

CULEGERI
DE
PROBLEME
DE
MATEMATICĂ
ȘI
FIZICĂ

SERIE

MONICA DUMITRESCU

DOREL FLOREA

CONSTANTIN TUDOR

**PROBLEME
DE TEORIA PROBABILITĂȚILOR
ȘI STATISTICĂ MATEMATICĂ**

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

MONICA DUMITRESCU

BOREL FLOREA

CONSTANTIN TUDOR

PROBLEME DE TEORIA PROBABILITĂȚILOR ȘI STATISTICĂ MATEMATICĂ

Seria

**Culegeri de probleme
de matematică și fizică**



Editura Tehnică

București - 1985

Culegerea de probleme de teoria probabilităților și statistică matematică este un instrument deosebit de util în învățarea sistematică și eficace a teoriei probabilităților și a statisticii matematice.

Licrarea este alcătuită din patru capitole: cimp de probabilitate; repartiții și variabile aleatoare; convergența variabilelor aleatoare; elemente de statistică matematică.

Fiecare capitol începe cu un rezumat teoretic în care sînt listate definițiile și rezultatele de bază necesare rezolvării problemelor.

Soluțiile problemelor sînt elaborate în spirit modern, riguros și în același timp într-o manieră deosebit de accesibilă cititorilor.

Cartea este adresată studenților de la institutele tehnice, universități, precum și inginerilor, economiștilor, matematicienilor și tuturor celor interesați de teoria probabilităților și de statistică matematică.

Control științific: Prof. Dr. GH. CIUCU
Redactor: LINA TICOȘ
Tehnoredactor: ELENA GERU

Bun de tipar: 6.11.1985.
Coll de tipar: 19.
C.Z. 519.21+519.24/.27

INTREPRINDEREA POLIGRAFICĂ CLUJ
Municipiul Cluj-Napoca
B-dul Lenin nr. 146.
Republica Socialistă România
Comanda nr. 281



PREFAȚĂ

În ultimile decenii rolul teoriei probabilităților și al statisticii matematice în științele naturii a crescut considerabil. În același timp relația strînsă dintre aceste două ramuri ale matematicii moderne și cererile activităților practice a constituit un important factor al dezvoltării lor. Avînd un rol important în fizică, chimie, biologie, metodele probabilistice și statistice sînt utilizate cu succes în diferite ramuri ale tehnicii, în economie, în organizarea producției, în medicină, în geologie etc. Menționăm că în țara noastră și pe plan mondial în cercetările geologice sînt folosite curent metode de investigare care fac apel la teoria probabilităților și la statistica matematică.

Pentru a-și însuși cu succes bazele teoretice ale teoriei probabilităților și ale statisticii matematice cititorul are la dispoziție o serie de tratate și manuale valoroase (unele dintre ele menționate la bibliografie). Dar, pentru înțelegerea și aprofundarea conceptelor unei teorii matematice, pentru a le putea utiliza cu ușurință în probleme practice, este nevoie și de culegeri de probleme. În acest scop a fost elaborată și această lucrare.

Lucrarea de față este împărțită în patru capitole. Fiecare capitol începe cu un rezumat teoretic în care sînt prezentate definițiile și rezultatele de bază necesare rezolvării cu succes a problemelor. Soluțiile problemelor sînt elaborate în spirit modern, riguros, dar în același timp și într-o manieră accesibilă cititorilor interesați.

În primul capitol, printr-o gamă largă de probleme (unele dintre ele devenite clasice) cititorul este familiarizat cu noțiunea fundamentală de probabilitate, cu proprietățile sale importante. Problemele din capitolul II sînt selectate astfel încît conceptele cele mai utilizate din teoria probabilităților (repartiții, momente, funcții caracteristice) să devină accesibile și ușor de manevrat. În capitolul III sînt exemplificate tipurile de

convergență larg utilizate în teoria probabilităților iar apoi sînt prezentate probleme care ilustrează cele două mari teoreme ale teoriei probabilităților (legea numerelor mari și teorema limită centrală).

Problemele din capitolele I, II și III au fost redactate de Constantin Tudor și Dorel Florea.

Capitolul IV abordează elementele fundamentale ale statisticii parametrice. În primul paragraf al capitolului apar probleme în care se stabilește completitudinea unor familii de repartiții și suficiența unor statistici. Importanța acestor noțiuni este pusă în evidență în paragrafele 3 și 4, în care sînt prezentate probleme care cer găsirea unor estimări nedepășate, de dispersie minimă și a testelor uniforme, cele mai puternice pentru verificarea unor ipoteze parametrice. Problemele din paragraful 2 abordează prima întrebare cu care este confruntat un statistician: cum trebuie ales un model probabilist și cum se verifică adecvarea modelului la datele experimentale.

Problemele din capitolul IV au fost redactate de Monica Dumitrescu.

Lucrarea este adresată inginerilor, economiștilor, matematicienilor, studenților de la institutele tehnice și de la universități și tuturor celor interesați de teoria probabilităților și de statistică matematică și care aplică în activitatea lor metodele probabilistice și statistice.

AUTORII

CUPRINS

Capitolul I. Cîmp de probabilitate	7
§1. Cîmp de probabilitate. Proprietăți	11
§2. Probabilitatea clasică și probabilitatea geometrică	36
Capitolul II. Repartiții și variabile aleatoare	60
§1. Repartiția variabilelor aleatoare	69
§2. Momentele variabilelor aleatoare	98
§3. Funcția caracteristică și funcția generatoare	127
Capitolul III. Convergența variabilelor aleatoare	160
§1. Tipuri de convergență a variabilelor aleatoare	163
§2. Legea numerelor mari. Teorema limită centrală	188
Capitolul IV. Elemente de statistică matematică	212
§1. Completitudine. Suficiență	220
§2. Selecție. Concordanța repartiției de selecție cu o repartiție teoretică	233
§3. Estimarea parametrilor	238
§4. Verificarea ipotezelor statistice. Teste parametrice	261
§5. <i>Anexe</i>	298

CAPITOLUL I

CÎMP DE PROBABILITATE

Obiectul de bază în teoria probabilităților este *cîmpul de probabilitate* care se definește ca un triplet (E, \mathcal{K}, P) , unde:

– E este o mulțime numită mulțimea *evenimentelor elementare*;

– \mathcal{K} este un *corp borelian* pe E , adică \mathcal{K} este o familie de părți ale lui E așa încît $E \in \mathcal{K}$, $A \in \mathcal{K}$ implică $\complement A \in \mathcal{K}$ și $A_n \in \mathcal{K}$ implică $\bigcup_{n=1}^{\infty} A_n \in \mathcal{K}$ (deci perechea (E, \mathcal{K}) este un spațiu măsurabil). Elementele lui \mathcal{K} se numesc *evenimente*.

– P este o *probabilitate* pe \mathcal{K} , adică $P: \mathcal{K} \rightarrow [0, 1]$ și are proprietățile $P(E) = 1$ și $P\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} P(A_i)$, dacă I este cel mult numărabilă și $A_i \cap A_j = \emptyset$ dacă $i \neq j$ (dacă această egalitate este adevărată numai pentru I finită, atunci P se numește *probabilitate finit aditivă*).

Proprietăți:

a) (Subaditivitatea): $P\left(\bigcup_{i \in I} A_i\right) \leq \sum_{i \in I} P(A_i)$, dacă I este cel mult numărabilă;

b) Dacă $A \supset B$, atunci $P(A \setminus B) = P(A) - P(B)$. În particular $P(\complement A) = 1 - P(A)$;

c) *Formula lui Poincaré*:
$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{L \subset \{1, \dots, n\}} (-1)^{\text{card } L+1} P\left(\bigcap_{i \in L} A_i\right) =$$

$$= \sum_{i=1}^n P(A_i) - \sum_{i < j} P(A_i \cap A_j) + \sum_{i < j < k} P(A_i \cap A_j \cap A_k) - \dots + (-1)^{n+1} P(A_1 \cap \dots \cap A_n)$$
;

d) Dacă $A_n \nearrow A$ (resp. $A_n \searrow A$), atunci $P(A_n) \nearrow P(A)$ (resp. $P(A_n) \searrow P(A)$). Dacă P este probabilitate finită aditivă, atunci fiecare din afirmațiile din (d) asigură că P este probabilitate (chiar mai mult în cazul $A_n \searrow A$ este suficient de luat numai situația $A = \emptyset$).

Dacă \mathfrak{M} este o familie de părți ale lui E , atunci definim *corpul borelian generat* de \mathfrak{M} prin $\mathfrak{B}(\mathfrak{M}) = \bigcap_{\mathcal{K} \supset \mathfrak{M}} \mathcal{K}$; deci $\mathfrak{B}(\mathfrak{M})$ este cel mai mic corp borelian ce include \mathfrak{M} ,

\mathcal{K} corp borelian

\mathfrak{M} se numește *sistem de generatori* pentru $\mathfrak{B}(\mathfrak{M})$.

În particular dacă E este spațiu topologic, atunci \mathfrak{B}_E va desemna corpul borelian generat de mulțimile deschise (respectiv închise).

Dacă $(\mathcal{X}_i)_{i \in I}$ este o familie de corpuri boreliene (\mathcal{X}_i fiind corp borelian E_i), atunci definim *produsul corpurilor boreliene* \mathcal{X}_i prin $\bigotimes_{i \in I} \mathcal{X}_i = \mathfrak{B} \left(\prod_{i \in I} A_i; A_i \in \mathcal{X}_i \text{ și } \{i; A_i \neq E_i\} \text{ este finită} \right)$, deci în particular $\mathcal{X}_1 \otimes \dots \otimes \mathcal{X}_n = \mathfrak{B}(A_1 \times \dots \times A_n; A_i \in \mathcal{H}_i)$. Dată fiind o familie de spații măsurabile $(E_i, \mathcal{X}_i)_{i \in I}$ și o familie de aplicații $f_i: E \rightarrow E_i$, atunci notăm $\mathfrak{B}(f_i; i \in I) = \mathfrak{B} \left(\bigcup_{i \in I} f_i^{-1}(\mathcal{X}_i) \right)$. Cu alte cuvinte $\mathfrak{B}(f_i; i \in I)$ este cel mai mic corp borelian pe E în raport cu care fiecare f_i este măsurabilă. Se observă că $\bigotimes_{i \in I} \mathcal{X}_i$ este cel mai mic corp borelian pe $\prod_{i \in I} E_i$ în raport cu care proiecțiile sînt măsurabile.

e) (*Teorema de unicitate*). Dacă două probabilități P_1, P_2 definite pe $\mathfrak{B}(U)$ coincid pe \mathfrak{M} , unde \mathfrak{M} este o familie de mulțimi închisă la intersecția finită, atunci ele coincid pe $\mathfrak{B}(\mathfrak{M})$.

Două exemple de cîmpuri de probabilitate. 1) *Cîmpul clasic* care se definește prin tripletul $(E, \mathfrak{B}(E), P)$ unde E este o mulțime finită iar P este definită prin $P(F) = \frac{\text{card } F}{\text{card } E}$ (P este numită *probabilitatea clasică*). Acest cîmp de probabilitate corespunde unui experiment aleator ale cărui experimente elementare sînt în număr finit și sintegal probabile $\left(\text{avem } P(\{e\}) = \frac{1}{\text{card } E} \text{ pentru orice } e \in E \right)$.

Conform probabilității clasice avem că probabilitatea unui eveniment F este egală cu raportul dintre numărul evenimentelor elementare favorabile lui F și numărul total al evenimentelor elementare.

Subliniem faptul că primul pas în calculul probabilităților cu ajutorul probabilității clasice îl constituie definirea corectă a mulțimii evenimentelor elementare E și a evenimentului F a cărui probabilitate trebuie calculată.

Al doilea pas îl constituie calculul lui $\text{card } E$, $\text{card } F$. O metodă eficientă de numărare este dată de analiza combinatorie. Redăm în continuare patru reguli de numărare care utilizate cu ingeniozitate permit calculul unei mari varietăți de probabilități.

f₁) *Principiul de bază al analizei combinatorii*: dacă obiectele x_1, \dots, x_k pot fi alese în m_1, \dots, m_k moduri (în ordinea scrisă), atunci k -truplul ordonat (x_1, \dots, x_k) poate fi ales în m_1, \dots, m_k moduri.

f₂) *Formula pentru numărul de k -trupluri ordonate*. Fie U o mulțime cu n elemente și k un întreg. Atunci există n^k k -trupluri ordonate distincte (u_1, \dots, u_k) de elemente ale lui U . Dacă $k < n$, atunci există A_n^k k -trupluri ordonate distincte de elemente ale lui U ce au componente distincte.

f₃) *Formula pentru numărul de submulțimi ale unei mulțimi*. Dacă U este o mulțime cu n elemente, atunci există C_n^k submulțimi ale lui U cu k elemente.

f₄) *Formula pentru numărul de partiții*. Fie U o mulțime cu n elemente și fie r_1, \dots, r_k întregi ce satisfac: $r_i \geq 0, \sum_{i=1}^k r_i = n$. Atunci există $\frac{n!}{r_1! \dots r_k!}$ partiții distincte (U_1, \dots, U_k) ale lui U cu $\text{card } U_i = r_i$.

g) *Cîmpul geometric.* Fie \mathfrak{M} o mulțime de sisteme de h puncte din R^n ai cărei parametri (coordonate) determină mulțimea boreliană mărginită D_0 de măsură Lebesgue nenulă și fie $\mathfrak{M}_1 \subset \mathfrak{M}$ ai cărei parametri determină submulțimea boreliană $D_1 \subset D_0$.

Definim $P(\mathfrak{M}_1) = \frac{m(D_1)}{m(D_0)}$ unde m este măsura Lebesgue în R^{nh} . Fie acum \mathfrak{M} o familie de drepte în plan de ecuații $x \cos \theta + y \sin \theta - \rho = 0$ unde (ρ, θ) variază într-o mulțime boreliană mărginită D_0 de măsură Lebesgue nenulă din R^2 și fie $\mathfrak{M}_1 \subset \mathfrak{M}$ ai cărei parametri (ρ, θ) variază în mulțimea boreliană D_1 .

Se definește $P(\mathfrak{M}_1) = \frac{m(D_1)}{m(D_0)}$, m fiind măsura Lebesgue în R^2 (este vizibil că putem extinde considerațiile și la spațiul euclidian R^n).

Cîmpul de probabilitate $(D_0, \mathfrak{S}_{D_0}, P)$ se numește cîmpul geometric de probabilitate asociat lui \mathfrak{M} , iar P se numește probabilitatea geometrică asociată lui \mathfrak{M} . Cu cîmpurile de probabilitate se pot face operațiile h, i, j de mai jos.

h) *Cîmpul indus.* Fie (E, \mathfrak{X}, P) un cîmp de probabilitate și B un eveniment din \mathfrak{X} cu $P(B) > 0$. Atunci pentru orice eveniment A din \mathfrak{X} se definește probabilitatea condiționată a lui A de B prin $P_B(A) = P(A|B) = \frac{P(A \cap B)}{P(B)}$.

Definim corpul borelian $\mathfrak{X}_B = \{A \in \mathfrak{X}; A \subset B\}$. Tripletul (B, \mathfrak{X}_B, P_B) este un cîmp de probabilitate (îl vom numi *urma* pe B a lui (E, \mathfrak{X}, P) sau *cîmpul indus* pe B de (E, \mathfrak{X}, P)).

Să observăm că P_B este probabilitate și pe \mathfrak{X} (nu numai pe \mathfrak{X}_B). Probabilitatea condiționată se bucură de următoarele proprietăți mai importante:

h_1) Dacă $B_1, \dots, B_n \in \mathfrak{X}$ și $P(B_1 \cap \dots \cap B_{n-1}) > 0$, atunci

$$\blacksquare P(B_1 \cap \dots \cap B_n) = P(B_1)P(B_2 | B_1) \dots P(B_n | B_1 \cap \dots \cap B_{n-1}).$$

h_2) (*Formula probabilității totale*). Fie $(B_i)_{i \in I}$ o partiție cel mult numărabilă a lui E cu elemente din \mathfrak{X} așa încît fiecare B_i are probabilitate nenulă. Atunci pentru $A \in \mathfrak{X}$ avem

$$P(A) = \sum_{i \in I} P(B_i)P(A|B_i).$$

h_3) (*Formula lui Bayes*). Cu notațiile și ipotezele din h_2 și dacă în plus $P(A) > 0$, atunci

$$P(B_i | A) = \frac{P(B_i)P(A | B_i)}{\sum_{j \in I} P(B_j)P(A | B_j)}.$$

i) *Produsul direct.* Fie $(E_i, \mathfrak{X}_i, P_i)_{1 \leq i \leq n}$, n cîmpuri de probabilitate și fie $\bigotimes_{i=1}^n (E_i, \mathfrak{X}_i, P_i) = \left(\prod_{i=1}^n E_i, \bigotimes_{i=1}^n \mathfrak{X}_i, \bigotimes_{i=1}^n P_i \right)$ produsul lor considerate ca spații cu măsură).

Reamintim că $\bigotimes_{i=1}^n \mathfrak{X}_i = \mathfrak{S}(A_1 \times \dots \times A_n, A_i \in \mathfrak{X}_i)$ și $\bigotimes_{i=1}^n P_i$ este unica măsură

pe $\bigotimes_{i=1}^n \mathcal{X}_i$ pentru care $\left(\bigotimes_{i=1}^n P_i \right) \left(\prod_{i=1}^n A_i \right) = \prod_{i=1}^n P_i(A_i)$, $A_i \in \mathcal{X}_i$. Cum $\left(\bigotimes_{i=1}^n P_i \right) \left(\prod_{i=1}^n E_i \right) = \prod_{i=1}^n P_i(E_i) = 1$ rezultă că $\bigotimes_{i=1}^n P_i$ este o probabilitate și prin urmare $\bigotimes_{i=1}^n (E_i, \mathcal{X}_i, P_i)$ este cîmp de probabilitate (îl vom numi *produsul cîmpurilor* $(E_i, \mathcal{X}_i, P_i)$).

j) *Suma directă*. Fie E o mulțime și E_1, \dots, E_n o partiție a lui E . Pentru orice i fie $(E_i, \mathcal{X}_i, P_i)$ un cîmp de probabilitate și fie $\alpha_i > 0$ cu $\sum_{i=1}^n \alpha_i = 1$.

