$$\mathbb{P}\{|\overline{Z}_n| \le a\} = \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} e^{-\frac{t^2}{2}}\, dt = 2\Phi(a) = \alpha.$$

Pe de altă parte,

$$\{|\overline{Z}_n| \le a\} = \left\{-a \le \frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} \le a\right\} = \left\{-\frac{a\sigma}{\sqrt{n}} \le \overline{X}_n - \mu \le \frac{a\sigma}{\sqrt{n}}\right\} =$$

$$= \left\{\overline{X}_n - \frac{a\sigma}{\sqrt{n}} \le \mu \le \overline{X}_n + \frac{a\sigma}{\sqrt{n}}\right\}.$$

Prin urmare,

$$\mathbb{P}\left\{\mu \in \left[\overline{X}_n - \frac{a\sigma}{\sqrt{n}}, \overline{X}_n + \frac{a\sigma}{\sqrt{n}}\right]\right\} = \alpha,$$

deci $\left[\overline{X}_n - \frac{a\sigma}{\sqrt{n}}, \overline{X}_n + \frac{a\sigma}{\sqrt{n}}\right]$ este un interval de încredere de nivel α pentru parametrul necunoscut μ.

4.6 Ipoteze statistice. Teste statistice

Pentru o populaţie P, analizăm statistic o caracteristică cantitativă (numerică) X.
Avem două alternative.
1. Presupunem, din anumite raţiuni, că variabila empirică X are distribuţie (lege) teoretică cunoscută. Frecvent, considerăm că X are o distribuţie Bernoulli, normală sau Poisson.
2. Nu avem argumente să încadrăm distribuţia lui X într-o familie de distribuţii teoretice cunoscute.

În primul caz, vom formula şi analiza statistic "ipoteze" privind valoarea parametrului/parametrilor (necunoscuţi) de care depinde distribuţia teoretică a caracteristicii X. Pentru studiul statistic întreprins, ne vom servi de rezultatele privind estimarea parametrului/parametrilor de care depinde distribuţia teoretică (presupusă) a caracteristicii X. De exemplu, estimăm media unei distribuţii normale cu dispersie cunoscută.

4.6.1 Descrierea tipurilor de ipoteze statistice

O **ipoteză statistică** reprezintă o presupunere făcută asupra uneia sau mai multor "însuşiri" ale populaţiei P, adică asupra tipului de distribuţie a caracteristicii X a populaţiei. În cazul când presupunem aprioric ca X este guvernată de o lege teoretică cunoscută, vom face presupuneri asupra valorilor parametrilor de care depinde distribuţia respectivă. Prezentăm şi exemplificăm principalele tipuri de ipoteze statistice.

a) Ipoteza statistică se numeşte **simplă** dacă include întreaga informaţie care determină distribuţia caracteristicii X.

Exemple.

1. Ipoteza statistică simplă H:
   Caracteristica X urmează o lege Bernoulli, de parametru $p = 0,4$.

2. Ipoteza statistică simplă H:
   Caracteristica X urmează o lege normală, de medie $\mu = 5$ şi dispersie $\sigma^2 = 4$.

b) Ipoteza statistică se numeşte **compusă** dacă se referă doar la o parte din informaţiile care determină distribuţia caracteristicii X.

Exemple.

1. Ipoteza statistică compusă H:
   Caracteristica X urmează o lege Poisson.

2. Ipoteza statistică compusă H:
   Caracteristica X urmează o lege normală, de medie $\mu = 20$.

In cazul ipotezelor compuse, parametrii neprecizaţi, care împreuna cu cei precizaţi în ipoteza statistică determină complet distributia lui X, se estimează prin examinarea unui eşantion al populaţiei, obţinut în urma unui proces de selecţie.

c) Ipoteza statistică se numeşte **exactă** dacă indică tipul distribuţiei teoretice a lui X şi o valoare precisă a unuia sau mai multor parametrii ai distribuţiei respective.

Exemplu.
Ipoteza statistică exactă H:
Caracteristica X urmează o lege Poisson de parametru $a = 2$.
d) Ipoteza statistică se numeşte **inexactă** dacă nu oferă valori precise.

Exemplu.
Ipoteza statistică inexactă H:
Caracteristica X are media $m \in [6, 8]$.
În statistică se operează simultan cu o pereche de ipoteze statistice alternative: ipoteza H_0 şi ipoteza alternativă H_1. Ipoteza alternativă H_1 nu reprezintă întotdeauna negaţia logică a ipotezei H_0.

Exemplu.
Ipoteza H_0: *media lui X este $m = 8,33$.*
Ipoteza H_1: *media lui X este $m < 8,33$.*

Ipoteza H_0 ocupă locul central. Se testează ***valabilitatea ipotezei*** H_0 ***împotriva valabilităţii ipotezei alternative*** H_1, astfel încât, la finalul testării, să fie acceptată ipoteza H_0, sau să fie respinsă ipoteza H_0 în favoarea ipotezei H_1.

Noţiuni generale despre teste statistice

Testul statistic reprezintă operatia de comparare a două ipoteze statistice alternative pe baza informaţiilor furnizate de sondaj (selecţie).

Un test statistic se numeşte:

1. **parametric**, dacă se presupune că X urmează o lege teoretică cunoscută, iar ipotezele statistice se referă la estimarea parametrilor legii (distribuţiei) respective;

2. **neparametric**, dacă ipotezele statistice se referă la estimarea distribuţiei variabilei empirice X sau a unor caracteristici numerice ale acesteia (medie, dispersie etc), în absenţa unei informaţii privind legea teoretică urmată de X.

Un test statistic presupune definirea unei **statistici de selecţie**, adică a unui şir de variabile aleatoare $g_n(X_1, X_2, \cdots, X_n)$, $n \in \mathbb{N}^*$, unde $g = (g_n)_{n \geq 1}$ este un şir de funcţii reale, iar $X_1, X_2, \ldots, X_n, \ldots$ sunt variabilele de selecţie, considerate independente şi identic distribuite. Anumite valori ale statisticii conduc la acceptarea ipotezei H_0, iar alte valori ale statisticii conduc la respingerea acestei ipoteze. Vom vorbi deci despre un **domeniu de respingere** (de neacceptare). Regula de decizie este dictată de specificarea domeniului de respingere al ipotezei H_0. Domeniul complementar domeniului de respingere reprezintă **domeniul de acceptare pentru ipoteza** H_0.