Definim $\bigoplus_{i=1}^n \mathcal{X}_i = \{A \subset E; A \cap E_i \in \mathcal{X}_i \text{ pentru orice } i\}$ și $\left(\bigoplus_{i=1}^n P_i \right) (A) = \sum_{i=1}^n \alpha_i P_i(A \cap E_i)$ pentru $A \in \bigoplus_{i=1}^n \mathcal{X}_i$. Tripletul $\left(E, \bigoplus_{i=1}^n \mathcal{X}_i, \bigoplus_{i=1}^n P_i \right)$ este cîmp de probabilitate și se numește *suma directă* a cîmpurilor $(E_i, \mathcal{X}_i, P_i)$ în raport cu ponderile α_i și se mai notează cu $\bigoplus_{i=1}^n (E_i, \mathcal{X}_i, P_i)$.

Corpurile boreliene $(\mathcal{X}_i)_{i \in I}$ sînt *independente* dacă

$$P \left(\bigcap_{i \in I} A_i \right) = \prod_{i \in I} P(A_i)$$

de îndată ce $A_i \in \mathcal{X}_i$ și $\{i, A_i \neq E\}$ este finită (avem fixat (E, \mathcal{X}, P) un cîmp de probabilitate și \mathcal{X}_i sînt corpuri boreliene incluse în \mathcal{X}). În particular rezultă că $\mathcal{X}_1, \dots, \mathcal{X}_n$ sînt independente dacă

$$P(A_1 \cap \dots \cap A_n) = P(A_1) \dots P(A_n)$$

pentru orice $A_i \in \mathcal{X}_i$.

Evenimentele $(A_i)_{i \in I}$ (respectiv partițiile lui E , $(\Delta_i)_{i \in I}$) sînt independente dacă corpurile boreliene $\{\mathfrak{B}(A_i)\}_{i \in I}$ (respectiv $\{\mathfrak{B}(\Delta_i)\}_{i \in I}$) sînt independente.

Următorul rezultat este foarte simplu dar extrem de util:

k) (*Lema Borel-Cantelli*). Fie (A_n) un șir de evenimente din \mathcal{X} .

1) Dacă $\sum_{n=1}^{\infty} P(A_n) < \infty$, atunci $P(\overline{\lim}_{n \rightarrow \infty} A_n) = 0$.

2) Dacă (A_n) sînt independente și $\sum_{n=1}^{\infty} P(A_n) = \infty$, atunci $P(\overline{\lim}_{n \rightarrow \infty} A_n) = 1$.

Are loc următorul criteriu de independență.

1) Corpurile boreliene $\mathfrak{B}(\mathfrak{N}_1), \dots, \mathfrak{B}(\mathfrak{N}_m)$, (unde fiecare \mathfrak{N}_i este închisă la intersecția finită și conține mulțimea E) sînt independente dacă și numai dacă $P(A_1 \cap \dots \cap A_m) = P(A_1) \dots P(A_m)$ pentru orice $A_i \in \mathfrak{N}_i$.

m) (*Asociativitatea independenței*). Dacă corpurile boreliene $(\mathcal{X}_i)_{i \in I}$ sînt independente și $(I_\alpha)_{\alpha \in A}$ este o partiție a lui I atunci corpurile boreliene $\left\{ \mathfrak{B} \left(\bigcup_{i \in I_\alpha} \mathcal{X}_i \right) \right\}_{\alpha \in A}$ sînt independente.

h) (*Dezasociativitatea independenței*). Fie $(I_\alpha)_{\alpha \in A}$ o partiție a mulțimii I și fie $(\mathcal{G}_\alpha)_{\alpha \in A}$ o familie de corpuri boreliene independente. Dacă pentru orice α , $(\mathcal{X}_i)_{i \in I_\alpha}$ este o familie de corpuri boreliene incluse în \mathcal{G}_α independente, atunci corpurile boreliene $(\mathcal{X}_i)_{i \in I}$ sînt independente.

§ 1. Cîmp de probabilitate. Proprietăți

1.1. Fie (E, \mathcal{X}, P) un cîmp de probabilitate și $A, B, C \in \mathcal{X}$. Să se demonstreze relațiile:

- a) $P(A \Delta B) = P(A) + P(B) - 2P(A \cap B)$;
- b) $\max [P(A), P(B)] \leq P(A \cup B) \leq 2 \max [P(A), P(B)]$;
- c) $|P(A \cap B) - P(A \cap C)| \leq P(B \Delta C)$;
- d) $P(A \cup B) P(A \cap B) \leq P(A)P(B)$;
- e) $P^2(A \cup B) + P^2(A \cap B) = P^2(A) + P^2(B) + 2P(A \cap C B) P(C A \cap B)$.

Soluție. a) Avem că $(A \Delta B) \cup (A \cap B) = A \cup B$ și $(A \Delta B) \cap (A \cap B) = \emptyset$, deci $P(A \Delta B) + P(A \cap B) = P(A \cup B) = P(A) + P(B) - P(A \cap B)$, de unde $P(A \Delta B) = P(A) + P(B) - 2P(A \cap B)$

b) $A \subset A \cup B$, $B \subset A \cup B$ deci $P(A) \leq P(A \cup B)$, $P(B) \leq P(A \cup B)$ și prin urmare

$$\max [P(A), P(B)] \leq P(A \cup B).$$

Apoi

$$P(A \cup B) \leq P(A) + P(B) \leq 2 \max [P(A), P(B)].$$

c) Observăm că $A \cap B \subset (A \cap C) \cup (B \Delta C)$, $A \cap C \subset (A \cap B) \cup (B \Delta C)$, așa că

$$P(A \cap B) \leq P(A \cap C) + P(B \Delta C); \quad P(A \cap C) \leq P(A \cap B) + P(B \Delta C)$$

de unde rezultă inegalitatea din enunț.

d) Deoarece $P(A) - P(A \cap B) \geq 0$, $P(B) - P(A \cap B) \geq 0$ și deci $[P(A) - P(A \cap B)][P(B) - P(A \cap B)] \geq 0$, avem că $[P(A) + P(B) - P(A \cap B)] P(A \cap B) \leq P(A) P(B)$ sau $P(A \cup B) P(A \cap B) \leq P(A) P(B)$.

e) Rezultă prin calcul direct observînd că

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

$$P(A \cap \complement B) = P(A) - P(A \cap B);$$

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

1.2.. Fie A, B două evenimente. Să se arate că

a) $|P(A \cap B) - P(A)P(B)| \leq \frac{1}{4}$.

b) $P^2(A \cap B) + P^2(\complement A \cap B) + P^2(A \cap \complement B) + P^2(\complement A \cap \complement B) \geq 1/4$, egalitatea avînd loc dacă și numai dacă $P(A) = P(B) = 1/2$ și $P(A \cap B) = 1/4$.

Soluție. a) Presupunem că $P(A) \geq P(B)$.

Deoarece $P(A \cap B) \leq P(B)$, avem

$$P(A \cap B) - P(A)P(B) \leq P(B)(1 - P(A)) \leq P(A)(1 - P(A)) \leq 1/4.$$

Să punem $x = P(A \cap B)$, $a = P(A \cap \complement B)$ și $b = P(\complement A \cap B)$ observăm că $a + b + x = P(A \cup B) \leq 1$

și cum $a \leq P(A)$; $b \leq P(\complement A) = 1 - P(A)$ $ab \leq P(A)(1 - P(A)) \leq 1/4$.

Avem în fine

$$\begin{aligned} P(A \cap B) - P(A)P(B) &= x - (a + x)(b + x) = \\ &= x - [ab + x(a + b + x)] \geq x - ab - x = -ab \geq -1/4. \end{aligned}$$

b) Evident

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

$$\begin{aligned} \text{Să punem } P(A \cap B) &= x + \frac{1}{4}, \quad P(A \cap \complement B) = y + \frac{1}{4}, \quad P(\complement A \cap B) = \\ &= z + \frac{1}{4}, \quad P(\complement A \cap \complement B) = t + \frac{1}{4}. \end{aligned}$$

Deci $x + y + z + t = 0$. Dar

$$\begin{aligned} \left(x + \frac{1}{4}\right)^2 + \left(y + \frac{1}{4}\right)^2 + \left(z + \frac{1}{4}\right)^2 + \left(t + \frac{1}{4}\right)^2 &= \frac{x^2 + y^2 + z^2 + t^2 +}{\geq 0} \\ &+ \underbrace{\frac{x + y + z + t}{2}}_{= 0} + \frac{1}{4} \geq \frac{1}{4}. \end{aligned}$$

egalitatea avînd loc dacă și numai dacă $x = y = z = t = 0$, ceea ce revine la faptul că $P(A) = P(B) = \frac{1}{2}$ și $P(A \cap B) = \frac{1}{4}$.

1.3. Într-un câmp de probabilitate (E, \mathcal{X}, P) fie evenimentele A_1, A_2, \dots, A_n astfel încît $\sum_{i=1}^n P(A_i) > n - 1$. Să se arate că $P\left(\bigcap_{i=1}^n A_i\right) > 0$.

Soluție. Avem că

$$\bigcap_{i=1}^n A_i = E \setminus \left(\bigcup_{i=1}^n \complement A_i\right).$$

Deci $P\left(\bigcap_{i=1}^n A_i\right) = 1 - P\left(\bigcup_{i=1}^n \complement A_i\right)$.

$$\begin{aligned} \text{Dar } P\left(\bigcup_{i=1}^n \complement A_i\right) &\leq \sum_{i=1}^n P(\complement A_i) = \sum_{i=1}^n (1 - P(A_i)) = n - \sum_{i=1}^n P(A_i) < \\ &< n - (n - 1) = 1, \end{aligned}$$

de unde rezultă $P\left(\bigcap_{i=1}^n A_i\right) > 0$.

1.4. Într-un câmp de probabilitate (E, \mathcal{X}, P) fie evenimentele A, A_1, A_2, \dots, A_n astfel încît $\bigcap_{i=1}^n A_i \subset A$. Să se arate că

$$P(A) \geq \sum_{i=1}^n P(A_i) - (n - 1).$$

Soluție. Să observăm întii că pentru orice evenimente B_1, \dots, B_n avem

$$\begin{aligned} P\left(\bigcup_{i=1}^n B_i\right) &= \sum_i P(B_i) - \sum_{i < j} P(B_i \cap B_j) + \sum_{i < j < k} P(B_i \cap B_j \cap B_k) + \dots \\ &\dots + (-1)^{n-1} P\left(\bigcap_{i=1}^n B_i\right) \leq \sum_{i=1}^n P(B_i) \end{aligned}$$

deci

$$\sum_{i < j} P(B_i \cap B_j) \geq \sum_{i < j < k} P(B_i \cap B_j \cap B_k) + \dots + (-1)^{n-1} P\left(\bigcap_{i=1}^n B_i\right)$$

Apoi prin ipoteză avem

$$\begin{aligned} P(A) &\geq P\left(\bigcap_{i=1}^n A_i\right) = 1 - P\left(\bigcup_{i=1}^n \bar{A}_i\right) = 1 - \sum_i P(\bar{A}_i) + \\ &+ \left(\sum_{i < j} P(\bar{A}_i \cap \bar{A}_j) - \sum_{i < j < k} P(\bar{A}_i \cap \bar{A}_j \cap \bar{A}_k) + \dots - \right. \\ &\left. - (-1)^{n-1} P\left(\bigcap_{i=1}^n \bar{A}_i\right)\right) \geq 1 - \sum_{i=1}^n P(\bar{A}_i) = \sum_{i=1}^n P(A_i) - (n-1). \end{aligned}$$

1.5. Să se arate că oricare ar fi evenimentele A_1, \dots, A_n are loc egalitatea

$$P(A_1 \cap \dots \cap A_n) = \sum_{L \subset \{1, \dots, n\}} (-1)^{\text{card } L+1} P\left(\bigcup_{i \in L} A_i\right).$$

Soluție. Aplicînd proprietățile *b* și *c* din rezumatul teoretic, obținem:

$$\begin{aligned} P\left(\bigcap_{i=1}^n A_i\right) &= P\left(\bar{C}\left(\bigcup_{i=1}^n (C A_i)\right)\right) = 1 - P\left(\bigcup_{i=1}^n C A_i\right) = 1 - \\ &- \sum_{L \subset \{1, \dots, n\}} (-1)^{\text{card } L+1} P\left(\bigcap_{i \in L} C A_i\right) = 1 - \\ &- \sum_{L \subset \{1, \dots, n\}} (-1)^{\text{card } L+1} P\left(\bar{C}\left(\bigcup_{i \in L} A_i\right)\right) = 1 - \\ &- \sum_{L \subset \{1, \dots, n\}} (-1)^{\text{card } L+1} \left[1 - P\left(\bigcup_{i \in L} A_i\right)\right] = \sum_{L \subset \{1, \dots, n\}} (-1)^{\text{card } L+1} \cdot \\ &\cdot P\left(\bigcup_{i \in L} A_i\right) + 1 - \sum_{i=1}^n \sum_{\substack{L \subset \{1, \dots, n\} \\ \text{card } L=i}} (-1)^{i+1} = \sum_{L \subset \{1, \dots, n\}} (-1)^{\text{card } L+1} \cdot \\ &\cdot P\left(\bigcup_{i \in L} A_i\right) + 1 - \sum_{i=1}^n (-1)^{i+1} C_n^i = \sum_{L \subset \{1, \dots, n\}} (-1)^{\text{card } L+1} P\left(\bigcup_{i \in L} A_i\right) + \\ &+ (1-1)^n = \sum_{L \subset \{1, \dots, n\}} (-1)^{\text{card } L+1} P\left(\bigcup_{i \in L} A_i\right). \end{aligned}$$

Lăsăm cititorului să găsească o demonstrație prin inducție după n .

1.6. Date fiind evenimentele A_1, \dots, A_n , să se calculeze probabilitatea realizării a r dintre aceste evenimente cunoscînd probabilitățile $P\left(\bigcap_{i \in L} A_i\right)$ pentru orice $L \subset \{1, \dots, n\}$.

Soluție. Evenimentul ce ne interesează se scrie sub forma

$$B = \bigcup_{\substack{J \subset \{1, \dots, n\} \\ \text{card } J = r}} \left\{ \left(\bigcap_{i \in J} A_i \right) \cap \left(\bigcap_{j \notin J} \bar{A}_j \right) \right\}$$

deci putem scrie

$$\begin{aligned} P(B) &= \sum_{\substack{J \subset \{1, \dots, n\} \\ \text{card } J = r}} P\left(\bigcap_{i \in J} A_i \cap \bigcap_{j \notin J} \bar{A}_j\right) = \sum_{\substack{J \subset \{1, \dots, n\} \\ \text{card } J = r}} \left[P\left(\bigcap_{i \in J} A_i\right) - \right. \\ &\quad \left. - P\left(\bigcup_{i \notin J} (A_j \cap \bigcap_{i \in J} A_i)\right) \right] = \sum_{\substack{J \subset \{1, \dots, n\} \\ \text{card } J = r}} \left[P\left(\bigcap_{i \in J} A_i\right) - \right. \\ &\quad \left. - \sum_{\substack{L \supset J \\ \text{card } L \geq r+1}} (-1)^{\text{card}(L-J)} P\left(\bigcap_{j \in L} A_j\right) \right] = \sum_{\substack{J \subset \{1, \dots, n\} \\ \text{card } J = r}} \sum_{L \supset J} (-1)^{\text{card } L - r} P\left(\bigcup_{i \in L} A_i\right) = \\ &= \sum_{\substack{L \subset \{1, \dots, n\} \\ \text{card } L \geq r}} \sum_{\substack{J \subset L \\ \text{card } J = r}} (-1)^{\text{card } L - r} P\left(\bigcap_{i \in L} A_i\right) = \\ &= \sum_{\substack{L \subset \{1, \dots, n\} \\ \text{card } L \geq r}} (-1)^{\text{card } L - r} C'_{\text{card } L} P\left(\bigcap_{i \in L} A_i\right) = \\ &= \sum_{k=r}^n \sum_{\substack{L \subset \{1, \dots, n\} \\ \text{card } L = k}} (-1)^{k-r} C'_k P\left(\bigcap_{i \in L} A_i\right). \end{aligned}$$

1.7. Fie p_1, p_2, p_{12} numere reale. Să se arate că există două evenimente A, B într-un câmp de probabilitate așa încît $P(A) = p_1, P(B) = p_2, P(A \cap B) = p_{12}$ dacă și numai dacă sînt verificate relațiile:

$$\begin{cases} 1 - p_1 - p_2 + p_{12} \geq 0 & p_1 - p_{12} \geq 0 \\ p_2 - p_{12} \geq 0 & p_{12} \geq 0 \end{cases} \quad (1)$$

Soluție. Presupunem că există A, B cu proprietățile din enunț. Atunci

$$\begin{aligned} 0 &\leq P(\bar{A} \cup \bar{B}) = 1 - P(A \cap B) = 1 - p_{12} \\ &\quad + P(A \cap B) = 1 - p_1 - p_2 + p_{12} \\ 0 &\leq P(A \setminus B) = P(A) - P(A \cap B) = p_1 - p_{12} \\ 0 &\leq P(B \setminus A) = P(B) - P(A \cap B) = p_2 - p_{12} \\ 0 &\leq P(A \cap B) = p_{12}. \end{aligned}$$

Presupunem acum că relațiile (1) sînt verificate. Rezultă în particular că $0 \leq p_{12} \leq p_1 \leq p_1 + p_2 - p_{12} \leq 1$; $0 \leq p_{12} \leq p_2 \leq p_1 + p_2 - p_{12} \leq 1$, deci p_1, p_2, p_{12} sînt cuprinse între 0 și 1.

Pe cîmpul de probabilitate $([0, 1], \mathfrak{B}_{[0, 1]}, m)$, m măsura Lebesgue, evenimentele $A = [0, p_1]$, $B = [p_1 - p_{12}, p_1 + p_2 - p_{12}]$ satisfac cerințele.