Ipoteza H_0, numită **ipoteza nulă**, se consideră **acceptabilă** dacă, în urma unei selecţii concrete $X_1 = x_1,\ X_2 = x_2, \cdots,\ X_n = x_n$, valoarea numerică rezultată $g_n(x_1, x_2, \cdots, x_n)$ a statisticii g aparţine domeniului de acceptare. Domeniul de acceptare este o mulţime

$A \subset \mathbb{R}$ stabilită teoretic prin luarea în considerare a ipozezei nule H_0. Astfel, dacă presupunem că ipoteza nulă H_0 (considerată ipoteză simplă) este valabilă, atunci distribuţia variabilei aleatoare $g_n(X_1, X_2, \cdots, X_n)$, unde $X_1, X_2, \cdots, X_n$ devin variabile aleatoare i.i.d., cu distribuţia comună cunoscută, este complet determinată.

Fie o constantă $\varepsilon \in (0,1)$ care va fi denumită în continuare **prag de semnificaţie** pentru statistica g. **Domeniul de acceptare de prag ε** este mulţinea $A_\varepsilon \subset \mathbb{R}$ definită prin condiţia:

$$\mathbb{P}\{g_n(X_1, X_2, \cdots, X_n) \in \mathbb{R} \setminus A_\varepsilon\} = \varepsilon,$$

sau, echivalent,

$$\mathbb{P}\{g_n(X_1, X_2, \cdots, X_n) \in A_\varepsilon\} = 1 - \varepsilon.$$

Mulţimea A_ε se interpretează ca o "zonă de o mare probabilitate teoretică". Domeniul de acceptare A_ε depinde de distribuţia variabilelor X_i (presupusă cunoscută), statistica g (definită), precum şi de parametrul ε (pragul de semnificaţie).

Etapele studiului.

1. Definirea ipotezei nule H_0 şi a ipotezei alternative H_1.

2. Definirea statisticii $g_n(X_1, \cdots, X_n)$ de estimare a unui parametru al distribuţiei comune a variabilelor aleatoare X_i.

3. Determinarea distribuţiei variabilei aleatoare $g_n(X_1, \cdots, X_n)$. Dacă variabila aleatoare $g_n(X_1, \cdots, X_n)$ este de tip continuu, se determină funcţia sa de densitate. Astfel, în ipoteza simplă H_0, statistica g este astfel o variabilă aleatoare cu distribuţie unic determinată.

4. Stabilirea pragului de semnificatie ε pentru variabila aleatoare $g_n(X_1, \cdots, X_n)$.

5. Determinarea domeniului de acceptare A_ε şi a domeniului de respingere $\mathbb{R} \setminus A_\varepsilon$ a ipotezei nule H_0. Adeseori aceste domenii se reprezintă grafic.

6. Calculul valorii numerice $g_n(x_1, x_2, \cdots, x_n)$ a statisticii g pentru o selecţie efectivă. Validarea ipotezei H_0 (cu pragul de semnificaţie ε) are loc dacă

$$g_n(x_1, x_2, \cdots, x_n) \in A_\varepsilon,$$

respectiv validarea ipotezei alternative H_1 are loc dacă

$$g_n(x_1, x_2, \cdots, x_n) \in \mathbb{R} \setminus A_\varepsilon.$$

Prezentăm detaliat şi exemplificăm două teste statistice importante: testul Z (parametric) şi testul χ^2 (neparametric).

4.6.2 Testul Z

Testul Z este un test parametric care se referă la estimarea mediei unei caracteristici X a unei populaţii P. Se presupune că X are o distribuţie normală, cu dispersia σ^2 cunoscută.

Presupunem aprioric că o caracteristică X a populaţiei P are o distribuţie normală, de medie μ (necunoscută) şi dispersie σ^2 (cunoscută). Considerm̆ variabilele aleatoare de selecţie $X_1,\ X_2, \ldots,\ X_n$. Notăm $\overline{X}_n = \dfrac{X_1 + X_2 + \ldots + X_n}{n}$.

1. Ne propunem să testăm ipoteza H_0: $\mu = \mu_0$, unde numărul μ_0 este specificat. Definim ipoteza alternativă $H_1 :\ \mu \neq \mu_0$

2. Consideram statistica $Z = (Z_n)_{n\geq 1}$ definită prin
$$Z_n = g_n(X_1, \cdots, X_n) = \frac{\sqrt{n}\left(\overline{X}_n - \mu_0\right)}{\sigma},\ n \in \mathbb{N}^*.$$
Conform Propoziţiei 4.5.1, Z_n este o variabilă aleatoare normală standard.

3. Stabilim pragul de semnificaţie ε, unde $0 < \varepsilon < 0,1$.

4. Determinăm domeniul de acceptare A_ε, astfel încât
$$\mathbb{P}\{Z_n \in A_\varepsilon\} = \mathbb{P}\{g_n(X_1, X_2, \cdots, X_n) \in A_\varepsilon\} = 1 - \varepsilon.$$
Astfel, notăm $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_0^x e^{-t^2/2} dt,\ x \in \mathbb{R}$.
Stabilim (din tabelul valorilor funcţiei Φ) **cuantila** $a_\varepsilon > 0$ astfel ca $\Phi\left(a_\varepsilon\right) = \dfrac{1-\varepsilon}{2}$.
Atunci domeniul de acceptare este $A_\varepsilon = \left[-a_\varepsilon, a_\varepsilon\right]$, deoarece
$$\mathbb{P}\{Z \in A_\varepsilon\} = \frac{1}{\sqrt{2\pi}} \int_{-a_\varepsilon}^{a_\varepsilon} e^{-t^2/2} dt = 2\Phi\left(a_\varepsilon\right) = 2 \cdot \frac{1-\varepsilon}{2} = 1 - \varepsilon.$$
Cuantila a_ε poate fi obţinută utilizând un soft de statistică (de ex. *Scilab*) din condiţia
$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a_\varepsilon} e^{-t^2/2} dt = 1 - \frac{\varepsilon}{2}.$$

5. Calculăm valoarea numerică $Z_n(\overline{x}) = \dfrac{\sqrt{n}\left(\overline{x} - \mu_0\right)}{\sigma}$ pentru selecţia (experimentală) $x_1, \cdots, x_n$, cu $\overline{x} = \dfrac{x_1 + \cdots + x_n}{n}$. Decidem că ipoteza H_0 este acceptată, cu pragul de semnificaţie ε, dacă $Z_n(\overline{x}) \in A_\varepsilon$, respectiv respinsă în favoarea ipotezei H_1 dacă $Z_n(\overline{x}) \in \mathbb{R} \setminus A_\varepsilon$.

Exemplu.
Fie populaţia P alcătuită din studenţii unui an care au transmis rezolvările la un test. Volumul populaţiei P este $N = 144$. Fie X caracteristica numerică: nota obţinută la test.

Presupunem aprioric că X are o distribuţie normală, de medie μ (necunoscută) şi dispersie $\sigma^2 = 2.1833014$ (cunoscută).