1.8. Dacă (A_n) este un șir de evenimente, să se probeze inegalitățile

$$P\left(\lim_n A_n\right) \leq \lim_n P(A_n) \leq \overline{\lim}_n P(A_n) \leq P\left(\overline{\lim}_n A_n\right)$$

Soluție. Avem:

$$P\left(\lim_n A_n\right) = P\left(\bigcap_n \bigcap_{m \geq n} A_m\right) = \lim_{n \rightarrow \infty} P\left(\bigcap_{m \geq n} A_m\right) \leq \lim_{n \rightarrow \infty} P(A_n)$$

$$P\left(\overline{\lim}_n A_n\right) = P\left(\bigcap_n \bigcup_{m \geq n} A_m\right) = \lim_{n \rightarrow \infty} P\left(\bigcup_{m \geq n} A_m\right) \geq \overline{\lim}_{n \rightarrow \infty} P(A_n)$$

am folosit faptul că șirul $\bigcap_{m \geq n} A_m$ este crescător și șirul $\bigcup_{m \geq n} A_m$ este descrescător și proprietatea d din rezumatul teoretic).

1.9. Pentru un șir de evenimente (A_n) să se demonstreze implicațiile

$$\sum P(A_n \Delta A_{n+1}) < \infty \Rightarrow P\left(\overline{\lim}_n A_n\right) = P\left(\lim_n A_n\right) \Rightarrow \lim_{m, n \rightarrow \infty} P(A_m \Delta A_n) = 0$$

Soluție. Avem:

$$\begin{aligned} \overline{\lim}_n A_n \setminus \lim_n A_n &= \bigcap_n \bigcup_{k \geq n} A_k \setminus \bigcup_k \bigcap_{k \geq n} A_k = \bigcap_k \bigcup_{m, n \geq k} A_m \Delta A_n = \\ &= \bigcap_k \bigcup_{n \geq k} A_n \Delta A_{n+1} \end{aligned}$$

deci

$$\begin{aligned} P\left(\overline{\lim}_n A_n \setminus \lim_n A_n\right) &= P\left(\bigcap_k \bigcup_{n \geq k} A_n \Delta A_{n+1}\right) = \lim_{k \rightarrow \infty} P\left(\bigcup_{n \geq k} A_n \Delta A_{n+1}\right) \leq \\ &\leq \lim_{k \rightarrow \infty} \sum_{n=k}^{\infty} P(A_n \Delta A_{n+1}) = 0 \text{ dacă } \sum_n P(A_n \Delta A_{n+1}) < \infty, \text{ cu alte cuvinte} \\ P\left(\overline{\lim}_n A_n\right) &= P\left(\lim_n A_n\right). \end{aligned}$$

Dacă $P(\overline{\lim} A_n) = P(\underline{\lim} A_n)$, atunci rezultă că $P\left(\bigcap_k \bigcup_{m, n \geq k} A_m \Delta A_n\right) = 0$ sau echivalent $\lim_{m, n \rightarrow \infty} P\left(\bigcup_{m, n \geq k} A_m \Delta A_n\right) = 0$. Apoi din incluziunea

$$A_m \Delta A_n \subset \bigcup_{k, l \geq \min(m, n)} A_k \Delta A_l$$

rezultă că $\lim_{m, n \rightarrow \infty} P(A_m \Delta A_n) \leq \lim_{m, n \rightarrow \infty} P\left(\bigcup_{k, l \geq \min(m, n)} A_k \Delta A_l\right) =$
 $= \lim_{n \rightarrow \infty} P\left(\bigcup_{k, l \geq n} A_k \Delta A_l\right) = 0$.

1.10. Să se arate că într-un câmp de probabilitate (E, \mathcal{A}, P) orice familie de evenimente disjuncte două câte două și de probabilități pozitive este cel mult numărabilă.

Soluție. Fie I o astfel de familie și fie $I_n = \left\{A \in I; P(A) > \frac{1}{n}\right\}$. Avem evident că $I = \bigcup_n I_n$, așa că este suficient să arătăm că I_n este cel mult numărabilă.

Vom arăta mai mult și anume că I_n are cel mult $n - 1$ evenimente. Prin absurd să presupunem că I_n are cel puțin $m \geq n$ evenimente, fie ele A_1, \dots, A_m . Atunci

$$1 = P(E) \geq P\left(\bigcup_{i=1}^m A_i\right) = \sum_{i=1}^m P(A_i) > \frac{m}{n} \geq 1,$$

deci o contradicție.

1.11. Fie P o probabilitate finit aditivă definită pe corpul borelian $\mathfrak{B}(\mathfrak{M})$. Dacă are loc egalitatea $P(A) = \sup_{\substack{B \subset A \\ B \in \mathfrak{M}}} P(B)$ și \mathfrak{M} este o familie compactă, în sensul că dacă $A_n \in \mathfrak{M}$ și $\bigcap_n A_n = \emptyset$, atunci există m așa încît $\bigcap_{n=1}^m A_n = \emptyset$, să se arate că P este probabilitate.

Soluție. Fie $A_n \searrow \emptyset$ și să arătăm că $P(A_n) \searrow 0$, deci că pentru $\varepsilon > 0$ există m așa încît $P(A_m) \leq \varepsilon$. Alegem $B_n \subset A_n$, $B_n \in \mathfrak{M}$ așa încît $P(A_n) - P(B_n) \leq \frac{\varepsilon}{2^n}$.

Avem $\bigcap_n B_n = \emptyset$, deci există m așa încît $\bigcap_{n=1}^m B_n = \emptyset$. Deoarece

$$A_m = \bigcap_{n=1}^m A_n \subset \bigcup_{n=1}^m (A_n \setminus B_n) \quad \text{obținem}$$

$$\begin{aligned} P(A_m) &\leq P\left(\bigcup_{n=1}^m A_n \setminus B_n\right) \leq \sum_{n=1}^m P(A_n \setminus B_n) = \sum_{n=1}^m [P(A_n) - P(B_n)] \leq \\ &\leq \sum_{n=1}^m \frac{\varepsilon}{2^n} \leq \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon. \end{aligned}$$

1.12. Să se arate că nu orice probabilitate finit aditivă definită pe un corp este σ -aditivă (reamintim că un corp este o familie de mulțimi nevidă, închisă la complementară și reuniunea finită).

Soluție. Fie Ω mulțimea numerelor raționale din $[0, 1]$ și \mathcal{C} corpul pe Ω format din mulțimile de forma $\{r; a \leq r \leq b\}$, $\{r; a \leq r < b\}$, $\{r; a < r \leq b\}$, $\{r; a < r < b\}$, a, b numere raționale și din reuniunile finite de astfel de mulțimi.

Pentru orice mulțime A de forma $\{r; a \leq r \leq b\}$, $\{r; a \leq r < b\}$, $\{r; a < r \leq b\}$, $\{r; a < r < b\}$ definim $P(A) = b - a$, iar dacă $B \in \mathcal{C}$ și $B = \bigcup_{i=1}^m A_i$, unde A_i sînt disjuncte două cîte două, definim

$$P(B) = \sum_{i=1}^m P(A_i).$$

Atunci P este bine definită și este finit aditivă. Arătăm că P nu este σ -aditivă.

Observăm că dacă r este un rațional din $[0, 1]$, atunci $\{r\} \in \mathcal{C}$ și $P(\{r\}) = 0$. Afirmarea rezultă din relațiile:

$$\Omega = \bigcup_{r \in \Omega} \{r\}; \quad P(\Omega) = 1 \neq \sum_{r \in \Omega} P(\{r\}) = 0.$$

1.13. După cum este cunoscut, teorema de unicitate a probabilităților afirmă că dacă două probabilități coincid pe un sistem de generatori închis la intersecția finită, atunci ele coincid pe corpul borelian generat de un astfel de sistem de generatori. Să se arate că nu se poate renunța la cerința ca sistemul de generatori să fie închis la intersecția finită.

Soluție. Fie $\Omega = \{l_1, l_2, l_3, l_4\}$, $\mathfrak{M} = \{\{l_1, l_2\}, \{l_1, l_3\}\}$. Este clar că $\mathfrak{B}(\mathfrak{M}) = \mathfrak{A}(\Omega)$ și \mathfrak{M} nu este închis la intersecția finită. Alegem $p, q > 0$ așa încît $p + q \leq 1$ și fie $\varepsilon = \min(p, q)$. Definim probabilitățile P_1, P_2 pe $\mathfrak{A}(\Omega)$ prin egalitățile:

$$P_1(\{l_1\}) = 0, P_1(\{l_2\}) = p, P_1(\{l_3\}) = q, P_1(\{l_4\}) = 1 - p - q;$$

$$P_2(\{l_1\}) = \varepsilon, P_2(\{l_2\}) = p - \varepsilon, P_2(\{l_3\}) = q - \varepsilon,$$

$$P_2(\{l_4\}) = 1 - p - q + \varepsilon.$$

Avem

$$P_1(\{l_1, l_2\}) = P_2(\{l_1, l_2\}), P_1(\{l_1, l_3\}) = P_2(\{l_1, l_3\}),$$

dar $P_1 \neq P_2$ căci $P_1(\{l_1\}) \neq P_2(\{l_1\})$.

1.14. Să se arate că două probabilități P_1, P_2 pe $\mathfrak{B}(\Delta)$ (unde Δ este o partiție cel mult numărabilă al unei mulțimi E) sînt egale dacă și numai dacă coincid pe Δ .

Soluție. Deci avem de arătat că dacă $P_1 = P_2$ pe Δ , atunci $P_1 = P_2$ pe $\mathfrak{B}(\Delta)$. Familia $\Delta' = \Delta \cup \{\emptyset\}$ este închisă la intersecția finită, satisface egalitatea $\mathfrak{B}(\Delta') = \mathfrak{B}(\Delta)$ și $P_1 = P_2$ pe Δ' . Din teorema de unicitate rezultă că $P_1 = P_2$ pe $\mathfrak{B}(\Delta')$.

Altă soluție. Arătăm întii că $\mathfrak{B}(\Delta) = \left\{ \bigcup_{A \in \Delta'} A; \Delta' \subset \Delta \right\}$. Este clar că totul se reduce la a arăta că $\mathfrak{F} = \left\{ \bigcup_{A \in \Delta'} A; \Delta' \subset \Delta \right\}$ este corp borelian. Avem că $E = \bigcup_{A \in \Delta} A \in \mathfrak{F}$ și dacă $B = \bigcup_{A \in \Delta'} A$ atunci $\mathfrak{C}B = \bigcup_{A \in \Delta \setminus \Delta'} A \in \mathfrak{F}$. Apoi dacă $B_n = \bigcup_{A \in \Delta_n} A$ atunci $\bigcup_n B_n = \bigcup_{A \in \bigcup_n \Delta_n} A \in \mathfrak{F}$. Dacă $B = \bigcup_{A \in \Delta'} A$ atunci $P_1(B) = \sum_{A \in \Delta'} P_1(A) = \sum_{A \in \Delta'} P_2(A) = P_2(B)$.

1.15. Fie E o mulțime cel mult numărabilă și \mathfrak{X} un corp borelian pe E . Să se arate că orice probabilitate P pe \mathfrak{X} este unic determinată de valorile ei într-o mulțime cel mult numărabilă de evenimente din \mathfrak{X} .

Soluție. Avînd în vedere problema precedentă observăm că este suficient să dovedim că avem $\mathfrak{X} = \mathfrak{B}(\Delta)$ cu Δ partiție (evident cel mult numărabilă) a lui E .

Fie pentru $x \in E$, $E_x = \bigcap_{\substack{x \in A \\ A \in \mathcal{X}}} A$ (avem $x \in E \in \mathcal{X}$). Vom demonstra că $E_x \in \mathcal{X}$ arătând că $\mathcal{C}E_x \in \mathcal{X}$.

Avem $\mathcal{C}E_x = \bigcup_{x \notin A} A$ și pentru orice $y \in \mathcal{C}E_x$ există $A_y \in \mathcal{X}$ așa încît $y \in A_y$ și $x \notin A_y$, deci $\mathcal{C}E_x = \bigcup_{y \in \mathcal{C}E_x} A_y \in \mathcal{X}$ ($\mathcal{C}E_x$ este cel mult numărabilă). Definim pe E_x relația de echivalență „ \sim ” prin $x \sim y \Leftrightarrow E_x = E_y$.

Arătăm că $\hat{x} = E_x$ pentru orice x . Fie $y \in \hat{x}$; atunci $E_y = E_x$ și cum $y \in E_y$, rezultă că $y \in E_x$. Reciproc, fie $y \in E_x$. Trebuie să arătăm că $y \in \hat{x}$, deci că $E_y = E_x$.

Evident $E_y \subset E_x$, căci $E_x \in \mathcal{X}$. Incluziunea reciprocă rezultă dacă arătăm că $x \in E_y$. Prin absurd presupunem că $x \in \mathcal{C}E_y$. Rezultă că există $A_x \in \mathcal{X}$ așa încît $x \in A_x$ și $y \notin A_x$, deci $y \notin E_x$, ceea ce nu se poate.

Fie A partiția cel mult numărabilă a lui E definită prin $\Delta = \{E_x; x \in E\}$. Vom arăta că are loc egalitatea $\mathcal{X} = \mathfrak{B}(A)$. Incluziunea $\mathfrak{B}(\Delta) \subset \mathcal{X}$ este imediată. Pentru incluziunea reciprocă se arată că orice $A \in \mathcal{X}$ se scrie sub forma $A = \bigcup_{x \in A} E_x \in \mathfrak{B}(\Delta)$.

1.16. Fie E o mulțime și Δ o partiție cel mult numărabilă a lui E . Să se arate că a da o probabilitate pe $\mathfrak{B}(\Delta)$ este echivalent cu a da un sistem de numere $(p_A)_{A \in \Delta}$, unde $p_A \geq 0$ și $\sum_{A \in \Delta} p_A = 1$.

Observație. În particular rezultă că dacă E este cel mult numărabilă, atunci a da o probabilitate pe $\mathfrak{B}(E)$ este echivalent cu a da un sistem de numere $(p_i)_{i \in E}$ unde $p_i \geq 0$ și $\sum_{i \in E} p_i = 1$ (se ia $\Delta = (\{i\})_{i \in E}$ și se observă că $\mathfrak{B}(\Delta) = \mathfrak{B}(E)$).

Soluție. Din problema 1.14 (vezi soluția) rezultă că $\mathfrak{B}(\Delta) = \left\{ \bigcup_{A \in \Delta'} A; \Delta' \subset \Delta \right\}$. Corespondența se obține astfel:

Dacă P este o probabilitate pe $\mathfrak{B}(\Delta)$, atunci definim $p_A = P(A)$. Este vizibil că $p_A \geq 0$ și $\sum_{A \in \Delta} p_A = P\left(\bigcup_{A \in \Delta} A\right) = P(E) = 1$.

Reciproc, fie (p_A) un sistem de numere cu proprietățile din enunț. Atunci pentru $B \in \mathfrak{B}(\Delta)$, deci $B = \bigcup_{A \in \Delta'} A$, definim $P(B) = \sum_{A \in \Delta'} p_A$. Avem $P(E) = \sum_{A \in \Delta} p_A = 1$ și dacă $B_n = \bigcup_{A \in \Delta_n} A$ sînt disjuncte două cîte două, atunci familiile Δ_n sînt disjuncte două cîte două și

$$P\left(\bigcup_n B_n\right) = P\left(\bigcup_n \bigcup_{A \in \Delta^n} A\right) = P\left(\bigcup_n \bigcup_{A \in \Delta^n} A\right) = \sum_n \sum_{A \in \Delta^n} p_A = \sum_n \sum_{A \in \Delta^n} p_A =$$

$$= \sum_n P(B_n),$$
 deci P este probabilitate pe $\mathfrak{B}(\Delta)$. Corespondența este evident bijectivă.

1.17. Fie $A_i, i = 1, \dots, n$ evenimente independente într-un câmp (E, \mathfrak{K}, P) de probabilitate cu $P(A_i) = p_i$ și A evenimentul ca să se producă cel puțin unul din evenimentele $A_i, P(A) = p$.

Să se arate că:

$$\sum_{i=1}^n p_i > p > 1 - e^{-\sum_{i=1}^n p_i}.$$

Soluție. Deoarece $A = \bigcup_{i=1}^n \left[A_i \cap \left(\bigcap_{j=1}^n \complement A_j \right) \right]$, avem pe de o parte

$$p = P(A) = \sum_{i=1}^n P\left[A_i \cap \left(\bigcap_{j=1}^n \complement A_j \right) \right] \leq \sum_{i=1}^n P(A_i) = \sum_{i=1}^n p_i.$$

Pe de altă parte, deoarece $\complement A = \bigcap_{i=1}^n \complement A_i$ și deoarece evenimentele $A_i, i = 1, 2, \dots, n$ sînt independente

$$p = 1 - P(\complement A) = 1 - (1 - p_1)(1 - p_2) \dots (1 - p_n) > 1 -$$

$$- \left(1 - \frac{\sum_{i=1}^n p_i}{n} \right)^n > 1 - e^{-\sum_{i=1}^n p_i}$$

(am ținut seama și de faptul că $\sum_{i=1}^n p_i \leq n$).

1.18. Pe un câmp de probabilitate (E, \mathfrak{K}, P) fie A_1, \dots, A_n, \dots o partiție a lui E cu elemente din \mathfrak{K} și fie A un eveniment oarecare.

Dacă A este independent de fiecare din evenimentele A_n , să se arate că A este independent de corpul borelian generat de partiția precedentă (în sensul că A este independent de orice eveniment din acest corp borelian).

Soluție. Din problema 1.14 rezultă că are loc egalitatea

$$\mathfrak{B}(A_n; n \geq 1) = \left\{ \bigcup_{i \in I} A_i; I \subset \{1, \dots, n, \dots\} \right\}$$

Avem de arătat că

$$P\left(A \cap \left(\bigcup_{i \in I} A_i\right)\right) = P(A)P\left(\bigcup_{i \in I} A_i\right).$$

Putem scrie

$$\begin{aligned} P\left(A \cap \left(\bigcup_{i \in I} A_i\right)\right) &= P\left(\bigcup_{i \in I} (A \cap A_i)\right) = \sum_{i \in I} P(A \cap A_i) = \\ &= \sum_{i \in I} P(A) P(A_i) = P(A) P\left(\bigcup_{i \in I} A_i\right). \end{aligned}$$

1.19. a) Să se arate că evenimentele A_1, \dots, A_n sînt independente dacă și numai dacă

$$P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j) \quad (1)$$

pentru orice $J \subset \{1, \dots, n\}$.

b) Să se arate că partițiile cel mult numărabile $\Delta_1, \dots, \Delta_n$ sînt independente dacă și numai dacă

$$P\left(\bigcap_{i=1}^n A_i\right) = \prod_{i=1}^n P(A_i) \quad (2)$$

pentru orice $A_i \in \Delta_i$.