Valoarea atribuită dispersiei a fost obţinută prin calculul dispersiei empirice a caracteristicii X pentru întreaga populaţie P:

$$\sigma^2(X) = \frac{1}{N}\sum_{i=1}^{N}\left(x_i - \frac{\sum_{j=1}^{N} x_j}{N}\right)^2 = 2,1833014,$$

pentru setul de date reprezentând notele obţinute la test, prezentat mai jos:
Catalog note=
7.75, 6, 6.75, 7, 7.5, 4.25, 5.75, 3, 7.75, 8, 9, 8.75, 3.25, 7.25, 7.5, 6.75, 8.25, 3.5, 5.5, 4.5, 7.5, 4, 5.75, 7.5, 7.75, 5.75, 6.5, 6.75, 8.5, 7.75, 8.25, 7.5, 7.25, 5.25, 6.5, 8, 8, 4, 7.5, 6.75, 8.5, 6.75, 7.75, 7.5, 3.5, 5.5, 6.75, 4.75, 7, 7.75, 6, 7.25, 5.75, 5, 5.75, 9, 7.25, 5.5, 5.25, 5.5, 4.5, 4.75, 8, 8.5, 4.25, 6.75, 5, 3, 6.25, 6.75, 4, 7.25, 6.25, 6.75, 8.5, 8, 8.5, 6, 6.25, 7.75, 6, 4.5, 9, 7, 8.75, 5.5, 6, 4.75, 4.75, 5.5, 6.75, 6, 8.25, 6.5, 6.75, 5.5, 7.25, 6.75, 5.5, 8, 8.75, 8.5, 7, 5.5, 5.25, 5.5, 8.5, 6.25, 8.25, 8, 7, 5.75, 7, 7.25, 6, 4, 5.5, 7, 5, 7.75, 4, 5, 7, 6.75, 8, 8.75, 8.25, 8.75, 7.75, 6, 4.75, 5.75, 7, 6.25, 6, 8.75, 3.5, 7.75, 5.25,5.75, 8.25, 7.75, 5.5, 8.75.
Astfel, avem $\sigma = \sqrt{2.1833014} = 1.4775999$.

1. Utilizând **testul Z**, ne propunem să testăm ipoteza H_0: $\mu = 6.55$.
   Definim ipoteza alternativă $H_1: \ \mu \neq 6.55$

2. Considerăm statistica $Z = (Z_n)_{n\geq 1}$ definită prin

$$Z_n = g_n(X_1, \cdots, X_n) = \frac{\sqrt{n}\left(\overline{X}_n - 6.55\right)}{1.4775999} \sim N(0,1).$$

3. Stabilim pragul de semnificaţie $\varepsilon = 0.02$.

4. Determinăm domeniul de acceptare A_ε, astfel încât

$$\mathbb{P}\{Z_n \in A_\varepsilon\} = \mathbb{P}\{g_n(X_1, X_2, \cdots, X_n) \in A_\varepsilon\} = 1 - \varepsilon.$$

   Utilizând softul Scilab, determinăm cuantila a_ε din condiţia

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a_\varepsilon} e^{-t^2/2}dt = 1 - \frac{\varepsilon}{2} = 0.99.$$

   Găsim $a_\varepsilon = 2.3263479$. Ca urmare, domeniul de acceptare al ipotezei nule este:

$$A_\varepsilon = [-2.3263479, 2.3263479].$$

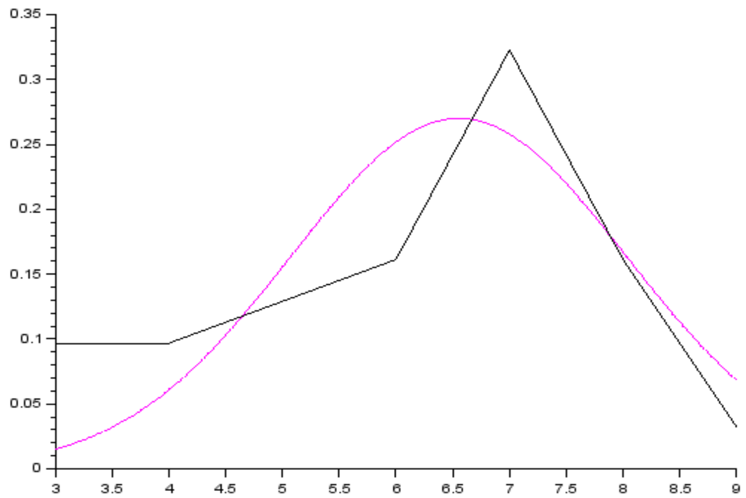
5. Fie selecţia de volum $n = 31$ formată din primele 31 de note din catalog.
   Calculăm valoarea numerică $Z_{31}(\overline{x}) = \dfrac{\sqrt{31}\,(\overline{x} - 6.55)}{1.4775999}$ pentru selecţia $x_1, \cdots, x_{31}$, având media empirică $\overline{x} = \dfrac{x_1 + \cdots + x_{31}}{31} = 6.5806452$. Obţinem

$$Z_{31}(\overline{x}) = 0.1154745 \in A_\varepsilon.$$

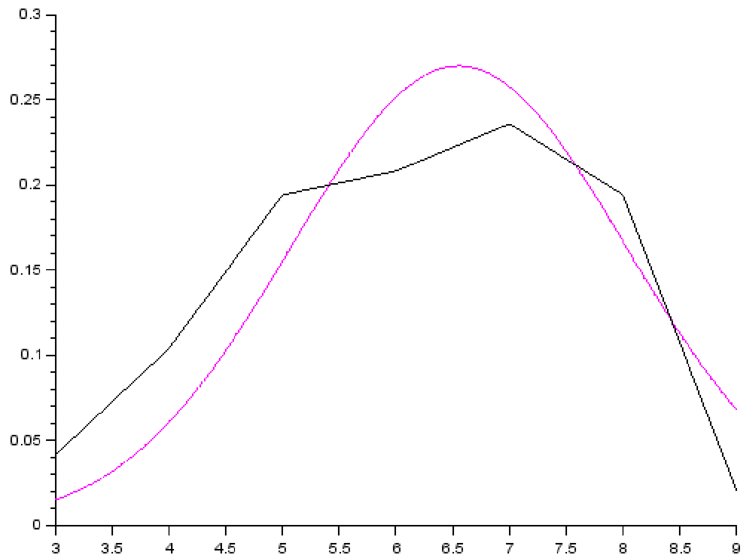
   Decidem că ipoteza H_0 este acceptată, cu pragul de semnificaţie $\varepsilon = 0.02$.

Ilustrări grafice.

1. Aproximarea normală $N(6.55, 2, 1833014)$ a frecvenţei relative a notelor selecţiei.



2. Aproximarea normală $N(6.5538194, 2, 1833014)$ a frecvenţei relative a notelor populaţiei.



4.6.3 Testul χ^2

Presupunem că o caracteristică X a unei populaţii P ia valori în mulţimea $\{v_1, v_2, \cdots, v_k\}$. **Testul χ^2 a lui Pearson** este un test neparametric care validează sau invalidează ipoteza nulă H_0:

$$\textit{caracteristica } X \textit{ are distribuţia } \; X : \begin{pmatrix} v_1 & v_2 & \cdots & v_k \\ p_1 & p_2 & \cdots & p_k \end{pmatrix},$$

unde $p_j = \mathbb{P}\{X = v_j\} > 0,\ j = 1, 2, \cdots, k$, cu $\sum_{j=1}^{k} p_j = 1$.

Rezultate teoretice

O variabilă aleatoare Z are distribuţia χ^2 cu k grade de libertate, unde $k \in \mathbb{N}^*$, dacă este pozitivă şi admite funcţia de densitate:

$$f_Z(x) = \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} x^{\frac{k}{2}-1} e^{-\frac{x}{2}},\ x > 0.$$

Astfel, Z are o distribuţie de tip Γ de parametri $n = \dfrac{k}{2}$ şi $a = \dfrac{1}{2}$, notat prin $Z \in \Gamma\left(\frac{k}{2}, \frac{1}{2}\right)$.

Testul χ^2 se bazează pe următoarele rezultate teoretice.

Teorema 4.6.1. *Dacă $Y_1, Y_2, \cdots, Y_k$ sunt k variabile aleatoare independente, cu distribuţie normală standard (deci $Y_i \sim N(0,1),\ i = 1, 2, \cdots, k$), atunci variabila aleatoare*

$$Z = \sum_{i=1}^{k} Y_i^2$$

are o distribuţie χ^2 cu k grade de libertate.

Demonstraţie. Fie $Y \sim N(0,1)$ a variabilă aleatoare normală standard. Determinăm functia de densitate f a variabilei aleatoare Y^2. Variabila aleatoare Y^2 este pozitivă şi are funcţia de repartiţie:

$$F(t) = \mathbb{P}\{Y^2 \le t\} = \mathbb{P}\{-\sqrt{t} \le Y \le \sqrt{t}\} = \frac{1}{\sqrt{2\pi}} \int_{-\sqrt{t}}^{\sqrt{t}} e^{-\frac{x^2}{2}}\, dx = \sqrt{\frac{2}{\pi}} \int_0^{\sqrt{t}} e^{-\frac{x^2}{2}}\, dx,\ \forall\, t \ge 0.$$

Atunci

$$f(t) = F'(t) = \sqrt{\frac{2}{\pi}} e^{-\frac{(\sqrt{t})^2}{2}} (\sqrt{t})' = \frac{1}{\sqrt{2\pi t}} e^{-\frac{t}{2}} = \frac{1}{2^{\frac{1}{2}}\Gamma\left(\frac{1}{2}\right)} t^{\frac{1}{2}-1} e^{-\frac{t}{2}},\ \forall\ t > 0.$$

Rezultă $Y^2 \in \Gamma\left(\frac{1}{2}, \frac{1}{2}\right)$, deci Y^2 are o distribuţie χ^2 cu 1 grad de libertate.

Conform Propoziţiei 2.4.1, dacă variabilele aleatoare U şi V sunt independente, având funcţiile de repartiţie F şi respectiv G, atunci funcţia de repartiţie a variabilei aleatoare $U + V$, notată $F * G$ şi denumită convoluţia funcţiilor de repartiţie F şi G, este dată de:

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

Dacă U şi V sunt variabile aleatoare de tip continuu, cu funcţiile de densitate $f = F'$ şi $g = G'$, atunci variabila aleatoare $U + V$ are funcţia de densitate $f * g = (F * G)'$, obţinută astfel:

$$(f * g)(t) = \frac{d}{dt} \int_{-\infty}^{\infty} F(t - x) dG(x) = \int_{-\infty}^{\infty} \frac{\partial}{\partial t}[F(t - x) g(x)] dx = \int_{-\infty}^{\infty} f(t - x) g(x)\, dx.$$

În particular, dacă U şi V sunt variabile aleatoare pozitive, atunci $U+V$ este de asemenea o variabilă aleatoare pozitivă. Pentru oricare $t>0$, din condiţia $f(x)=g(x)=0,\ \forall\, x<0$, obţinem

$$(f*g)(t)=\int_{-\infty}^{0} f(t-x)g(x)\,dx+\int_{0}^{t} f(t-x)g(x)\,dx+\int_{t}^{\infty} f(t-x)g(x)\,dx=\int_{0}^{t} f(t-x)g(x)\,dx.$$

Fie $(Y_k)_{k\geq 1}$ un şir de variabile aleatoare normale standard, independente. Demonstrăm prin inducţie după $k\in\mathbb{N}^*$ că $\sum_{i=1}^{k} Y_i^2\in\Gamma\left(\frac{k}{2},\frac{1}{2}\right)$. Proprietatea a fost demonstrată pentru $k=1$. Presupunem că, pentru un număr natural nenul k avem $\sum_{i=1}^{k} Y_i^2\in\Gamma\left(\frac{k}{2},\frac{1}{2}\right)$. Notăm f_k funcţia de densitate a variabilei aleatoare $\sum_{i=1}^{k} Y_i^2$. Variabila aleatoare Y_{k+1}^2 are funcţia de densitate f, descrisă anterior. Atunci, datorită independenţei variabilelor aleatoare pozitive $\sum_{i=1}^{k} Y_i^2$ şi Y_{k+1}^2, variabila aleatoare $\sum_{i=1}^{k+1} Y_i^2$ are funcţia de densitate $f_{k+1}=f*f_k$. Avem

$$f_{k+1}(t)=(f*f_k)(t)=\int_0^t f(t-x)f_k(x)\,dx=\int_0^t\left[\frac{1}{\sqrt{2\pi(t-x)}}e^{-\frac{t-x}{2}}\right]\cdot\left[\frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)}x^{\frac{k}{2}-1}e^{-\frac{x}{2}}\right]dx=$$

$$=\frac{e^{-\frac{t}{2}}}{2^{\frac{k+1}{2}}\Gamma\left(\frac{k}{2}\right)\sqrt{\pi}}\int_0^t x^{\frac{k}{2}-1}(t-x)^{\frac{1}{2}-1}dx\overset{x=ty}{=}\frac{e^{-\frac{t}{2}}}{2^{\frac{k+1}{2}}\Gamma\left(\frac{k}{2}\right)\sqrt{\pi}}\int_0^1 t^{\frac{k+1}{2}-1}y^{\frac{k}{2}-1}(1-y)^{\frac{1}{2}-1}dy=$$

$$=\frac{t^{\frac{k+1}{2}-1}e^{-\frac{t}{2}}}{2^{\frac{k+1}{2}}\Gamma\left(\frac{k}{2}\right)\sqrt{\pi}}\,\beta\left(\frac{k}{2},\frac{1}{2}\right)=\frac{t^{\frac{k+1}{2}-1}e^{-\frac{t}{2}}}{2^{\frac{k+1}{2}}\Gamma\left(\frac{k}{2}\right)\sqrt{\pi}}\,\frac{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{k+1}{2}\right)}=\frac{t^{\frac{k+1}{2}-1}e^{-\frac{t}{2}}}{2^{\frac{k+1}{2}}\Gamma\left(\frac{k+1}{2}\right)},\ \forall\, t>0.$$