Soluție. a) Familia $\mathfrak{M} = \{A_j, \emptyset\}$ este închisă la intersecția finită și $\mathfrak{B}(\mathfrak{M}_j) = \mathfrak{B}(A_j)$

De asemenea este vizibil că (1) este valabilă și dacă în loc de un A_j punem \emptyset . Se aplică acum teorema de unicitate a probabilităților.

Altă soluție. Avem de arătat:

$$P\left(\bigcap_{j \in J} B_j\right) = \prod_{j \in J} P(B_j)$$

dacă $B_j \in \mathfrak{B}(A_j) = \{(A_j, \mathcal{C}A_j, \emptyset, E)\}$. Se observă că totul se reduce la cazul: dacă $J = \{j_1, \dots, j_l\}$, $j_1 < \dots < j_l$, atunci $P(\mathcal{C}A_{j_1} \cap \dots \cap \mathcal{C}A_{j_k} \cap A_{j_{k+1}} \cap \dots \cap A_{j_l}) = P(\mathcal{C}A_{j_1}) \dots P(\mathcal{C}A_{j_k}) P(A_{j_{k+1}}) \dots P(A_{j_l})$.

Aplicind inducția după k , rămîne de considerat cazul $k = 1$.

Avem :

$$\begin{aligned} P(\complement A_{j_1} \cap A_{j_2} \cap \dots \cap A_{j_l}) &= P(A_{j_2} \cap \dots \cap A_{j_l}) - \\ &- P(A_{j_1} \cap A_{j_2} \cap \dots \cap A_{j_l}) = P(A_{j_2}) \dots P(A_{j_l}) - \\ &- P(A_{j_1}) P(A_{j_2}) \dots P(A_{j_l}) = [1 - P(A_{j_1})] P(A_{j_2}) \dots P(A_{j_l}) = \\ &= P(\complement A_{j_1}) P(A_{j_2}) \dots P(A_{j_l}) \end{aligned}$$

b) Se consideră $\mathfrak{A}_j = \{\Delta_j, \emptyset\}$ și se aplică teorema de unicitate a probabilităților.

Altă soluție. Din problema 1.14 se știe că $\mathfrak{B}(\Delta_j) = \left\{ \bigcup_{A_j \in \Delta} A_j ; \Delta \subset \Delta_j \right\}$, deci avem de arătat că $P(B_1 \cap \dots \cap B_n) = P(B_1) \dots P(B_n)$ pentru $B_j \in \mathfrak{B}(\Delta_j)$. Scriind $B_j = \bigcup_{A_j \in \Delta_j} A_j$ obținem :

$$\begin{aligned} P(B_1 \cap \dots \cap B_n) &= P\left(\bigcup_{A_1 \in \Delta_1} \dots \bigcup_{A_n \in \Delta_n} (A_1 \cap \dots \cap A_n) \right) = \\ &= \sum_{A_j \in \Delta_j} P(A_1 \cap \dots \cap A_n) = \sum_{A_j \in \Delta_j} P(A_1) \dots P(A_n) = \\ &= \left[\sum_{A_1 \in \Delta_1} P(A_1) \right] \dots \left[\sum_{A_n \in \Delta_n} P(A_n) \right] = P\left(\bigcup_{A_1 \in \Delta_1} A_1 \right) \dots P\left(\bigcup_{A_n \in \Delta_n} A_n \right) = \\ &= P(B_1) \dots P(B_n). \end{aligned}$$

1.20. Fie $(E_i, \mathfrak{E}(E_i), P_i)_{1 \leq i \leq n}$, n cîmpuri clasice de probabilitate.

a) Să se arate că cîmpul de probabilitate produs $\otimes (E_i, \mathfrak{E}(E_i), P_i)$ este cîmpul clasic de probabilitate atașat mulțimii $\prod_{i=1}^n E_i$.

b) Fie $A_i \subset E_i$ și $B_i = E_1 \times \dots \times E_{i-1} \times A_i \times E_{i+1} \times \dots \times E_n$. Să se arate că evenimentele B_1, \dots, B_n sînt independente.

Observație. Punctul 2) este adevărat în general în sensul că se poate înlocui cîmpul clasic $(E_i, \mathfrak{E}(E_i), P_i)$ cu un cîmp de probabilitate oarecare $(\Omega_i, \mathfrak{F}_i, Q_i)$ (în acest caz $B_i \in \bigotimes_i \mathfrak{F}_i$).

Soluție. a) Avem de arătat că $\bigotimes_{i=1}^n \mathfrak{E}(E_i) = \mathfrak{E}\left(\prod_{i=1}^n E_i\right)$ și că $\bigotimes_{i=1}^n P_i$ este probabilitatea clasică pe $\prod_{i=1}^n E_i$. Incluziunea $\bigotimes_{i=1}^n \mathfrak{E}(E_i) \subset \mathfrak{E}\left(\prod_{i=1}^n E_i\right)$ este imediată.

Pentru incluziunea inversă este suficient să observăm (întrucît $\prod_{i=1}^n E_i$ este finită) că $\{(l_1, \dots, l_n)\} \in \bigotimes_{i=1}^n \mathfrak{E}(E_i)$

Cum $\mathfrak{E}\left(\prod_{i=1}^n E_i\right) = \mathfrak{B}(\Delta)$, unde Δ este partiția cu elementele $\{l\}_{l \in \prod_{i=1}^n E_i}$ rezultă că $\bigotimes_{i=1}^n P_i$ coincide cu probabilitatea clasică dacă și numai dacă $\left(\bigotimes_{i=1}^n P_i\right)(\{(l_1, \dots, l_n)\}) = \frac{1}{\prod_i \text{card } E_i}$ (vezi problema 1.14).

Or

$$\left(\bigotimes_{i=1}^n P_i\right)(\{(l_1, \dots, l_n)\}) = \prod_{i=1}^n P_i(\{l_i\}) = \prod_{i=1}^n \frac{1}{\text{card } E_i} = \frac{1}{\prod_i \text{card } E_i}.$$

b) Dacă $J \subset \{1, \dots, n\}$, atunci $\bigcap_{i \in J} B_i = C_1 \times \dots \times C_n$, unde $C_i = A_i$ dacă $i \in J$ și $C_i = E_i$ în caz contrar. Avem deci

$$P\left(\bigcap_{i \in J} B_i\right) = P(C_1 \times \dots \times C_n) = \prod_{i=1}^n P_i(C_i) = \prod_{i \in J} P_i(A_i) = \prod_{i \in J} P(B_i),$$

deci evenimentele B_1, \dots, B_n sînt independente.

1.21. Fie E o mulțime finită cu n elemente. Să se arate că dacă $n > 3$, atunci există o probabilitate P pe $\mathfrak{E}(E)$ și n evenimente A_1, \dots, A_n independente două cîte două cu $0 < P(A_i)$ pentru orice i și $P(A_j) < 1$ pentru $2 \leq j \leq n$.

Soluție. Fie $n > 3$, $E = \{1, \dots, n\}$ și P probabilitatea pe $\mathfrak{E}(E)$ definită prin $P(\{1\}) = 1 - (n-1)x$, $P(\{i\}) = x$, $i = 2, \dots, n$, unde $0 < x < 1$ este soluția ecuației $1 - (n-1)x = [1 - (n-2)x]^2$, adică $x = \frac{n-3}{(n-2)^2}$.

Definim evenimentele

$$A_1 = \{1, \dots, n\} \quad A_i = \{1, i\}, \quad i = 2, \dots, n.$$

Avem $P(A_1) = 1$, $P(A_j) = 1 - (n-2)x$, și evident $P(A_j) > 0$ pentru orice $1 \leq j \leq n$ și $P(A_j) < 1$ dacă $2 \leq j \leq n$. Este vizibil că pentru orice $2 \leq i \leq n$ evenimentele A_1, A_i sînt independente. Să arătăm că pentru orice $2 \leq i, j \leq n$ evenimentele A_i, A_j sînt independente.

Într-adevăr

$$P(A_i \cap A_j) = P(\{1\}) = 1 - (n-1)x$$

$$P(A_i)P(A_j) = [1 - (n-2)x]^2$$

deci $P(A_i \cap A_j) = P(A_i)P(A_j)$ dacă avem în vedere ecuația ce o verifică x .

1.22. Dacă într-un câmp de probabilitate (E, \mathcal{K}, P) (unde E este finită) există evenimentele independente A_1, \dots, A_n cu $0 < P(A_i) < 1$ pentru orice i , să se arate că $n \leq \lg_2(\text{card } E)$

Să se construiască apoi un câmp de probabilitate (E, \mathcal{K}, P) cu proprietățile din enunț.

Soluție. Dacă punem $A_i^0 = A_i, A_i^1 = \bar{A}_i$ atunci pentru orice vector $(\varepsilon_1, \dots, \varepsilon_n)$, unde $\varepsilon_i = 0$ sau 1 , avem

$$P\left(\bigcap_{i=1}^n A_i^{\varepsilon_i}\right) = \prod_{i=1}^n P(A_i^{\varepsilon_i}) > 0, \text{ deci în particular } \bigcap_{i=1}^n A_i^{\varepsilon_i} \neq \emptyset.$$

Deoarece pentru $(\varepsilon_1, \dots, \varepsilon_n) \neq (\varepsilon'_1, \dots, \varepsilon'_n)$ evenimentele $\bigcap_{i=1}^n A_i^{\varepsilon_i}$ și

$\bigcap_{i=1}^n A_i^{\varepsilon'_i}$ sînt disjuncte și sînt în număr de 2^n rezultă că $2^n \leq$

$\leq \text{card}\left(\bigcup_{\varepsilon_i} A_i^{\varepsilon_i}\right) \leq \text{card } E$, deci $n \leq \lg_2(\text{card } E)$.

În continuare fie $E = \{(\varepsilon_1, \dots, \varepsilon_n); \varepsilon_i = 0 \text{ sau } 1\}$ și considerăm câmpul clasic de probabilitate $(E, \mathfrak{E}(E), P)$.

Atunci evenimentele A_1, \dots, A_n unde $A_i = \{(\varepsilon_1, \dots, \varepsilon_n); \varepsilon_i = 0\}$ satisfac cerințele (adică sînt independente și $0 < P(A_i) < 1$).

Într-adevăr cu notația $\{0, 1\}_j = \{0, 1\}$ avem

$$P(A_i) = \frac{\text{card } A_i}{\text{card } E} = \frac{\text{card} \prod_{\substack{j=1 \\ j \neq i}}^n \{0, 1\}_j}{2^n} = \frac{\prod_{i \neq j} \text{card } \{0, 1\}_j}{2^n} = \frac{2^{n-1}}{2^n} = \frac{1}{2}$$

$$\begin{aligned}
 P(A_{i_1} \cap \dots \cap A_{i_m}) &= P(\{(\varepsilon_1, \dots, \varepsilon_n); \varepsilon_{i_1} = \dots = \varepsilon_{i_m} = 0\}) = \\
 &= \frac{\text{card} \prod_{j \neq i_1, \dots, i_m} \{0, 1\}_j}{2^n} = \frac{\prod_{j \neq i_1, \dots, i_m} \text{card} \{0, 1\}_j}{2^n} = \frac{2^{n-m}}{2^n} = \frac{1}{2^m} = \\
 &= P(A_{i_1}) \dots P(A_{i_m}), \text{ deci } A_1, \dots, A_n \text{ sînt independente.}
 \end{aligned}$$

1.23. Evenimentele A_1, \dots, A_n satisfac condițiile $P\left(\bigcap_{j=1}^i A_j\right) = \prod_{j=1}^i P(A_j)$ pentru orice i .

Sînt aceste evenimente independente?

Soluție. Din problema 1.19 rezultă că în cazul $n = 2$ răspunsul este afirmativ. Arătăm printr-un exemplu că afirmația este falsă dacă $n \geq 3$. Considerăm cîmpul de probabilitate $((0, 1), \mathfrak{B}_{(0, 1)}, m)$, m măsura Lebesgue și evenimentele $A_1 = \left(0, \frac{1}{2}\right)$, $A_2 = \left(\frac{1}{4}, \frac{3}{4}\right)$, $A_3 = \left(\frac{3}{8}, \frac{7}{8}\right)$. Avem $m(A_i) = \frac{1}{2}$ și $m(A_1 \cap A_2) = m\left(\left(\frac{1}{4}, \frac{1}{2}\right)\right) = \frac{1}{4} = m(A_1)m(A_2)$

$$m(A_1 \cap A_2 \cap A_3) = m\left(\left(\frac{3}{8}, \frac{1}{2}\right)\right) = \frac{1}{8} = m(A_1)m(A_2)m(A_3),$$

deci A_1, A_2, A_3 satisfac cerințele din enunțul problemei.

Dar A_1, A_2, A_3 nu sînt independente deoarece

$$m(A_1 \cap A_3) = m\left(\left(\frac{3}{8}, \frac{1}{2}\right)\right) = \frac{1}{8} \neq m(A_1)m(A_3) = \frac{1}{4}.$$

1.24. Să se arate că dacă în problema 1.19 egalitatea (1) nu este adevărată pentru orice J , nu rezultă că evenimentele A_1, \dots, A_n sînt independente.

Soluție. Următorul exemplu datorat lui Bernstein justifică acest lucru. Considerăm cîmpul clasic de probabilitate $(\{1, 2, 3, 4\}, \mathfrak{B}(\{1, 2, 3, 4\}), P)$ și evenimentele $A_i = \{i, 4\}$; $i = 1, 2, 3$. Dacă $i \neq j$, avem

$$P(A_i \cap A_j) = P(\{4\}) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(A_i)P(A_j).$$

Însă

$$P(A_1 \cap A_2 \cap A_3) = P(\{4\}) = \frac{1}{4} \neq \frac{1}{8} = P(A_1)P(A_2)P(A_3),$$

deci A_1, A_2, A_3 nu sînt independente.

1.25. Fie (E, \mathcal{K}) un spațiu măsurabil, $(\mathfrak{F}_i)_{i \in I}$ o familie de corpuri boreliene incluse în \mathcal{K} , P_1, P_2 două probabilități pe \mathcal{K} așa încît $P_1 = P_2$ pe orice \mathfrak{F}_i și $(\mathfrak{F}_i)_i$ să fie independente atît relativ la P_1 cît și la P_2 .

Să se arate că $P_1 = P_2$ pe $\mathfrak{B}\left(\bigcup_i \mathfrak{F}_i\right)$.

Soluție. Observăm că $\mathfrak{B}\left(\bigcup_i \mathfrak{F}_i\right) = \mathfrak{B}\left(\bigcap_i A_i; A_i \in \mathfrak{F}_i \text{ și } \{i; A_i \neq E\} \text{ este finită}\right)$.

Conform teoremei de unicitate a probabilităților avem de arătat că

$$P_1\left(\bigcap_i A_i\right) = P_2\left(\bigcap_i A_i\right)$$

dacă $A_i \in \mathfrak{F}_i$ și $\{i; A_i \neq E\}$ este finită. Avem din independență

$$P_1\left(\bigcap_i A_i\right) = \prod_i P_1(A_i) = \prod_i P_2(A_i) = P_2\left(\bigcap_i A_i\right).$$

1.26. Fie A_1, \dots, A_n evenimente independente cu $P(A_i) = p$ pentru orice i . Să se arate că probabilitatea realizării a r dintre aceste evenimente este $C_n^r p^r (1-p)^{n-r}$.

Soluție. Se poate aplica rezultatul problemei 1.6, însă este mai ușor de făcut calculul direct.

Evenimentul a cărui probabilitate se cere se scrie sub forma

$$F = \bigcup_{\text{card } L=r} B_L, \text{ unde}$$

$$B_L = \left(\bigcap_{i \in L} A_i\right) \cap \left(\bigcap_{i \notin L} \mathcal{C}A_i\right).$$

Avem

$$P(B_L) = \prod_{i \in L} P(A_i) \prod_{i \notin L} P(\mathcal{C}A_i) = p^{\text{card } L} (1-p)^{n-\text{card } L},$$

deci

$$P(F) = \sum_{\text{card } L=r} P(B_L) = \sum_{\text{card } L=r} p^r (1-p)^{n-r} = C_n^r p^r (1-p)^{n-r}.$$

1.27. Într-un câmp de probabilitate (E, \mathcal{X}, P) considerăm o partiție A_1, \dots, A_r a evenimentelor elementare (deci dacă $P(A_i) = p_i$, atunci $p_i \geq 0$ și $p_1 + \dots + p_r = 1$). Presupunem că repetăm de n ori (în mod independent) experimentul aleator ξ care este modelat de (E, \mathcal{X}, P) .

Fie întregii $n_i \geq 0$ cu $n_1 + \dots + n_r = n$. Să se arate că probabilitatea ca pentru orice i evenimentul A_i să se producă de n_i ori, în cele n experimente este egală cu $\frac{n!}{n_1! \dots n_r!} p_1^{n_1} \dots p_r^{n_r}$. (Această problemă generalizează pe cea precedentă).

Soluție. În continuare prin $\alpha = (\alpha_1, \dots, \alpha_r)$ vom desemna o partiție a lui $\{1, \dots, n\}$, cu card $\alpha_i = n_i$.

Fie $A_{ij} = E \times \dots \times E \times A_i \times E \times \dots \times E$ și $B_\alpha = \bigcap_{i=1}^r \bigcap_{j \in \alpha_i} A_{ij}$. Atunci evenimentul a cărui probabilitate se cere este $B = \bigcup_{\alpha} B_\alpha$.

Câmpul de probabilitate ce modelează experimentul constituit din repetarea independentă de n ori a lui este produsul $\bigotimes_{i=1}^n (E, \mathcal{X}, P)_i$,

unde $(E, \mathcal{X}, P)_i = (E, \mathcal{X}, P)$ pentru orice i . Din problema 1.20 rezultă că corpurile boreliene $\mathcal{X}_1, \dots, \mathcal{X}_n$, unde $\mathcal{X}_i = \mathcal{X}$, sînt independente și cum prin ipoteză pentru j fixat evenimentele A_{1j}, \dots, A_{rj} sînt independente rezultă din dezasociativitatea independenței că evenimentele $(A_{ij})_{i,j}$ sînt independente (vezi n din rezumatul teoretic).

Aplicînd apoi asociativitatea independenței (vezi m din rezumatul teoretic) rezultă că evenimentele $\left(\bigcap_{j \in \alpha_i} A_{ij}\right)_i$ sînt independente. Putem deci scrie

$$P(B_\alpha) = \prod_{i=1}^r \prod_{j \in \alpha_i} P(A_{ij}) = \prod_{i=1}^r p_i^{n_i},$$

$$P(B) = \sum_{\alpha} P(B_\alpha) = \sum_r \prod_{i=1}^r p_i^{n_i} = \frac{n!}{n_1! \dots n_r!} \prod_{i=1}^r p_i^{n_i}$$

(am aplicat la ultima egalitate formula f_4 din rezumatul teoretic).

1.28. Fie A_1, \dots, A_n evenimente independente cu $P(A_i) = p$ pentru orice i . Să se arate că probabilitatea ca să se realizeze A_n și $r-1$ dintre A_1, \dots, A_{n-1} este $C_{n-1}^{r-1} p^r (1-p)^{n-r}$.

Soluție. Fie A evenimentul că exact $r - 1$ dintre A_1, \dots, A_{n-1} se realizează. Atunci evenimentul a cărui probabilitate se cere este $B = A \cap A_n$. Cum din asociativitatea independenței rezultă că evenimentele A și A_n sînt independente, putem scrie

$$P(B) = P(A)P(A_n) = C_{n-1}^{r-1} p^{r-1} (1-p)^{n-r} \cdot p = C_{n-1}^{r-1} p^r (1-p)^{n-r}$$

(am folosit mai sus problema 1.26).

1.29. Să se arate că dacă $P(\overline{\lim} A_n) = 0$, nu rezultă că $\sum_n P(A_n) < \infty$ (lema Borel-Cantelli afirmă că afirmația reciprocă este adevărată).

Soluție. Considerăm cîmpul de probabilitate $((0,1), \mathfrak{B}_{(0,1)}, m)$, m fiind măsura Lebesgue, și fie $A_n = (0, \frac{1}{n})$.

Deoarece $A_{n+1} \subset A_n$ rezultă că $\overline{\lim}_n A_n = \bigcap_{n=1}^{\infty} A_n = \emptyset$, deci $m(\overline{\lim}_n A_n) = 0$. Însă $\sum_n m(A_n) = \sum_n \frac{1}{n} = \infty$.

1.30. Dacă (A_n) este un șir de evenimente cu proprietatea că există un subșir (A_{n_j}) de evenimente independente așa încît $\sum_j P(A_{n_j}) = \infty$, să se arate că $P(\overline{\lim}_{n \rightarrow \infty} A_n) = 1$.

Soluție. Din lema Borel-Cantelli rezultă că $P(\overline{\lim}_i A_{n_i}) = 1$. Apoi din incluziunea $\overline{\lim}_i A_{n_i} \subset \overline{\lim}_n A_n$ rezultă că $P(\overline{\lim}_n A_n) = 1$.

1.31. Fie (A_n) un șir de evenimente independente. Să se arate că $P(\bigcup_n A_n) = 1$ dacă $\sum_n P(A_n) = \infty$.

Soluție. $P(\bigcup_n A_n) = 1 \Leftrightarrow P(\mathcal{C}(\bigcup_n A_n)) = 0 \Leftrightarrow P(\bigcap_n \mathcal{C} A_n) = 0 \Leftrightarrow \prod_{n=1}^{\infty} P(\mathcal{C} A_n) = 0$. Or avînd în vedere inegalitatea $1 - \alpha \leq e^{-\alpha}$, $0 \leq \alpha \leq 1$, rezultă că

$$\prod_n P(\mathcal{C} A_n) = \prod_n [1 - P(A_n)] = 0 \text{ dacă } \sum_n P(A_n) = \infty$$

1.32. Fie (A_n) un șir de evenimente independente și fie $B_k^m = A_k \cap \dots \cap A_{k+m}$.

a) Să se arate că pentru m fixat avem $P(\overline{\lim}_k B_k^m) = 1$ dacă și numai dacă

$$\sum_k \prod_{j=0}^m P(A_{k+j}) = \infty. \quad (1)$$

Să se deducă apoi că $P\left(\bigcap_m \bigcup_k B_k^m\right) = 1$ dacă (1) este satisfăcută.

b) Dacă $\sum_n P(\complement A_n) = \infty$ să se arate că $P\left(\bigcap_m \bigcup_k B_k^m\right) = 0$.

Soluție. a) Faptul că $P\left(\overline{\lim}_k B_k^m\right) = 1$ implică (1) rezultă din lema Borel-Cantelli.

Dacă (1) are loc rezultă în particular că $\sum_k P(A_{k+j}) = \infty$ pentru orice $0 \leq j \leq m$.

Să arătăm că $P\left(\overline{\lim}_k B_k^m\right) = 0$ sau echivalent că $P\left(\lim_k \complement B_k^m\right) = 0$.

Se observă că totul se reduce la a arăta că $P\left(\bigcap_{k \geq n} \bigcup_{j=0}^m \complement A_{k+j}\right) = 0$ sau încă la $P\left(\bigcap_{k \geq n} \complement A_{k+j}\right) = 0$ pentru orice $0 \leq j \leq m$. Se aplică acum problema precedentă.

Mai departe tot din problema precedentă rezultă că $P\left(\bigcup_k B_k^m\right) = 1$ dacă (1) este verificată, deci și $P\left(\bigcap_m \bigcup_k B_k^m\right) = 1$ (intersecția numărabilă de evenimente cu probabilitatea 1 are probabilitatea 1).

b) Dacă (2) are loc, atunci din lema Borel-Cantelli rezultă că $P\left(\overline{\lim}_n \complement A_n\right) = 1$ și cum are loc incluziunea

$$\overline{\lim}_m \complement A_n \subset \bigcup_m \bigcap_k \bigcup_{j=0}^m \complement A_{k+j} = \complement \left(\bigcap_m \bigcup_k B_k^m \right),$$

rezultă că $P\left(\complement \left(\bigcap_m \bigcup_k B_k^m \right)\right) = 1$ sau echivalent $P\left(\bigcap_m \bigcup_k B_k^m\right) = 0$.

1.33. Fie $C([0, 1], R)$ considerat ca spațiu metric separabil și complet în raport cu metrica $d(f, g) = \sup |f(t) - g(t)|$ și fie $\pi_t: C([0, 1], R) \rightarrow R$ proiecția canonică definită prin $\pi_t(f) = f(t)$. Să se arate că

$$\mathfrak{B}_{C([0, 1], R)} = \mathfrak{B}(\pi_t; t \in [0, 1]).$$

Să se deducă apoi că două probabilități P_1, P_2 pe $\mathfrak{B}_{C([0, 1], R)}$ coincid dacă și numai dacă coincid pe mulțimile [de forma $\bigcap_{i=1}^m \{\pi_{t_i} \in A_i\}$, unde $A_i \in \mathfrak{B}_R$.

Soluție. Cum fiecare π_t este continuă, deci măsurabilă, rezultă că avem $\mathfrak{B}(\pi_t, t) \subset \mathfrak{B}_{C([0, 1], R)}$.

Deoarece $C([0, 1], R)$ este separabil, rezultă că orice mulțime deschisă este reuniune numărabilă de sfere închise, deci pentru incluziunea inversă este suficient să probăm că sferile închise sînt în $\mathfrak{B}(\pi_t; t)$.

Fie $f_0 \in C([0, 1], R)$ și $r < 0$. Avem

$$\begin{aligned} S_r(f_0) &= \{f; d(f, f_0) \leq r\} = \bigcap_{t \in Q \cap [0, 1]} \{f; |f(t) - f_0(t)| \leq r\} = \\ &= \bigcap_{t \in Q \cap [0, 1]} \{f; |\pi_t(f) - \pi_t(f_0)| \leq r\} \in \mathfrak{B}(\pi_t; t) \end{aligned}$$

pentru că $\{\pi_t - \pi_t(f_0)\}$ este $\mathfrak{B}(\pi_u; u)$ — măsurabilă și intersecția este numărabilă.

Este vizibil că are loc și egalitatea $\mathfrak{B}(\pi_t; t) = \mathfrak{B}\left(\bigcap_{i=1}^m \{\pi_{t_i} \in A_i\}; n \text{ arbitrar și } A_i \in \mathfrak{B}_R\right)$ și cum familia $\mathfrak{M} = \left\{\bigcap_{i=1}^m \{\pi_{t_i} \in A_i\}; m \text{ arbitrar și } A_i \in \mathfrak{B}_R\right\}$ este închisă la intersecția finită, rezultă din teorema de unicitate a probabilităților că P_1 și P_2 coincid dacă și numai dacă coincid pe \mathfrak{M} .

1.34. Fie E un spațiu metric și P_1, P_2 probabilități pe \mathfrak{B}_E . Să se arate că dacă:

- $P_1(G) = P_2(G)$ pentru orice mulțime deschisă G sau
- $\int f dP_1 = \int f dP_2$ pentru orice $f: E \rightarrow R$ continuă și mărginită;

Rezultă :

$$P_1 = P_2.$$

Dacă E este local compact și cu bază numărabilă și

a') $P_1(K) = P_2(K)$ pentru orice mulțime compactă K sau

b') $\int f dP_1 = \int f dP_2$ pentru orice $f: E \rightarrow R$ continuă și cu suport compact, atunci $P_1 = P_2$.

Soluție. Deoarece mulțimile deschise generează pe \mathfrak{B}_E și sînt închise la intersecția finită rezultă din teorema de unicitate a probabilităților că a) implică $P_1 = P_2$.

Să presupunem acum că are loc b) și să arătăm că $P_1(G) = P_2(G)$ pentru orice deschis G .

Funcțiile continue $f_n(x) = \min(1, nd(x, \mathcal{C}G))$ converg crescător către λ_G și $0 \leq f_n \leq 1$. Prin ipoteză $\int f_n dP_1 = \int f_n dP_2$ de unde aplicînd teorema de convergență dominată obținem $\int \lambda_G dP_1 = \int \lambda_G dP_2$ adică $P_1(G) = P_2(G)$.

Să presupunem acum că a') este satisfăcută. Să alegem un șir de compacti K_n așa încît $K_n \subset K_{n+1}$ și $E = \bigcup_n K_n$.

Atunci dacă F este mulțime închisă, putem scrie $F = \bigcup (K_n \cap F)$, unde $K_n \cap F$ sînt compacte, de unde

$$P_1(F) = \lim_{n \rightarrow \infty} P_1(K_n \cap F) = \lim_{n \rightarrow \infty} P_2(K_n \cap F) = P_2(F)$$

deci $P_1 = P_2$ dacă avem în vedere teorema de unicitate a probabilităților. În fine presupunem că (b') are loc. Conform punctului b) este suficient să arătăm că $\int f dP_1 = \int f dP_2$ pentru f continuă și mărginită.

Fie $\varphi_n: E \rightarrow R$ continuă așa încît $0 \leq \varphi_n \leq 1$, $\varphi_n = 1$ pe K_n și $\varphi_n = 0$ pe CK_n . Atunci $f\varphi_n \nearrow f$, $f\varphi_n$ sînt continue și cu suport compact și uniform mărginite.

Deci în egalitatea $\int f\varphi_n dP_1 = \int f\varphi_n dP_2$ putem trece la limita sub integrală pe baza teoremei de convergență dominată și se obține $\int f dP_1 = \int f dP_2$.

1.35. Dacă E este spațiu metric, să se arate că \mathfrak{B}_E este cea mai mică clasă \mathcal{A} de submulțimi ale lui E care conține mulțimile deschise (sau închise) și este închisă la intersecțiile și reuniunile numărabile.

Aplicație. Dată fiind o probabilitate P pe \mathfrak{B}_E , să se probeze egalitățile:

$$P(A) = \sup_{\substack{F \subset A \\ F \text{ închisă}}} P(F) = \inf_{\substack{D \supset A \\ D \text{ deschisă}}} P(D)$$

Soluție. Deoarece \mathfrak{B}_E are proprietățile din definiția lui \mathcal{A} , rezultă că $\mathfrak{B}_E \supset \mathcal{A}$.

Fie $\mathcal{A}' = \{A; A \in \mathcal{A}, \complement A \in \mathcal{A}\}$. Este ușor de văzut că \mathcal{A}' este corp borelian și cum orice mulțime închisă este intersecție numărabilă de mulțimi deschise (respectiv orice mulțime deschisă este reuniune numărabilă de mulțimi închise) rezultă că $\mathcal{A}' \supset \mathfrak{B}_E$. Dar cum $\mathcal{A}' \supset \mathcal{A}$, deci $\mathcal{A}' = \mathcal{A}$, rezultă că $\mathcal{A} \supset \mathfrak{B}_E$.

Aplicație. Fie $\mathcal{A} = \{A \in \mathfrak{B}_E; P(A) = \sup_{\substack{F \subset A \\ F \text{ închisă}}} P(F)\}$. Vom arăta că \mathcal{A} este închisă la intersecțiile și reuniunile numărabile și evident include mulțimile închise (deci va rezulta că $\mathcal{A} = \mathfrak{B}_E$).

Fie deci A_n așa încît $P(A_n) = \sup_{\substack{F \subset A_n \\ F \text{ închisă}}} P(F)$, deci pentru $\varepsilon > 0$ există $F_n \subset A_n$, F_n închisă așa încît $P(A_n) - P(F_n) \leq \frac{\varepsilon}{2^n}$. Fie mulțimea închisă

$\bigcap_n F_n$ care este inclusă în $\bigcap_n A_n$. Deoarece $\bigcap_n A_n \setminus \bigcap_n F_n \subset \bigcup_n (A_n \setminus F_n)$ rezultă că

$$P\left(\bigcap_n A_n\right) - P\left(\bigcap_n F_n\right) \leq \sum_n P(A_n \setminus F_n) \leq \sum_n \frac{\varepsilon}{2^n} = \varepsilon$$

cu alte cuvinte $\bigcap_n A_n \in \mathcal{A}$. Pe de altă parte, deoarece $P\left(\bigcup_n A_n\right) = \lim_m P\left(\bigcup_{n=1}^m A_n\right)$ rezultă că există m așa încît $P\left(\bigcup_{n=1}^\infty A_n\right) - P\left(\bigcup_{n=1}^m A_n\right) \leq \varepsilon$.

Avem $P\left(\bigcup_{n=1}^m A_n\right) - P\left(\bigcup_{n=1}^m F_n\right) \leq P\left(\bigcup_{n=1}^m (A_n \setminus F_n)\right) \leq \sum_{n=1}^m P(A_n \setminus F_n) \leq \sum_{n=1}^m \frac{\varepsilon}{2^n} \leq \varepsilon$ și prin urmare $P\left(\bigcup_{n=1}^\infty A_n \setminus \bigcup_{n=1}^m F_n\right) \leq P\left(\bigcup_{n=1}^\infty A_n\right) - P\left(\bigcup_{n=1}^m A_n\right) + P\left(\bigcup_{n=1}^m A_n\right) - P\left(\bigcup_{n=1}^m F_n\right) \leq 2\varepsilon$, cu alte cuvinte $\bigcup_{n=1}^\infty A_n \in \mathcal{A}$.

Mai departe avem :

$$\begin{aligned}
 P(\mathcal{C}A) = 1 - P(A) &= \sup_{\substack{F \subset \mathcal{C}A \\ F \text{ închisă}}} P(F) = \sup_{\substack{\mathcal{C}F \supset A \\ F \text{ închisă}}} [1 - \mathcal{C}(CF)] = \\
 &= \sup_{\substack{D \supset A \\ D \text{ deschisă}}} [1 - P(D)] = 1 - \inf_{\substack{D \supset A \\ D \text{ deschisă}}} P(D)
 \end{aligned}$$

cu alte cuvinte $P(A) = \inf_{\substack{D \supset A \\ D \text{ deschisă}}} P(D)$.

1.36. Fie (E, \mathcal{X}, P) un câmp de probabilitate cu proprietatea că pentru orice $A \in \mathcal{X}$ cu $P(A) > 0$ există $B \in \mathcal{X}$ așa încît $0 < P(B) < P(A)$ (un astfel de câmp de probabilitate se numește neatomic).

Să se arate că $P(\mathcal{X}) = [0, 1]$.

Soluție. Pasul 1. Arătăm că pentru orice $A \in \mathcal{X}$ cu $P(A) > 0$ și $\varepsilon > 0$ există $B \in \mathcal{X}$ cu $B \subset A$ și $0 < P(B) \leq \varepsilon$.

Într-adevăr, dacă $P(A) \leq \varepsilon$, luăm $B = A$. Dacă $P(A) > 0$, atunci fie $B \subset A$ cu $0 < P(B) < P(A)$.

Deoarece $P(B) + P(A \setminus B) = P(A)$, rezultă că B sau $A \setminus B$ are probabilitatea $\leq \frac{1}{2} P(A)$. Fie B'_1 acela dintre ei, deci $0 < P(B'_1) \leq \frac{P(A)}{2}$. Dacă $\frac{P(A)}{2} \leq \varepsilon$, luăm mulțimea căutată egală cu B'_1 . Dacă $\frac{P(A)}{2} > \varepsilon$, atunci reluăm raționamentul și obținem $B_1 \subset B'_1$ așa încît unul dintre evenimentele $B'_1 \setminus B_1$ și B_1 (fie acela B'_2) are proprietatea că $0 < P(B'_2) \leq \frac{P(B_1)}{2}$.

Dacă $\frac{P(A)}{2^2} \leq \varepsilon$, deci dacă $\frac{P(B'_2)}{2^2} \leq \varepsilon$, luăm mulțimea căutată egală cu B'_2 . Dacă $\frac{P(A)}{2^2} > \varepsilon$, continuăm raționamentul.

Deoarece există un întreg r așa încît $\frac{P(A)}{2^r} \leq \varepsilon$, rezultă că, după un număr finit de pași, vom găsi $B \subset A$ cu $0 < P(B) \leq \varepsilon$.

Pasul 2. Arătăm că pentru orice $A \in \mathcal{X}$ cu $P(A) > 0$ și $\varepsilon > 0$ există o partiție finită A_1, \dots, A_n a lui A așa încît $0 < P(A_i) \leq \varepsilon$ pentru orice i .