Rezultă că $\sum_{i=1}^{k+1} Y_i^2$ are distribuţia χ^2 cu $k+1$ grade de libertate, adică $\sum_{i=1}^{k+1} Y_i^2\in\Gamma\left(\frac{k+1}{2},\frac{1}{2}\right)$.

Astfel, teorema este demonstrată. □

Definiţia 4.6.1. *Matricea de covarianţă asociată unui vector aleator* $\mathbf{X}=(X_1,X_2,\cdots,X_k)$ *se defineşte prin*

$$\Sigma=(\lambda_{i,j})_{i,j=1,\cdots,k}=\begin{pmatrix}\lambda_{1,1} & \lambda_{1,2} & \cdots & \lambda_{1,k}\\ \lambda_{2,1} & \lambda_{2,2} & \cdots & \lambda_{2,k}\\ \cdots & \cdots & \cdots & \cdots\\ \lambda_{k,1} & \lambda_{k,1} & \cdots & \lambda_{k,k}\end{pmatrix}\in\mathcal{M}_k(\mathbb{R}),$$

unde $\lambda_{i,j}=\lambda(X_i,X_j)=\mathbb{E}(X_iX_j)-\mathbb{E}(X_i)\mathbb{E}(X_j)$ *este corelaţia variabilelor aleatoare* X_i *şi* X_j*, pentru* $i,j\in\{1,2,\cdots,k\}$.

Matricea de covarianţă asociată vectorului aleator $\mathbf{X}=(X_1,\cdots,X_k)$ este simetrică şi semipozitiv definită, adică toţi minorii săi principali sunt pozitivi. Pe diagonala principală se află dispersiile componentelor vectorului X:

$$\lambda_{i,i}=\lambda(X_i,X_i)=\mathbb{E}(X_i^2)-\mathbb{E}^2(X_i)=V(X_i),\ i=1,2,\cdots,k.$$

Definiţia 4.6.2. *Un vector aleator* $\mathbf{Z} = (Z_1, Z_2, \cdots, Z_k)$*, cu media* $\boldsymbol{\mu}= (\mu_1, \mu_2, \cdots, \mu_k)$*, unde* $\mu_i = \mathbb{E}(Z_i)$*, având matricea de covarianţă* Σ *pozitiv definită, adică toţi minorii săi principali sunt strict pozitivi, are o distribuţie mutivariată normală dacă admite funcţia de densitate*

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^k \det(\Sigma)}} e^{-\frac{1}{2}(\mathbf{x}-\mu)\Sigma^{-1}(\mathbf{x}-\mu)^T}, \ \forall\ \mathbf{x} = (x_1, x_2, \cdots, x_k) \in \mathbb{R}^k.$$

Notăm $\mathbf{Z} \sim N_k(\boldsymbol{\mu}, \Sigma)$*.*

În particular, dacă $\Sigma = I_k$ *(matricea unitate de ordinul k) şi* $\boldsymbol{\mu}= \mathbf{0}$*, atunci*

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^k}} e^{-\frac{1}{2}\mathbf{x}\mathbf{x}^T} = \frac{1}{\sqrt{(2\pi)^k}} e^{-\frac{1}{2}\sum_{i=1}^k x_i^2} = \prod_{i=1}^k \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}} \right), \ \forall\ \mathbf{x} = (x_1, \cdots, x_k) \in \mathbb{R}^k.$$

Notăm $\mathbf{Z} \sim N_k(\mathbf{0}, I_k)$ *şi spunem ca vectorul aleator* $\mathbf{Z}$ *are o distribuţie mutivariată normală standard.*

Dacă $\Sigma^* = (a_{i,j})_{i,j=\overline{1,k}}$ este adjuncta matricei de corelaţie Σ, atunci avem

$$(\mathbf{x}-\mu)\Sigma^{-1}(\mathbf{x}-\mu)^T = \frac{1}{\det(\Sigma)}(\mathbf{x}-\mu)\Sigma^*(\mathbf{x}-\mu)^T =$$

$$= \frac{1}{\det(\Sigma)} \left[\sum_{i=1}^k a_{i,i}(x_i-\mu_i)^2 + 2 \sum_{1\leq i<j\leq k} a_{i,j}(x_i-\mu_i)(x_j-\mu_j) \right].$$

Dacă $\Sigma = I_k$, atunci variabilele aleatoare $Z_1, Z_2, \cdots, Z_k$ sunt necorelate şi au o distribuţie normală standard.

Teorema 4.6.2 (Teorema limită centrală multidimensională). *Fie* $(\mathbf{X}_n)_{n\geq 1}$ *un şir de vectori aleatori* $\mathbf{X}_n = (X_{n,1}, X_{n,2}, \cdots, X_{n,k})$ *de dimensiune k, independenţi şi identic distribuiţi, cu media comună* $\boldsymbol{\mu}= (\mu_1, \mu_2, \cdots, \mu_k)$ *şi matricea de covarianţă comună* Σ*. Presupunem că* Σ *este pozitiv definită şi considerăm matricea simetrică* $\sqrt{\Sigma^{-1}} \in \mathcal{M}_k(\mathbb{R})$ *cu proprietăţile* $\det\left(\sqrt{\Sigma^{-1}}\right) > 0$ *şi* $\left(\sqrt{\Sigma^{-1}}\right)^2 = \Sigma^{-1}$*. Fie* $\overline{\mathbf{X}}_n = \dfrac{1}{n}\displaystyle\sum_{i=1}^n \mathbf{X}_i,\ n \in \mathbb{N}^*$*. Atunci*

$$\sqrt{n}(\overline{\mathbf{X}}_n - \boldsymbol{\mu})\sqrt{\Sigma^{-1}} \ \xrightarrow{d} \mathbf{Z},$$

unde $\mathbf{Z} \sim N_k(\mathbf{0}, I_k)$ *este un vector aleator cu o distribuţie mutivariată normală standard.*

Teorema 4.6.3 (Formula Sherman-Morrison). *Fie* $k \in \mathbb{N},\ k \geq 2,\ A \in \mathcal{M}_k(\mathbb{R})$ *o matrice invertibilă, iar* $u, v \in \mathcal{M}_{1,k}$ *doua matrice linie. Atunci matricea* $B = A - u^T v$ *din* $\mathcal{M}_k(\mathbb{R})$ *este invertibilă dacă şi numai dacă* $vA^{-1}u^T \neq 1$*. În caz afirmativ,*

$$B^{-1} = A^{-1} + \frac{1}{1 - vA^{-1}u^T} A^{-1}u^T v A^{-1}.$$

Demonstraţie. Avem

$$(A - u^T v)\left(A^{-1} + \frac{1}{1 - vA^{-1}u^T} A^{-1}u^T v A^{-1}\right) =$$

$$= I_k - u^T vA^{-1} + \frac{1}{1 - vA^{-1}u^T}\left(u^T vA^{-1}\right) - \frac{1}{1 - vA^{-1}u^T}\left(u^T vA^{-1}u^T vA^{-1}\right) =$$

$$= I_k + \frac{1}{1 - vA^{-1}u^T}\left[-u^T vA^{-1} + (vA^{-1}u^T)u^T vA^{-1} + u^T vA^{-1} - u^T\left(vA^{-1}u^T\right)vA^{-1}\right] =$$

$$= I_k + \frac{vA^{-1}u^T}{1 - vA^{-1}u^T}\left(u^T vA^{-1} - u^T vA^{-1}\right) = I_k. \quad \square$$

Teorema 4.6.4 (**Baza teoretică al testului χ^2**). *Fie* $\mathbf{X} = (X_1, X_2, \cdots, X_k)$ *un vector aleator care ia valorile* $(\underbrace{1, 0, 0 \cdots, 0}_{k})$, $(\underbrace{0, 1, 0, \cdots, 0}_{k}), \cdots, (\underbrace{0, 0, 0, \cdots, 1}_{k})$ *cu probabilităţile strict pozitive* $p_1, p_2, \cdots, p_k$, *astfel ca* $\sum_{i=1}^{k} p_i = 1$. *Prin urmare, variabila aleatoare* X_i *ia valori în mulţimea* $\{0, 1\}$, *pentru* $i = 1, 2, \cdots, k$, *astfel ca* $\sum_{i=1}^{k} X_i = 1$. *Avem*

$$\mathbb{P}\{X_i = 1\} = \mathbb{P}\left\{\mathbf{X} = (\underbrace{0, \cdots, \overset{i}{1}, \cdots, 0}_{k})\right\} = p_i > 0, \quad i = 1, 2, \cdots, k.$$

Considerăm un şir $(\mathbf{X}_n)_{n\geq 1}$ *de vectori aleatori, independenţi şi identic distribuiţi, având distribuţia lui* $\mathbf{X}$. *Fie* $\mathbf{X}_n = (X_{n,1}, X_{n,2}, \cdots, X_{n,k})$, $n \in \mathbb{N}^*$. *Notăm*

$$S_{n,i} = \sum_{m=1}^{n} X_{m,i}, \ i = 1, 2, \cdots, k, \ n \in \mathbb{N}^*.$$

Definim şirul de variabile aleatoare

$$\Phi_n = \frac{(S_{n,1} - np_1)^2}{np_1} + \frac{(S_{n,2} - np_2)^2}{np_2} + \cdots + \frac{(S_{n,k} - np_k)^2}{np_k}, \ n \in \mathbb{N}^*.$$

Atunci are loc convergenţa în distribuţie $\Phi_n \xrightarrow{d} Z$, *unde* Z *este o variabilă aleatoare cu distribuţia* χ^2 *cu* $k - 1$ *grade de libertate.*

Demonstraţie. Determinăm matricea de covarianţă $\Sigma \in \mathcal{M}_k(\mathbb{R})$ a vectorului aleator $\mathbf{X} = (X_1, X_2, \cdots, X_k)$. Variabilele X_i au o distribuţie Bernoulli. Astfel

$$X_i : \begin{pmatrix} 0 & 1 \\ 1 - p_i & p_i \end{pmatrix}, \ i = 1, \cdots, k.$$

Atunci $V(X_i) = \mathbb{E}(X_i^2) - \mathbb{E}^2(X_i) = p_i - p_i^2$, $i = 1, \cdots, k$. Corelaţia variabilelor X_i şi X_j, unde $i \neq j$, este $\lambda(X_i, X_j) = \mathbb{E}(X_i X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j) = -p_i p_j$, deoarece $X_i X_j = 0$. Rezultă

$$\Sigma = \begin{pmatrix} p_1 - p_1^2 & -p_1 p_2 & \cdots & -p_1 p_k \\ -p_2 p_1 & p_2 - p_2^2 & \cdots & -p_2 p_k \\ \cdots & \cdots & \cdots & \cdots \\ -p_k p_1 & -p_k p_2 & \cdots & p_k - p_k^2 \end{pmatrix}.$$

Considerăm matricea $\Sigma^{(1)}$ formată din primele $k-1$ linii şi coloane ale matricei Σ:

$$\Sigma^{(1)} = \begin{pmatrix} p_1 - p_1^2 & -p_1p_2 & \cdots & -p_1p_{k-1} \\ -p_2p_1 & p_2 - p_2^2 & \cdots & -p_2p_{k-1} \\ \cdots & \cdots & \cdots & \cdots \\ -p_{k-1}p_1 & -p_{k-1}p_2 & \cdots & p_{k-1} - p_{k-1}^2 \end{pmatrix}.$$

$\Sigma^{(1)}$ reprezintă matricea de covarianţă a vectorului aleator $\mathbf{X}^{(1)} = (X_1, X_2, \cdots, X_{k-1})$, cu media $\mathbf{p}^{(1)} = (p_1, p_2, \cdots, p_{k-1})$. Matricea $\Sigma^{(1)} \in \mathcal{M}_{k-1}(\mathbb{R})$ se poate rescrie

$$\Sigma^{(1)} = \begin{pmatrix} p_1 & 0 & \cdots & 0 \\ 0 & p_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & p_{k-1} \end{pmatrix} - \begin{pmatrix} p_1 \\ p_2 \\ \cdots \\ p_{k-1} \end{pmatrix} \begin{pmatrix} p_1 & p_2 & \cdots & p_{k-1} \end{pmatrix}.$$

Fie $A = \begin{pmatrix} p_1 & 0 & \cdots & 0 \\ 0 & p_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & p_{k-1} \end{pmatrix}$ şi $u = \begin{pmatrix} p_1 & p_2 & \cdots & p_{k-1} \end{pmatrix}$. Astfel, $\Sigma^{(1)} = A - u^T u$.