Pentru aceasta, considerăm întâi cazul $A = E$. Aplicînd pasul 1 pentru $A = E$ obținem un eveniment A_1 , așa încît $0 < P(A_1) \leq \varepsilon$.

Să notăm $\mu_\varepsilon(B) = \sup_{\substack{C \subset B \\ P(C) \leq \varepsilon}} P(C)$, $B \in \mathcal{X}$. Evident $\mu_\varepsilon(B) > 0$, dacă $P(B) > 0$ (deoarece din pasul 1 există $C \subset B$ cu $0 < P(C) \leq \varepsilon$).

Fie $A_2 \subset \mathcal{C}A_1$ cu $\varepsilon \geq P(A_2) \geq \frac{1}{2} \mu_\varepsilon(\mathcal{C}A_1)$. Apoi fie $A_3 \subset \mathcal{C}(A_1 \cup A_2) \cup A_2$ așa încît $\varepsilon \geq P(A_3) \geq \frac{1}{2} \mu_\varepsilon(\mathcal{C}(A_1 \cup A_2))$ și în general fie $A_{n+1} \subset \mathcal{C}(A_1 \cup \dots \cup A_n)$ așa încît $\varepsilon \geq P(A_{n+1}) \geq \frac{1}{2} \mu_\varepsilon(\mathcal{C}(A_1 \cup \dots \cup A_n))$.

Deoarece A_1, \dots, A_n, \dots sînt disjuncte, rezultă $\sum_{n=1}^{\infty} P(A_n) = P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq 1$, deci în particular $P(A_n) \rightarrow 0$ și prin urmare $\mu_\varepsilon(\mathcal{C}(A_1 \cup \dots \cup A_n)) \rightarrow 0$. Luînd $B = \bigcup_{n=1}^{\infty} A_n$ și avînd în vedere că μ_ε este crescătoare obținem că $\mu_\varepsilon(\mathcal{C}B) = 0$, deci $P(\mathcal{C}B) = 0$.

Fie $A'_1 = A_1 \cup \mathcal{C}B$; atunci $A'_1 \cup \left(\bigcup_{n=2}^{\infty} A_n\right) = E$ și $0 < P(A_n) \leq \varepsilon$. Fie N suficient de mare așa încît $\sum_{n=N}^{\infty} P(A_n) \leq \varepsilon$.

Atunci $A'_1, A_2, \dots, A_{N-1}, \bigcup_{n=N}^{\infty} A_n$ este partiția căutată. În general pentru A oarecare se consideră partiția $A \cap A_1, \dots, A \cap A_n$, unde A_1, \dots, A_n este partiția corespunzătoare lui E .

Pasul 3. Vom arăta că pentru orice $0 < x < 1$ există un eveniment A cu $P(A) = x$ (cazurile $x = 0, 1$ sînt imediate).

Fie A_{11}, \dots, A_{1n_1} o partiție a lui E , așa încît $0 < P(A_{1j}) \leq \frac{x}{2}$ și fie

$$x_{1,r} = \sum_{j=1}^r P(A_{1j}), \quad 1 \leq r \leq n_1 - 1.$$

Atunci x aparține la unul din intervalele $[x_{1,r}, x_{1,r+1})$, $1 \leq r \leq n_1 - 1$.

Fie $[x_{1,r}, x_{1,r+1})$ acest interval. Dacă $x = x_{1,r_1}$ atunci se ia $A = \bigcup_{j=1}^{r_1} A_{1j}$.

Dacă $x > x_1, r_1$, fie A_{21}, \dots, A_{2n} , o partiție a lui A_{1, r_1+1} așa încît

$$0 < P(A_{2j}) \leq \frac{x - x_1, r_1}{2} \text{ și } x_{2r} = P\left(\bigcup_{j=1}^{r_1} A_{1j} \cup \bigcup_{j=1}^r A_{2j}\right), \quad 1 \leq r \leq n_2 - 1.$$

Atunci x aparține la unul din intervalele $[x_2, r, x_2, r+1)$. Continuînd procedeul, vom obține evenimentul

$$A = \bigcup_{i=1}^{\infty} \left(\bigcup_{j=1}^{r_1} A_{ij} \right) \text{ ce are proprietatea că } P(A) = x.$$

§2. Probabilitatea clasică și probabilitatea geometrică

1.37. Să se afle probabilitatea ca alegînd la întîmplare un număr din mulțimea $\{1, \dots, n\}$ acesta să se dividă prin numărul natural k .

Soluție. Evenimentele elementare (posibile) sînt $E = \{1, \dots, n\}$ iar evenimentele favorabile sînt $F = \{a \in \{1, \dots, n\}; a \text{ se divide cu } k\}$. Rezultă

$$P(F) = \frac{\text{card } F}{\text{card } E} = \frac{\left[\frac{n}{k} \right]}{n}.$$

1.38. Din mulțimea $\{1, \dots, n\}$ se extrage la întîmplare un număr a ; să se afle probabilitatea ca $a^2 - 1$ să se dividă cu 10.

Soluție. Evenimentele elementare sînt $E = \{1, \dots, n\}$. Deoarece $a^2 \equiv 1 \pmod{10}$ dacă și numai dacă $a \equiv 1 \pmod{10}$ sau $a \equiv 9 \pmod{10}$ rezultă că evenimentele favorabile sînt $F = \{a \in \{1, \dots, n\}; a \equiv 1 \pmod{10} \text{ sau } a \equiv 9 \pmod{10}\} = \{a \in \{1, \dots, n\}; a \text{ are cifre unităților } 1 \text{ sau } 9\}$.

Fie $n = 10k + l$, $0 \leq l < 9$.

Cazul $l = 0$. În acest caz $F = \{1, 10l - 1, 10l + 1, 1 \leq l \leq k - 1, 10(k - 1) + 9\}$, deci $\text{card } F = 2k$ și prin urmare $P(F) = \frac{\text{card } F}{\text{card } E} = \frac{2k}{n}$.

Cazul $0 < l < 9$. În acest caz $F = \{1, 10l \pm 1, 1 \leq l \leq k\}$, deci $\text{card } F = 2k + 1$, și prin urmare $P(F) = \frac{2k + 1}{n}$.

Cazul $l = 9$. Avem $F = \{1, 10l \pm 1, 1 \leq l \leq k, 10k + 9\}$, deci $\text{card } F = 2k + 2$ și prin urmare $P(F) = \frac{2k + 2}{n}$.

1.39. Dintr-o urnă ce conține n bile de diferite culori se extrag la întâmplare k bile cu întoarcere.

Să se determine probabilitatea ca cele k bile extrase să fie de culori diferite.

Soluție. Evenimentele elementare E sînt toate k -truplurile (u_1, \dots, u_k) ce pot fi formate cu cele n bile, așa că avem $\text{card } E = n^k$. Evenimentele favorabile F sînt k -truplurile (u_1, \dots, u_k) cu elemente distincte, deci $\text{card } F = A_n^k$.

$$\text{Rezultă că } P(F) = \frac{\text{card } F}{\text{card } E} = \frac{A_n^k}{n^k} = \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{k-1}{n}\right).$$

1.40. Un șofer a fost amendat pentru parcare interzisă de 12 ori numai în zilele de luni și marți. Din această cauză șoferul s-a decis să nu folosească mașina în aceste două zile. Este justificată această decizie?

Dacă cele 12 amenzi au fost date în zilele lucrătoare ale săptămîinii, rezultă că nu se dau amenzi duminică?

Soluție. În primul caz evenimentele elementare sînt date $E = \{\text{zilele săptămîinii}\}^{12}$, iar evenimentele favorabile sînt $F = \{\text{luni, marți}\}^{12}$. Rezultă că $P(F) = \frac{\text{card } F}{\text{card } E} = \left(\frac{2}{7}\right)^{12} = 0,0000003$, deci probabilitatea este extrem de mică ca să se dea amenzi numai luna și marța, așa că decizia șoferului nu este justificată.

La fel ca mai înainte rezultă că probabilitatea ca să se dea amenzi numai în zilele lucrătoare (deci nu duminică) este $\left(\frac{6}{7}\right)^{12} \approx \frac{1}{6}$.

Deoarece această probabilitate este intermediară (nici foarte mică, nici foarte mare) rezultă că nu se poate trage nici o concluzie.

1.41. Să se arate că este mai probabil ca aruncînd 4 zaruri să se obțină cel puțin o dată fața unu decît aruncînd de 24 de ori cîte două zaruri să se obțină cel puțin o dată o dublă de unul (paradoxul lui Mère).

Soluție. Pentru aruncarea a 4 zaruri evenimentele elementare sînt $E_1 = \{1, \dots, 6\}^4$ iar pentru aruncarea de 24 de ori a două zaruri evenimentele elementare sînt $E_2 = \{(1, 1), \dots, (6, 6)\}^{24}$ (deci avem de comparat două probabilități a două evenimente din cîmpuri de probabilitate diferite).

Fie A_1 evenimentul ca să se obțină cel puțin o dată fața unu. Avem egalitatea $\mathcal{C}A_1 = \{2, 3, \dots, 6\}^4$, deci

$$P(A_1) = 1 - \frac{\text{card } \mathcal{C}A_1}{\text{card } E_1} = 1 - \left(\frac{5}{6}\right)^4 = 0,517747.$$

Fie A_2 evenimentul ca să se obțină cel puțin o dublă de unu. Are loc egalitatea $\mathcal{C}A_2 = \{(1, 2), (1, 3), \dots, (6, 6)\}^{24}$, așa că

$$P(A_2) = 1 - P(\mathcal{C}A_2) = 1 - \frac{\text{card } \mathcal{C}A_2}{\text{card } E_2} = 1 - \left(\frac{35}{36}\right)^{24} = 0,491404.$$

Se observă că $P(A_1) > P(A_2)$.

1.42. Se imaginează un joc în care jucătorul X aruncă 6 zaruri și câștigă jocu dacă obține cel puțin o dată fața unu, iar jucătorul Y aruncă 12 zaruri și câștigă jocul dacă obține cel puțin de două ori fața unu.

Care dintre cei doi jucători are șanse mai mari de a câștiga jocul?

Soluție. Evenimentele elementare în cazul cînd se aruncă 6 zaruri sînt date de $E_1 = \{1, \dots, 6\}^6$ iar în cazul cînd se aruncă 12 zaruri sînt date de $E_2 = \{1, \dots, 6\}^{12}$.

Fie A_1 evenimentul ca X să obțină cel puțin o dată fața unu. La fel ca la problema precedentă se obține că $P(A_1) = 1 - \left(\frac{5}{6}\right)^6 = 0,66510$.

Fie A_2 evenimentul ca Y să obțină cel puțin de două ori fața unu. Avem $\mathcal{C}A_2 = B_1 \cup B_2$ unde B_1 este evenimentul ca Y să obțină exact o dată fața unu și B_2 este evenimentul Y să nu obțină niciodată

fața unu. De asemenea observăm că $B_1 = \bigcup_{i=1}^{12} C_i$ unde C_i este evenimentul ca anul să apară la al i -lea zar. Avem $B_2 = \{2, \dots, 6\}^{12}$, deci $P(B_2) = \left(\frac{5}{6}\right)^{12}$ și $C_i = \{2, \dots, 6\}^{i-1} \times \{1\} \times \{2, \dots, 6\}^{12-i}$,

$$P(C_i) = \left(\frac{5}{6}\right)^{11} \cdot \frac{1}{6}, \text{ așa că } P(B_1) = \sum_{i=1}^{12} P(C_i) = 12 \cdot \left(\frac{5}{6}\right)^{11} \cdot \frac{1}{6} = 2 \left(\frac{5}{6}\right)^{11}.$$

$$\text{În fine, } P(A_2) = 1 - P(\mathcal{C}A_2) = 1 - P(B_1) - P(B_2) = 1 - \left(\frac{5}{6}\right)^{12} - 2 \left(\frac{5}{6}\right)^{11} = 0,61867.$$

Deci că avem $P(A_1) > P(A_2)$, cu alte cuvinte X are șansa mai mare să câștige jocul.

1.43. Dintr-o urnă ce conține m bile roșii și $n - m$ bile albe se extrag la întâmplare l bile cu întoarcere.

Să se găsească probabilitatea ca bila i să fie roșie.

Soluție. Evenimentele elementare E sînt l - truplurile (u_1, \dots, u_l) ce se pot forma cu mulțimea bilor, deci $\text{card } E = n^l$.

Dacă A_i este evenimentul ca bila i să fie roșie atunci A_i este dat de l - truplurile (u_1, \dots, u_l) din E cu n_i bilă roșie. Există m posibilități de alegere ale lui u_i și n posibilități de alegere a lui u_j pentru $j \neq i$, deoarece u_j nu are nici o restricție.

Conform principiului de bază, avem $\text{card } A_i = mn^{l-1}$, așa că rezultă

$$P(A_i) = \frac{\text{card } A_i}{\text{card } E} = \frac{mn^{l-1}}{n^l} = \frac{m}{n}.$$

Se va vedea în problema 1.59 că același rezultat se obține și dacă extracția se face fără întoarcere.

1.44. Din mulțimea $\{1, \dots, n\}$, $n \geq 3$, se extrag la întâmplare două numere a și b .

Să se arate că este mai posibil ca $a^2 - b^2$ să se dividă cu 3 decât cu 2.

Soluție. Mulțimea evenimentelor elementare este $E = \{1, \dots, n\}^2$, deci $\text{card } E = n^2$. Fie $F_1 = \{(a, b) \in E; a^2 - b^2 \text{ este divizibil cu } 2\}$, $F_2 = \{(a, b) \in E; a^2 - b^2 \text{ este divizibil cu } 3\}$.

Avem

$$F_1 = \{(a, b) \in E; a, b \text{ sînt pare sau } a, b \text{ sînt impare}\}$$

$$F_2 = \{(a, b) \in E; a, b \text{ se divid cu } 3 \text{ sau } a, b \text{ nu se divid cu } 3\}$$

deci

$$P(F_1) = P(\{(a, b) \in E; a, b \text{ pare}\}) + P(\{(a, b) \in E; a, b \text{ impare}\}) =$$

$$= \frac{\text{card } \{(a, b) \in E; a, b \text{ pare}\}}{\text{card } E} + \frac{\text{card } \{(a, b) \in E; a, b \text{ impare}\}}{\text{card } E} =$$

$$= \frac{\left[\frac{n}{2}\right]^2}{n^2} + \frac{\left(n - \left[\frac{n}{2}\right]\right)^2}{n^2} = 1 - \frac{2}{n} \left[\frac{n}{2}\right] + \frac{2}{n^2} \left[\frac{n}{2}\right]^2$$

$$P(F_2) = P(\{(a, b) \in E; a, b \text{ se divid cu } 3\}) + P(\{(a, b) \in E; a, b \text{ nu se}$$

$$\text{divid cu } 3\}) = \frac{\left[\frac{n}{3}\right]^2}{n^2} + \frac{\left(n - \left[\frac{n}{3}\right]\right)^2}{n^2} = 1 - \frac{2}{n} \left[\frac{n}{3}\right] + \frac{2}{n^2} \left[\frac{n}{3}\right]^2$$

Arătăm că $P(F_1) < P(F_2)$ dacă $n \geq 3$. Avem

$$\begin{aligned} P(F_2) - P(F_1) &= \frac{2}{n} \left(\left\lfloor \frac{n}{2} \right\rfloor - \left\lfloor \frac{n}{3} \right\rfloor \right) + \frac{2}{n^2} \left(\left\lfloor \frac{n}{3} \right\rfloor^2 - \left\lfloor \frac{n}{2} \right\rfloor^2 \right) = \\ &= \frac{2}{n} \left(\left\lfloor \frac{n}{2} \right\rfloor - \left\lfloor \frac{n}{3} \right\rfloor \right) \left[1 - \frac{1}{n} \left(\left\lfloor \frac{n}{3} \right\rfloor + \left\lfloor \frac{n}{2} \right\rfloor \right) \right] > 0. \end{aligned}$$

1.45. Perechile de numere (x, y) , unde x, y iau valorile întregi între $-n$, și n , sînt scrise pe niște cartonașe (pe fiecare cartonaș este scris numai o pereche). Aceste cartonașe se introduc într-o urnă și se extrag apoi la întîmplare trei cartonașe fără întoarcere.

Considerînd cele trei perechi scrise pe cele trei cartonașe ca fiind coordonatele punctelor M_1, M_2, M_3 , să se calculeze probabilitatea ca M_1 și M_2 să fie simetrice în raport cu M_3 .

Soluție. Mulțimea evenimentelor elementare E este dată de 3-truplurile ordonate cu componente distincte formate cu cele $(2n+1)^2$ cartonașe. Conform punctului f_2 din rezumatul teoretic, avem egalitatea

$$\text{card } E = A_{(2n+1)}^3 = 4n(n+1)\{[4n(n+1)]^2 - 1\}.$$

Pentru $|d_1| \leq n, |d_2| \leq n$, fie $F_{d_1, d_2} = \{(x_1, y_1), (x_2, y_2), (x_3, y_3)\}; x_3 - x_1 = x_2 - x_3 = d_1, y_3 - y_1 = y_2 - y_3 = d_2\}$.

Atunci evenimentele favorabile sînt date de $F = \bigcup_{\substack{d_1, d_2 = -n \\ d_1^2 + d_2^2 \neq 0}}^n F_{d_1, d_2}$,

Aplicația $f: F_{d_1, d_2} \rightarrow \{(x, y); |x| \leq n - |d_1|, |y| \leq n - |d_2|\}$ definită prin $f((x_1, y_1), (x_2, y_2), (x_3, y_3)) = (x_3, y_3)$ este o bijecție, deci $\text{card } F_{d_1, d_2} = \text{card } \{(x, y); |x| \leq n - |d_1|, |y| \leq n - |d_2|\} = \text{card } \{x; |x| \leq n - |d_1|\} \text{ card } \{y; |y| \leq n - |d_2|\} = (2n + 1 - |d_1|)(2n + 1 - |d_2|)$

Rezultă

$$\begin{aligned} \text{card } F &= \sum_{\substack{d_1, d_2 = -n \\ d_1^2 + d_2^2 \neq 0}}^n \text{card } F_{d_1, d_2} = \sum_{\substack{d_1, d_2 = -n \\ d_1^2 + d_2^2 \neq 0}}^n (2n + 1 - |d_1|)(2n + 1 - |d_2|) = \\ &= (2n + 1) 2 \sum_{d_2=1}^n (2n + 1 - 2d_2) + 2 \sum_{d_1=1}^n (2n + 1 - \\ &- 2d_1) \left(2n + 1 + 2 \sum_{d_2=1}^n (2n + 1 - 2d_2) \right) = 2(2n + 1)(n(2n + 1) - \\ &- n(n + 1)) + 2(n(2n + 1) - n(n + 1))(2n + 1 + 2(n(2n + 1) - \\ &- n(n + 1))) = 4n^2(n + 1)^2 \end{aligned}$$

deci

$$P(F) = \frac{\text{card } F}{\text{card } E} = \frac{4n^2(n+1)^2}{4n(n+1) \cdot \{[4n(n+1)]^2 - 1\}} = \frac{n(n+1)}{[4n(n+1)]^2 - 1}.$$

1.46. Un grup de $2n$ băieți și $2n$ fete este împărțit la întâmplare în două grupuri egale.

Să se găsească probabilitatea p ca fiecare grup să aibă același număr de băieți și fete și apoi să se estimeze p folosind formula lui Stirling.

Soluție. Fie A mulțimea băieților și B cea a fetelor. Atunci evenimentele elementare sînt $E = \{L \subset A \cup B; \text{card } L = 2n\}$, deci conform punctului f_3 din rezumatul teoretic avem $\text{card } E = C_{4n}^{2n}$.

Evenimentele favorabile sînt $F = \{L \subset A \cup B; \text{card } L = 2n, \text{card}(L \cap A) = \text{card}(L \cap B) = n\}$.

Aplicația $L \rightarrow (L \cap A, L \cap B)$ este o bijecție între F și $\{L_1 \subset A; \text{card } L_1 = n\} \times \{L_2 \subset B; \text{card } L_2 = n\}$, deci $\text{card } F = C_{2n}^n \cdot C_{2n}^n = (C_{2n}^n)^2$.

Rezultă

$$p = P(F) = \frac{\text{card } F}{\text{card } E} = \frac{(C_{2n}^n)^2}{C_{4n}^{2n}}.$$

Formula lui Stirling afirmă că $n! \sim \sqrt{2\pi n}^{n+\frac{1}{2}} e^{-n} (a_n \sim b_n$ înseamnă că $\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = 1)$. Avem

$$\begin{aligned} p &= \left[\frac{(2n)!}{(n!)^2} \right]^2 \cdot \frac{(2n)!^2}{4n!} = \frac{(2n)!^4}{4n!(n!)^4} \sim \\ &\sim \frac{\left[\sqrt{2\pi(2n)}^{2n+\frac{1}{2}} e^{-8n} \right]^4}{\sqrt{2\pi(4n)}^{4n+\frac{1}{2}} e^{-4n} \left[\sqrt{2\pi n}^{n+\frac{1}{2}} e^{-n} \right]^4 e^{4n}} = \sqrt{\frac{2}{\pi}}. \end{aligned}$$

1.47. (*Schema bilei neîntoarse*). Dintr-o urnă ce conține m bile roșii și $n - m$ bile albe se extrag la întâmplare k bile fără întoarcere (extracție se face simultan sau secvențial).

Să se arate că probabilitatea ca cele k bile extrase să conțină exact r bile roșii este $\frac{C_m^r C_{n-m}^{k-r}}{C_n^k}$.

Soluție. Presupunem întâi că bilele se extrag simultan. Fie R mulțimea bilelor roșii și A cea a bilelor albe.

Evenimentele elementare sînt $E = \{L \subset R \cup A; \text{card } L = k\}$, deci $\text{card } E = C_n^k$. Evenimentele favorabile sînt $F = \{L \subset R \cup A, \text{card } L = k, \text{card } (L \cap R) = r\}$.

Aplicația $L \rightarrow (L \cap A, L \cap R)$ este o bijecție între F și $\{L_1 \subset A; \text{card } L_1 = k - r\} \times \{L_2 \subset R; \text{card } L_2 = r\}$, deci $\text{card } F = C_{n-m}^{k-r} C_m^r$ și prin urmare $P(F) = \frac{\text{card } F}{\text{card } E} = \frac{C_m^r C_{n-m}^{k-r}}{C_n^k}$.

Presupunem în continuare că bilele se extrag secvențial. În acest caz evenimentele elementare sînt $E = \{u = (u_1, \dots, u_k); u_i \in R \cup A, u$ are elemente distincte}, deci conform punctului f_2 din rezumatul teoretic avem $\text{card } E = A_n^k$.

Evenimentele favorabile sînt $F = \{u = (u_1, \dots, u_k) \in E; \text{card } \{i; u_i \in R\} = r\}$.

Dacă pentru $L = \{i_1, \dots, i_r\} \subset \{1, \dots, k\}$ cu $i_1 < \dots < i_r$, definim

$$F_L = \{u = (u_1, \dots, u_k) \in E; u_{ij} \in R, 1 \leq j \leq r\}$$

atunci

$$F = \bigcup_L F_L.$$

Mulțimile F_L și $\{v = (v_{i_1}, \dots, v_{i_r}); v_{i_j} \in R, v$ are elemente distincte} \times $\{w = (w_{j_1}, \dots, w_{j_{k-r}}); w_{i_j} \in A, w$ are elemente distincte}, unde $j_1 < \dots < j_{k-r}$ sînt indicii din $\{1, \dots, n\} \setminus L$, sînt în corespondență bijectivă prin aplicația

$$u = (u_1, \dots, u_k) \rightarrow ((u_{i_1}, \dots, u_{i_r}), (u_{j_1}, \dots, u_{j_{k-r}}))$$

Rezultă că avem $\text{card } F = \sum_L \text{card } F_L = \sum_j A_m^r A_{n-m}^{k-r} = C_k^r A_m^r A_{n-m}^{k-r}$

și prin urmare $P(F) = \frac{C_k^r A_m^r A_{n-m}^{k-r}}{A_n^k}$.

Mai departe avem

$$\frac{C_m^r A_m^r A_{n-m}^{k-r}}{A_n^k} = \frac{k!}{r!(k-r)!} \cdot \frac{A_m^r A_{n-m}^{k-r}}{A_n^k} = \frac{A_m^r}{r!} \cdot \frac{A_{n-m}^{k-r}}{(k-r)!} \cdot \frac{k!}{A_n^k} = \frac{C_m^r C_{n-m}^{k-r}}{C_n^k}$$

1.48. (Schema bilei întoarse). Dintr-o urnă ce conține m bile roșii și $n - m$ bile albe se extrag la întâmplare k bile cu întoarcere.

Să se arate că probabilitatea ca cele k bile să conțină exact r bile roșii este $\frac{C_m^r m^r (n-m)^{k-r}}{n^k}$.

Soluție. Fie R mulțimea bilelor roșii și A cea a bilelor albe. Evenimentele elementare sînt $E = (R \cup A)^k$, deci $\text{card } E = n^k$.

Observăm că E se mai scrie sub forma

$$E = \bigcup_{G_i = A \text{ sau } R} G_1 \times \dots \times G_k. \text{ Evenimentele favorabile sînt}$$

$$F = \bigcup_{\substack{G_i = A \text{ sau } R \\ \text{card } \{i; G_i = R\} = r}} G_1 \times \dots \times G_k,$$

$$\text{deci } \text{card } F = \sum_{\substack{G_i = A \text{ sau } R \\ \text{card } \{i; G_i = R\} = r}} \text{card } (G_1 \times \dots \times G_k) =$$

$$= m^r (n-m)^{k-r} \text{card } \{(G_1 \dots G_k); G_i = A \text{ sau } R, \text{card } \{i; G_i = R\} = r\},$$

Aplicația $(G_1, \dots, G_k) \rightarrow \{i; G_i = R\}$ este o bijecție între $\{(G_1, \dots, G_k); G_i = A \text{ sau } R, \text{card } \{i; G_i = R\} = r\}$ și $\{L \subset \{1, \dots, k\}; \text{card } L = r\}$, fapt ce implică egalitatea următoare: $\text{card } \{(G_1, \dots, G_k); G_i = A \text{ sau } R, \text{card } \{i; G_i = R\} = r\} = C_k^r$. Prin urmare $\text{card } F =$

$$= C_k^r m^r (n-m)^{k-r}, \text{ de unde } P(F) = \frac{\text{card } F}{\text{card } E} = \frac{C_k^r m^r (n-m)^{k-r}}{n^k}.$$

Altă soluție. Conform problemei 1.20 din paragraful precedent probabilitatea clasică pe E este produsul probabilităților clasice pe $(R \cup A)_i$; $i = 1, \dots, k$, unde $(R \cup A)_i = R \cup A$.

Dacă notăm cu F_i evenimentul că la extragerea i am obținut o bilă roșie, atunci $P(F_i) = \frac{m}{n}$ și din problema 1.20 rezultă că evenimentele $\bar{F}_i = (R \cup A)_1 \times \dots \times (R \cup A)_{i-1} \times F_i \times (R \cup A)_{i+1} \cup \dots \cup (R \cup A)_k$ sînt independente.

Evenimentul F ca în cele k bile extrase să avem r bile roșii înseamnă realizarea a r dintre evenimentele \bar{F}_i , deci conform problemei 1.26 rezultă că

$$P(F) = C_k^r \left(\frac{m}{n}\right)^r \left(1 - \frac{m}{n}\right)^{k-r} = C_k^r \frac{m^r (n-m)^{k-r}}{n^k}$$

1.49. Pe niște bilete sînt scrise numerele de la 000000 la 999999. Care este probabilitatea ca extrăgînd la întîmplare un bilet acesta să aibă suma primelor trei cifre egală cu suma ultimelor trei cifre.

Soluție. Evenimentele elementare sînt $E = \{(a_1, \dots, a_6); 0 \leq a_i \leq 9\}$, deci $\text{card } E = 10^6$.

Evenimentele favorabile sînt $F = \{(a_1, \dots, a_6); 0 \leq a_i \leq 9, a_1 + a_2 + a_3 = a_4 + a_5 + a_6\} = \bigcup_{l=0}^{27} \{(a_1, \dots, a_6); a \leq a_i \leq 9;$

$$a_1 + a_2 + a_3 = a_4 + a_5 + a_6 = l\} = \bigcup_{l=0}^{27} F_l.$$

Cum aplicația $f: F_l \rightarrow \{(a_1, a_2, a_3); 0 \leq a_i \leq 9, a_1 + a_2 + a_3 = l\}^2$ definită prin $f((a_1, \dots, a_6)) = (a_1, a_2, a_3), (a_4, a_5, a_6)$ este o bijectie, rezultă

$$\text{card } F_l \stackrel{\text{def}}{=} b_l^2 = [\text{card}\{(a, b, c); 0 \leq a, b, c \leq 9, a + b + c = l\}]^2.$$

Atunci

$$P(F) = \frac{\text{card } F}{\text{card } E} = \frac{\sum_{l=0}^{27} \text{card } F_l}{10^6} = 10^{-6} \left(\sum_{l=0}^{27} b_l^2 \right)$$

Rămîn de calculat numerele b_0, \dots, b_{27} .

$$\begin{aligned} \text{Calculul lui } b_0, \dots, b_9. \quad b_l &= \text{card} \bigcup_{a=0}^l \{(b, c); 0 \leq b, c \leq 9, b + c = \\ &= l - a\} = \sum_{a=0}^l (l - a + 1) = \frac{(l+1)(l+2)}{2} \end{aligned}$$

$$\begin{aligned} \text{Calculul } b_{10}, \dots, b_{18}. \quad \text{Avem } b_{l+9} &= \sum_{a=0}^9 \text{card} \{(b, c); 0 \leq b, c \leq 9, \\ b + c = 9 + l - a\} &= \sum_{a=0}^{l-1} \text{card} \{(b, c); 0 \leq b, c \leq 9, b + c = 9 + l - a\} + \\ &+ \sum_{a=l}^9 \text{card} \{(b, c); 0 \leq b, c \leq 9, b + c = 9 + l - a\}. \end{aligned}$$

Deoarece pentru $a \leq l - 1$ avem $b + c = 9 + l - a$ dacă și numai dacă: $b = 9, c = l - a; b = 8, c = l - a + 1; \dots; b = l - a, c = 9$ și pentru $a \geq l$ avem $b + c = 9 + l - a$ dacă și numai dacă: $b = 0, c = 9 + l - a; b = 1, c = 9 + l - a - 1; \dots; b = 9 + l - a, c = 0$ rezultă că avem $\text{card} \{(b, c); 0 \leq b, c \leq 9, b + c = 9 + l - a\} = 10 - (l - a)$ dacă $a \leq l - 1$ și $\text{card} \{(b, c); 0 \leq b, c \leq 9, b + c = 9 + l - a\} = 10 + l - a$ dacă $a \geq l$.

$$\begin{aligned} \text{Prin urmare } b_{l+9} &= \sum_{a=0}^{l-1} (10 - l + a) + \sum_{a=l}^9 (10 + l - a) = \\ &= \frac{9 + 10 - l}{2} \cdot l + \frac{10 + l + 1}{2} (10 - l) = 9l - l^2 + 55. \end{aligned}$$

Calculul lui b_{19}, \dots, b_{27} . Avem $b_{18+l} = \sum_{a=l}^9 \text{card} \{(b, c); 0 \leq b, c \leq 9, b + c = 18 + l - a\} = \sum_{a=l}^9 (a - l + 1) = \frac{(10 - l)(11 - l)}{2}$.

1.50. Dacă n persoane, printre care se află A și B , se așază la întâmplare într-un rînd, care este probabilitatea ca să existe exact r persoane între A și B ?

Dacă cele n persoane se așază într-un cerc, să se arate că probabilitatea precedentă este independentă de r și deci egală cu $\frac{1}{(n-1)}$ (în așezare în cerc se consideră numai arcurile AB în direcție pozitivă).

Soluție. Evenimentele elementare E sînt date de n -truplurile ordonate distincte cu componente distincte formate cu cele n persoane deci $\text{card } E = n!$. A și B pot să stea pe locurile $(i, i + r + 1)$, $1 \leq i \leq n - r - 1$, deci A și B se pot așeza în $2(n - r - 1)$ moduri (A și B au roluri simetrice).

Evenimentele favorabile F sînt date de n -truplurile (u_1, \dots, u_n) distincte cu elemente distincte cu proprietatea că $u_i = A, u_{i+r+1} = B$, sau $u_i = B, u_{i+r+1} = A$ pentru $1 \leq i \leq n - r - 1$, deci

$$F = \bigcup_{\substack{u_i = A, u_{i+r+1} = B \\ \text{sau} \\ u_i = B, u_{i+r+1} = A \\ 1 \leq i \leq n - r - 1}} \{u = (u_1, \dots, u_n); u_i \text{ din mulțimea celor } n \text{ persoane, } u \text{ are componente distincte}\}.$$

Rezultă că avem $\text{card } F = \sum_{\substack{u_i=A, u_{i+r+1}=B \\ \text{sau} \\ u_i=B, u_{i+r+1}=A \\ 1 \leq i \leq n-r-1}} \text{card} \{u = u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_{i+r}, \dots, u_{i+r+2}, \dots, u_n\}; u \text{ are componente distincte} =$

$$= \sum_{\substack{u_i=A, u_{i+r+1}=B \\ \text{sau} \\ u_i=B, u_{i+r+1}=A \\ 1 \leq i \leq n-r-1}} (n-2)! = 2(n-r-1)(n-2)!$$

Prin urmare $P(F) = \frac{\text{card } F}{\text{card } E} = \frac{2(n-r-1)(n-2)!}{n!} = \frac{2(n-r-1)}{n(n-1)}$.

Presupunem acum că persoanele se așază în cerc.

Fie $\{\bar{1}, \dots, \bar{n}\}$ mulțimea persoanelor. În acest caz evenimentele elementare sînt $E = \{f: \{1, \dots, n\} \rightarrow \{\bar{1}, \dots, \bar{n}\}; f \text{ bijectivă}\}$ deci $\text{card } E = n!$ Pentru $1 \leq i \leq n$ fie $A_i = \{f \in E; f(i) = A, f(i+r+1) = B\}$.

Avem $\text{card } A_i = (n-2)!$ și evenimentele sînt date de $F = \bigcup_{i=1}^n A_i$,

deci $\text{card } F = \sum_{i=1}^n \text{card } A_i = n(n-2)!$, așa că $P(F) = \frac{\text{card } F}{\text{card } E} = \frac{1}{n-1}$.

1.51. Din submulțimile mulțimii $\{1, \dots, n\}$ se extrag la întâmplare r submulțimi A_1, \dots, A_r cu întoarcere.

Să se afle probabilitatea ca cele r submulțimi să fie disjuncte două câte două.

Soluție. Evenimentele elementare sînt $E_r = (\mathcal{P}(\{1, \dots, n\}))^r$, deci $\text{card } E_r = 2^{nr}$. Evenimentele favorabile sînt $F_r = \{(A_1, \dots, A_r) \in E_r; A_i \text{ sînt disjuncte două câte două}\}$. Pentru a calcula $\text{card } F_r$, aplicăm inducția după r .