Matricea A este invertibilă, cu inversa $A^{-1} = \begin{pmatrix} 1/p_1 & 0 & \cdots & 0 \\ 0 & 1/p_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1/p_{k-1} \end{pmatrix}$. Avem

$$uA^{-1}u^T = \begin{pmatrix} p_1 & p_2 & \cdots & p_{k-1} \end{pmatrix} \begin{pmatrix} 1/p_1 & 0 & \cdots & 0 \\ 0 & 1/p_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1/p_{k-1} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \cdots \\ p_{k-1} \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \cdots \\ p_{k-1} \end{pmatrix} = \sum_{i=1}^{k-1} p_i = 1 - p_k \neq 1.$$

Atunci, pe baza Teoremei Sherman-Morrison, matricea $\Sigma^{(1)}$ este invertibilă (şi pozitiv definită), cu inversa

$$\left(\Sigma^{(1)}\right)^{-1} = A^{-1} + \frac{1}{1 - uA^{-1}u^T} A^{-1}u^T u A^{-1} = A^{-1} + \frac{1}{1 - uA^{-1}u^T} \left(A^{-1}u^T\right)\left(uA^{-1}\right) =$$

$$= \begin{pmatrix} 1/p_1 & 0 & \cdots & 0 \\ 0 & 1/p_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1/p_{k-1} \end{pmatrix} + \frac{1}{p_k} \begin{pmatrix} 1 \\ 1 \\ \cdots \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} 1/p_1 + 1/p_k & 1/p_k & \cdots & 1/p_k \\ 1/p_k & 1/p_2 + 1/p_k & \cdots & 1/p_k \\ \cdots & \cdots & \cdots & \cdots \\ 1/p_k & 1/p_k & \cdots & 1/p_{k-1} + 1/p_k \end{pmatrix}.$$

Pe de altă parte, din relaţia $X_k = 1 - \sum_{i=1}^{k-1} X_i$ obţinem $S_{n.k} = n - \sum_{i=1}^{k-1} S_{n,i}$. Apoi avem $np_k = n\left(1 - \sum_{i=1}^{k-1} p_i\right) = n - \sum_{i=1}^{k-1} np_i$. Atunci

$$\Phi_n = \sum_{i=1}^{k} \frac{(S_{n,i} - np_i)^2}{np_i} = \sum_{i=1}^{k-1} \frac{(S_{n,i} - np_i)^2}{np_i} + \frac{\left[\sum_{i=1}^{k-1}(np_i - S_{n,i})\right]^2}{np_k} =$$

$$= n\left\{\sum_{i=1}^{k-1} \frac{(S_{n,i}/n - p_i)^2}{p_i} + \frac{\left[\sum_{i=1}^{k-1}(S_{n,i}/n - p_i)\right]^2}{p_k}\right\} =$$

$$= n\left[\sum_{i=1}^{k-1}\left(\frac{S_{n,i}}{n} - p_i\right)^2\left(\frac{1}{p_i} + \frac{1}{p_k}\right) + 2\sum_{1 \le i < j \le k-1}\left(\frac{S_{n,i}}{n} - p_i\right)\left(\frac{S_{n,j}}{n} - p_j\right)\frac{1}{p_k}\right].$$

Asociem şirului $\mathbf{X}_n = (X_{n,1}, X_{n,2}, \cdots, X_{n,k})$ şirul $\mathbf{X}_n^{(1)} = (X_{n,1}, X_{n,2}, \cdots, X_{n,k-1}),\ n \in \mathbb{N}^*$, de vectori independenţi, cu distribuţia vectorului aleator $\mathbf{X}^{(1)} = (X_1, \ldots, X_{k-1})$. Fie

$$\overline{\mathbf{X}}_n^{(1)} = \frac{1}{n}\sum_{m=1}^{n} \mathbf{X}_m^{(1)} = \left(\frac{S_{n,1}}{n}, \frac{S_{n,2}}{n}, \cdots, \frac{S_{n,k-1}}{n}\right),\ n \in \mathbb{N}^*.$$

Avem

$$\overline{\mathbf{X}}_n^{(1)} - \mathbf{p}^{(1)} = \left(\frac{S_{n,1}}{n} - p_1, \frac{S_{n,2}}{n} - p_2, \cdots, \frac{S_{n,k-1}}{n} - p_{k-1}\right),\ n \in \mathbb{N}^*.$$

Din expresiile lui $\left(\Sigma^{(1)}\right)^{-1}$ şi Φ_n, rezultă următoarea reprezentare maticeala a variabilei aleatoare Φ_n:

$$\Phi_n = n\left(\overline{\mathbf{X}}_n^{(1)} - \mathbf{p}^{(1)}\right)\left(\Sigma^{(1)}\right)^{-1}\left(\overline{\mathbf{X}}_n^{(1)} - \mathbf{p}^{(1)}\right)^T =$$

$$= \left[\sqrt{n}\left(\overline{\mathbf{X}}_n^{(1)} - \mathbf{p}^{(1)}\right)\sqrt{\left(\Sigma^{(1)}\right)^{-1}}\right]\left[\sqrt{n}\left(\overline{\mathbf{X}}_n^{(1)} - \mathbf{p}^{(1)}\right)\sqrt{\left(\Sigma^{(1)}\right)^{-1}}\right]^T,$$

unde s-a utilizat proprietatea $\sqrt{\left(\Sigma^{(1)}\right)^{-1}} = \left(\sqrt{\left(\Sigma^{(1)}\right)^{-1}}\right)^T$, care decurge din relaţia de simetrie $\left(\Sigma^{(1)}\right)^{-1} = \left(\left(\Sigma^{(1)}\right)^{-1}\right)^T$. Conform Teoremei limită centrală multidimensională,

$$\sqrt{n}\left(\overline{\mathbf{X}}_n^{(1)} - \mathbf{p}^{(1)}\right)\sqrt{\left(\Sigma^{(1)}\right)^{-1}} \xrightarrow{d} (Z_1, Z_2, \cdots, Z_{k-1}),$$

unde $(Z_1, Z_2, \cdots, Z_{k-1}) \sim N_{k-1}(\mathbf{0}, I_{k-1})$. Rezultă

$$\Phi_n \xrightarrow{d} (Z_1, Z_2, \cdots, Z_{k-1})\begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_{k-1} \end{pmatrix} = \sum_{i=1}^{k-1} Z_i^2.$$

Cum $Z_1, Z_2, \cdots, Z_{k-1}$ sunt variabile aleatoare normale standard, necorelate, variabila aleatoare $Z = \sum_{i=1}^{k-1} Z_i^2$ are o distribuţie χ^2 cu $k-1$ grade de libertate (Teorema 4.6.1).

Astfel, teorema este demonstrată. □

Descrierea testului χ^2

O caracteristică X a unei populaţii P poate lua una din valorile $v_1, v_2, \cdots, v_k$. Fie numerele strict pozitive $p_1, p_2, \cdots, p_k$, cu $\sum_{i=1}^{k} p_i = 1$.

1. Testăm ipoteza nulă H_0: *caracteristica X are distribuţia* $\begin{pmatrix} v_1 & v_2 & \cdots & v_k \\ p_1 & p_2 & \cdots & p_k \end{pmatrix}$.

2. Fie variabilele de selecţie $X_1, X_2, \cdots, X_n$. Definim statistica

$$\Phi_n = g_n(X_1, X_2, \cdots, X_n) = \sum_{i=1}^{k} \frac{(S_{n,i} - np_i)^2}{np_i},$$

unde $S_{n,i} = |\{m \in \{1, \cdots, n\} : \ X_m = v_i\}|$, pentru $i = 1, 2, \cdots, k$.
Conform Teoremei 4.6.4, dacă ipoteza H_0 este adevărată, atunci are loc convergenţa în distribuţie: $g_n(X_1, X_2, \cdots, X_n) = \Phi_n \xrightarrow{d} Z$, unde Z este o variabilă aleatoare cu distribuţia χ^2 cu $k-1$ grade de libertate.

3. Stabilim pragul de semnificaţie $\varepsilon \in (0, 1/10)$.

4. Fie F funcţia de repartiţie a variabilei aleatoare Z. Determinăm domeniul de acceptare $A_\varepsilon = [0, a_\varepsilon]$ astfel ca $\mathbb{P}\{Z \in A_\varepsilon\} = \mathbb{P}\{Z \leq a_\varepsilon\} = F(a_\varepsilon) = 1 - \varepsilon$.

Pentru numărul gradelor de libertate $k-1 = d \in \{1, 2, \cdots, 20\}$ şi $\varepsilon \in \{0.05, 0.01, 0.001\}$, cuantila a_ε se citeşte pe linia d şi coloana ε în tabelul de mai jos.

Critical values of the Chi-square distribution with *d* degrees of freedom

	Probability of exceeding the critical value						
d	0.05	0.01	0.001	*d*	0.05	0.01	0.001
1	3.841	6.635	10.828	11	19.675	24.725	31.264
2	5.991	9.210	13.816	12	21.026	26.217	32.910
3	7.815	11.345	16.266	13	22.362	27.688	34.528
4	9.488	13.277	18.467	14	23.685	29.141	36.123
5	11.070	15.086	20.515	15	24.996	30.578	37.697
6	12.592	16.812	22.458	16	26.296	32.000	39.252
7	14.067	18.475	24.322	17	27.587	33.409	40.790
8	15.507	20.090	26.125	18	28.869	34.805	42.312
9	16.919	21.666	27.877	19	30.144	36.191	43.820
10	18.307	23.209	29.588	20	31.410	37.566	45.315

***INTRODUCTION TO POPULATION GENETICS*, Table D.1**

Pentru alte valori ale lui ε se poate utiliza un soft statistic. Reţinem că, pentru n "mare", are loc aproximarea:

$$\mathbb{P}\{g_n(X_1, X_2, \cdots, X_n) \in A_\varepsilon\} \approx \mathbb{P}\{Z \in A_\varepsilon\} = F(a_\varepsilon) = 1 - \varepsilon,$$

5. Dacă valoarea numerică a lui Φ_n calculată pentru datele concrete ale eşantionului $X_1, X_2, \cdots, X_n$ aparţine mulţimii $A_\varepsilon = [0, a_\varepsilon]$, atunci H_0 este acceptată cu pragul de semnificaţie ε. În caz contrar, ipoteza H_0 este respinsă.

Exemplu
Examenul de *Teoria probabilităţilor şi statistică matematică* din sesiunea de vară 2020, susţinut de 204 studenţi de la programele de studii *Electronică*, *Calculatoare* şi *Tehnologia informaţiei* a avut următoarea distribuţie a rezultatelor

$$\begin{pmatrix} v_1: \textit{ abs sau neprom.} & v_2: \textit{ nota } 5 \textit{ sau } 6 & v_3: \textit{ nota } 7 \textit{ sau } 8 & v_4: \textit{ nota } 9 \textit{ sau } 10 \\ p_1 = 0.17 & p_2 = 0.20 & p_3 = 0.44 & p_4 = 0.19 \end{pmatrix}.$$

În aceeaşi sesiune, la examenul de *Probabilităţi şi statistică*, susţinut de 73 de studenţi de la programul *Informatică* s-a înregistrat următoarea situaţie: 5 studenţi au fost absenţi (valoarea v_1), 14 studenţi au luat notele 5-6 (valoarea v_2), 31 studenţi au luat notele 7-8 (valoarea v_3) şi 23 de studenţi au luat notele 9-10 (valoarea v_4). Utilizând testul χ^2 ne propunem să stabilim valabilitatea ipotezei H_0:

> rezultatele înregistrate la examenul de de *Probabilităţi şi statistică* au aceeaşi distribuţie cu cele înregistrate la examenul de *Teoria probabilităţilor şi statistică matematică.*

Volumul eşantionului este $n = 73$, iar $k = 4$. Alegem pragul de semnificaţie $\varepsilon = 0.01$. Fie Z o variabilă aleatoare χ^2 cu $k - 1 = 3$ grade de libertate. Conform tabelului de mai sus, obţinem cuantila $a_\varepsilon = 11.345$, deci domeniul de acceptare a ipotezei nule este $A_\varepsilon = [0, 11.345]$. Pentru datele menţionate, avem

$$\Phi_{73} = \frac{(5 - 73 \cdot p_1)^2}{73 \cdot p_1} + \frac{(14 - 73 \cdot p_2)^2}{73 \cdot p_2} + \frac{(31 - 73 \cdot p_3)^2}{73 \cdot p_3} + \frac{(23 - 73 \cdot p_4)^2}{73 \cdot p_4} = 10.498086.$$

Ca urmare, ipoteza H_0 este acceptată (la limită!) cu pragul de semnificaţie $\varepsilon = 0.01$.

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 $\beta(a, b)$ 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 $\dfrac{\theta}{1-\theta}$ 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 $T = \overline{X}_n(1 - \overline{X}_n)$ nu este fără abatere, dar există un multiplu al său având această proprietate.

4.8 Rezumat

Statistica matematica utilizează rezultatele oferite de teoria probabilităţilor pentru a generealiza 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 defineşte 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:

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

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

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

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.