Fie $r = 2$. Atunci $F_2 = \bigcup_{k=0}^n \bar{F}^k$ unde $\bar{F}^k = \{(A_1, A_2); \text{card } A_1 = k, A_1 \cap A_2 = \emptyset\} = \bigcup_{\substack{A_1 \subset \{1, \dots, n\} \\ \text{card } A_1 = k}} \{(A_1, A_2); A_1 \cap A_2 = \emptyset\}$, deci

$$\text{card } F_2 = \sum_{k=0}^n \text{card } \bar{F}^k = \sum_{k=0}^n \sum_{\substack{A_1 \subset \{1, \dots, n\} \\ \text{card } A_1 = k}} \text{card} \{(A_1, A_2); A_1 \cap A_2 = \emptyset\} =$$

$$= \sum_{k=0}^n \sum_{\substack{A_1 \subset \{1, \dots, n\} \\ \text{card } A_1 = k}} 2^{n-k} = \sum_{k=0}^n C_n^k 2^{n-k} = 3^n \text{ (am folosit faptul că pentru$$

A_1 fixat cu card $A_1 = k$ mulțimile $\{(A_1, A_2); A_1 \cap A_2 = \emptyset\}$ și $\mathfrak{A}(A_1)$ au același cardinal; se vede că aplicația $(A_1, A_2) \rightarrow A_2$ este o bijecție între aceste mulțimi).

$$\text{Rezultă că } P(F_2) = \frac{\text{card } F_2}{\text{card } E_2} = \frac{3^n}{2^{2n}} = \left(\frac{3}{4}\right)^n.$$

Presupunem acum că avem card $F_{r-1} = r^n$ și să arătăm că card $F_r = (r+1)^n$. Avem $F_r = \bigcup_{k=0}^n \tilde{F}^k$ unde $\tilde{F}^k = \{(A_1, \dots, A_r); \text{card } A_1 = k, A_i \text{ sînt disjuncte două cîte două}\} = \bigcup_{\substack{A_1 \subset \{1, \dots, n\} \\ \text{card } A_1 = k}} G_{A_1}$, unde

$G_{A_1} = \{(A_1, \dots, A_r); A_i \text{ sînt disjuncte două cîte două}\}$. Dar prin aplicația bijectivă $((A_1, \dots, A_r) \rightarrow (A_2, \dots, A_r))$ se observă că G_{A_1} are același cardinal cu F_{r-1} construit plecînd de la mulțimea \mathfrak{A}_{A_1} , deci conform ipotezei inducției avem card $G_{A_1} = r^{n - \text{card } A_1}$. Rezultă

$$\begin{aligned} \text{card } F_r &= \sum_{k=0}^n \sum_{\substack{A_1 \subset \{1, \dots, n\} \\ \text{card } A_1 = k}} \text{card } G_{A_1} = \sum_{k=0}^n \sum_{\substack{A_1 \subset \{1, \dots, n\} \\ \text{card } A_1 = k}} r^{n-k} = \\ &= \sum_{k=0}^n C_n^k r^{n-k} = (r+1)^n. \end{aligned}$$

$$\text{Prin urmare } P(F_r) = \frac{\text{card } F_r}{\text{card } E_r} = \frac{(r+1)^n}{2^{nr}} = \left(\frac{r+1}{2r}\right)^n.$$

1.52. Numerele $1, \dots, n$ sînt scrise pe niște cartonașe. Se introduc cartonașele într-o urnă și se extrag la întîmplare l cartonașe (cele l cartonașe se extrag simultan). Să se afle probabilitatea ca în cele l cartonașe să nu fie două numere consecutive.

Soluție. Evenimentele elementare sînt $E = \{L \subset \{1, \dots, n\}; \text{card } L = l\}$, deci avem card $L = C_n^l$.

Evenimentele favorabile sînt $F = \{L \in E; L \text{ nu conține două numere consecutive}\}$. Definim aplicația $f: F \rightarrow \{L \subset \{0, 1, \dots, n-l\}; \text{card } L = l\}$ prin $f(L) = \{i_1 - 1, \dots, i_l - l\}$, dacă $L = \{i_1, \dots, i_l\}$ $i_1 < \dots < i_l$.

Este ușor de văzut că f este o bijecție, deci card $F = C_{n-l+1}^l$.

$$\text{Prin urmare } P(F) = \frac{\text{card } F}{\text{card } E} = \frac{C_{n-l+1}^l}{C_n^l}.$$

1.53. (Problema mesei rotunde). În jurul unei mese rotunde cu $2n$ scaune se așază în mod aleator n perechi soț-soție astfel încît să nu existe două persoane de același sex pe scaune alăturate.

□

Se cere probabilitatea ca nici un bărbat să nu aibă ca vecină soția sa (nu se disting două configurații ce diferă printr-o rotație).

Soluție. Fie $\{1, \dots, n\}$ mulțimea bărbaților și $\{\bar{1}, \dots, \bar{n}\}$ mulțimea femeilor.

Evenimentele elementare (așezările posibile) sînt $E = \{f: \{1, \dots, n\} \rightarrow \{\bar{1}, \dots, \bar{n}\}; f \text{ bijectivă}\}$.

Într-adevăr la o așezare (dacă notăm cu $1, \dots, n$ locurile pe care stau bărbații) îi facem să corespundă aplicația bijectivă $f: \{1, \dots, n\} \rightarrow \{\bar{1}, \dots, \bar{n}\}$ definită prin $f(i) = \text{soția ce stă în dreapta lui } i$. Invers dacă $f: \{1, \dots, n\} \rightarrow \{\bar{1}, \dots, \bar{n}\}$ este o funcție bijectivă îi facem să corespundă așezarea definită astfel: în dreapta lui i pun $f(i)$.

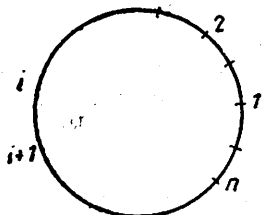


Fig. 1.

Fie $A_{2i-1} = \{f \in E; f(i) = \bar{i}\}$, $1 \leq i \leq n$, $A_{2i} = \{f \in E; f(i) = \bar{i+1}\}$, $1 \leq i \leq n-1$, $A_{2n} = \{f \in E; f(n) = \bar{1}\}$.

Evenimentele favorabile sînt $F = \complement(A_1 \cup \dots \cup A_{2n})$, deci $P(F) = 1 - P(A_1 \cup \dots$

$$\dots \cup A_{2n}) = 1 - \sum_{L \subset \{1, \dots, 2n\}} (-1)^{\text{card } L+1} P\left(\bigcap_{i \in L} A_i\right) = 1 - \sum_{L \subset \{1, \dots, 2n\}} \sum_{\text{card } L=i} (-1)^{i+1} P\left(\bigcap_{i \in L} A_i\right).$$

Deci pentru $L \subset \{1, \dots, 2n\}$ $\text{card } L = l$, avem de calculat $P\left(\bigcap_{i \in L} A_i\right)$

Arătăm pentru aceasta că:

1) Dacă $l \geq n+1$, atunci L conține doi întregi consecutivi din $\{1, \dots, 2n, 1\}$ (aici 1 este considerat consecutiv lui $2n$).

2) Dacă L conține doi întregi consecutivi din $\{1, \dots, 2n, 1\}$, atunci $\bigcap_{i \in L} A_i = \emptyset$.

1) Presupunem deci că $L \supset \{i_1, \dots, i_{n+1}\}$ cu $i_1 < \dots < i_{n+1}$ și $i_{j+1} - i_j > 1$. Rezultă că între i_j și i_{j+1} se află cel puțin un număr din $\{1, \dots, 2n\}$, deci în total cel puțin n numere din $\{1, \dots, 2n\}$ se află intercalate între numerele i_1, \dots, i_{n+1} , fapt ce ar implica $\{1, \dots, 2n\}$ ar avea mai mult de $2n$ elemente (i_1, \dots, i_{n+1} și cele n intervale), ceea ce nu se poate.

2) Presupunem că L conține doi întregi consecutivi și prin absurd fie $f \in \bigcap_{i \in L} A_i$. Fie $(2i, 2i + 1)$, $(2n, 1)$ una din perechile de numere consecutive pe care le conține L .

Dacă $f \in A_{2i} \cap A_{2i+1}$ atunci $f \in A_{2i}$ implică $f(i) = \overline{i + 1}$ și $f \in A_{2i+1}$ implică $f(i + 1) = \overline{i + 1}$, fapt ce contrazice injectivitatea lui f .

Dacă $f \in A_{2i-1} \cap A_{2i}$, atunci $f(i) = \overline{i + 1}$, $f(i) = \overline{i}$, ceea ce ar arăta că f nu este univoc definită.

Dacă $f \in A_{2n} \cap A_1$, atunci $f(n) = \overline{1}$, $f(1) = \overline{1}$, fapt ce contrazice injectivitatea lui f .

Rezultă deci că

$$P(F) = 1 - \sum_{l=1}^{2n} \sum_{\substack{L \subset \{1, \dots, 2n\} \\ \text{card } L = l \text{ și } L \\ \text{nu conține doi} \\ \text{întregi consecu-} \\ \text{tivi din} \\ \{1, \dots, 2n, 1\}}} (-1)^{l+1} P\left(\bigcap_{i \in L} A_i\right)$$

Dar $P\left(\bigcap_{i \in L} A_i\right) = \frac{(n - \text{card } L)!}{n!}$ așa că

$$P(F) = 1 - \sum_{l=1}^n (-1)^{l+1} \frac{(n-l)!}{n!} \text{card } \{L \subset \{1, \dots, 2n\}, \text{card } L = l \text{ și } L \text{ nu conține doi întregi consecutivi din } 1, \dots, 2n, 1\}$$

Fie $\mathcal{A} = \{L \subset \{1, \dots, 2n\}; \text{card } L = l, L \text{ nu conține doi întregi consecutivi din } 1, \dots, 2n, 1\}$

$\mathcal{A}_1 = \{L_1 \subset \{2, \dots, 2n-2\}; \text{card } L_1 = l-1, L_1 \text{ nu conține doi întregi consecutivi din } 2, \dots, 2n-2\}$

$\mathcal{A}_2 = \{L_2 \subset \{1, \dots, 2n-1\}; \text{card } L_2 = l \text{ și } L_2 \text{ nu conține doi întregi consecutivi din } 1, \dots, 2n-1\}$.

Avem $\text{card } \mathcal{A}_1 = C_{2n-1}^{l-1}$, $\text{card } \mathcal{A}_2 = C_{2n-1}^l$ și este vizibil că $\mathcal{A}_1 \cap \mathcal{A}_2 = \emptyset$ Aplicația $\varphi: \mathcal{A} \rightarrow \mathcal{A}_1 \cup \mathcal{A}_2$ definită prin

$$\varphi(L) = \begin{cases} L \setminus \{1, 2n-1, 2n\} & \text{dacă } 2n \in L \\ L & \text{dacă } 2n \notin L \end{cases}$$

este o bijecție, așa că avem $\text{card } \mathcal{A} = \text{card } \mathcal{A}_1 + \text{card } \mathcal{A}_2 = \frac{2n}{2n-1} C_{2n-1}^l$.

În fine obținem că

$$P(F) = 1 - \sum_{l=1}^n (-1)^{l+1} \frac{(n-l)!}{n!} \frac{2n}{2n-l} C_{2n-1}^l$$

1.54. Fie f, g funcții reale definite pe mulțimea părților mulțimii finite E . Să se arate că următoarele două afirmații sînt echivalente :

a) pentru orice $A \subset E$ avem $f(A) = \sum_{B \supset A} g(B)$;

b) pentru orice $A \subset E$ avem $g(A) = \sum_{B \supset A} (-1)^{\text{card}(B \setminus A)} f(B)$
(principiul includerii și al excluderii).

Aplicație. La o sărbătoare participă n copii și m adulți. Fiecare adult atribuie la întîmplare copiilor cîte r cadouri (un adult nu dă mai mult de un cadou la un copil și deci $r \leq n$).

Care este probabilitatea ca fiecare copil să primească cel puțin un cadou?

Soluție. Observăm întîi că dacă $A \subset C \subset E$ și $\text{card}(C \setminus A) = n > 0$ atunci există C_n^k mulțimi B așa încît $A \subset B \subset C$ și $\text{card}(B \setminus A) = k$.

Deci

$$\sum_{A \subset B \subset C} (-1)^{\text{card}(B \setminus A)} = \sum_{k=0}^n C_n^k (-1)^k = (1 - 1)^n = 0.$$

$$\sum_{A \subset B \subset C} (-1)^{\text{card}(C \setminus B)} = \sum_{k=0}^n C_n^k (-1)^{n-k} = (-1 + 1)^n = 0.$$

Rezultă că $f(B) = \sum_{B \subset C} g(C)$ pentru orice B implică

$$\begin{aligned} \sum_{A \subset B} (-1)^{\text{card}(B \setminus A)} f(B) &= \sum_{A \subset B} (-1)^{\text{card}(B \setminus A)} \sum_{B \subset C} g(C) = \\ &= \sum_{A \subset C} g(C) \sum_{A \subset B \subset C} (-1)^{\text{card}(B \setminus A)} = g(A). \end{aligned}$$

Invers dacă $g(B) = \sum_{B \subset C} (-1)^{\text{card}(C \setminus B)} f(C)$ pentru orice B atunci

$$\begin{aligned} \sum_{A \subset B} g(B) &= \sum_{A \subset B} \sum_{B \subset C} (-1)^{\text{card}(C \setminus B)} f(C) = \\ &= \sum_{A \subset C} f(C) \sum_{A \subset B \subset C} (-1)^{\text{card}(C \setminus B)} = f(A) \end{aligned}$$

Aplicație. Dacă E este mulțimea copiilor atunci pentru $A \subset E$ definim $f(A)$ ca fiind probabilitatea ca cel puțin copiii din mulțimea A să nu fi primit cadouri și $g(A)$ ca fiind probabilitatea ca numai copiii din mulțimea A să nu fi primit cadouri.

Evenimentele elementare sînt date de $\bar{E} = \{L \subset E; \text{card } L = r\}^m$, deci $\text{card } \bar{E} = (C_n^r)^m$.

Fie F evenimentul ca cel puțin copii din A să nu primească cadouri. Avem $F = \emptyset$ dacă $\text{card } A > n - r$ și $F = \{L \subset \mathcal{C}A; \text{card } L = r\}^m$ dacă $\text{card } A \leq n - r$, deci

$$P(F) = 0 \text{ dacă } \text{card } A > n - r \text{ și } f(A) = \frac{\text{card } F}{\text{card } E} = \left[\frac{C_{n-r}^r}{C_n^r} \right]^m$$

dacă $\text{card } A \leq n - r$. Deoarece pentru $B \neq B'$, $g(B)$ și $g(B')$ sînt probabilități de evenimente disjuncte, rezultă că $f(A) = \sum_{A \subset B} g(B)$.

Avem de calculat $g(\emptyset)$ care, din principiul includerii și al excluderii, este dat de $g(\emptyset) = \sum_B (-1)^{\text{card } B} f(B) = \frac{[(n-r)!]^m}{[n!]^{m-1}} \sum_{k=0}^{n-r} \frac{[(n-k)!]^{m-1}}{[(n-r-k)!]^m} \frac{(-1)^k}{k!}$

Lășăm cititorului să dea o soluție directă aplicației.

1.55. (Problema concordanțelor). n perechi soț-soție participă la un dans unde partenerii sînt aleși prin tragere la sorți.

Care este probabilitatea ca nici un bărbat să nu danseze cu soția sa?

Soluție. Fie $\{i, \bar{i}\}_{1 \leq i \leq n}$ cele n perechi. Evenimentele elementare (perechile de dans posibile) sînt date de $E = \{f: \{1, \dots, n\} \rightarrow \{\bar{1}, \dots, \bar{n}\}; f \text{ bijectivă}\}$, deci $\text{card } E = n!$. Fie $A_i = \{f \in E; f(i) = \bar{i}\}$. Atunci evenimentul a cărui probabilitate se cere este $F = \mathcal{C}(A_1 \cup \dots \cup A_n)$, deci

$$P(F) = 1 - P\left(\bigcup_{i=1}^n A_i\right) = 1 - \sum_{L \subset \{1, \dots, n\}} (-1)^{\text{card } L+1} P\left(\bigcap_{i \in L} A_i\right).$$

Dar $\bigcap_{i \in L} A_i$ este în bijecție cu $\{f: \{1, \dots, n\} \setminus L \rightarrow \{\bar{1}, \dots, \bar{n}\} \setminus \bar{L};$

$f \text{ bijecție}\}$ prin aplicația $f \rightarrow f|_{\{1, \dots, n\} \setminus L}$, deci $\text{card}\left(\bigcap_{i \in L} A_i\right) =$

$$= (n - \text{card } L)! \text{ și prin urmare } P\left(\bigcap_{i \in L} A_i\right) = \frac{(n - \text{card } L)!}{n!}$$

Rezultă că

$$\begin{aligned} P(F) &= 1 - \sum_{L \subset \{1, \dots, n\}} (-1)^{\text{card } L+1} \frac{(n - \text{card } L)!}{n!} = \\ &= 1 - \sum_{l=1}^n \sum_{\substack{L \subset \{1, \dots, n\} \\ \text{card } L=l}} (-1)^{l+1} \frac{(n-l)!}{n!} = \\ &= 1 - \sum_{l=1}^n (-1)^{l+1} C_n^l \frac{(n-l)!}{n!} = \sum_{l=0}^n (-1)^l C_n^l \frac{(n-l)!}{n!} = \sum_{l=0}^n \frac{(-1)^l}{l!}. \end{aligned}$$

Este amuzant de văzut că $P(F) \rightarrow e^{-1} = 0,36$ când $n \rightarrow \infty$. Putem aplica și principiul includerii și al excluderii pentru rezolvarea problemei. Pentru $A \subset \{1, \dots, n\}$ definim $f(A)$ ca fiind probabilitatea ca cel puțin dansatorii din A să danseze cu soțiile lor și $g(A)$ ca fiind probabilitatea ca numai dansatorii din A să danseze cu soțiile lor.

Conform celor precedente, avem $f(A) = \frac{(n - \text{card } A)!}{n!}$ și, cum pentru $B' \neq B''$ $g(B')$ și $g(B'')$ sînt probabilități de evenimente disjuncte rezultă că $f(A) = \sum_{A \subset B} g(B)$. Avem de calculat $g(\emptyset)$ care conform principiului includerii și al excluderii este dată de egalitatea

$$g(\emptyset) = \sum_B (-1)^{\text{card } B} \frac{(n - \text{card } B)!}{n!} = \sum_{l=0}^n (-1)^l C_n^l \frac{(n-l)!}{n!} = \sum_{l=0}^n \frac{(-1)^l}{l!}.$$

1.56. Printre n bilete de examen, m sînt „favorabile”. Studenții vin pe rînd să tragă cîte un bilet.

Dintre primii doi studenți, care are șansa mai mare de a trage un bilet „favorabil”?

Soluție. Fie A evenimentul că primul student trage un bilet „favorabil” și B evenimentul că al doilea student trage un bilet „favorabil”.

Avem $P(A) = \frac{m}{n}$, iar din formula probabilității totale (vezi punctul h_2 din rezumatul teoretic) rezultă că

$$P(B) = P(A)P(B|A) + P(\bar{A})P(B|\bar{A}) = \frac{m}{n}P(B|A) + \left(1 - \frac{m}{n}\right)P(B|\bar{A})$$

Dar

$$P(B|A) = \frac{m-1}{n-1}, \quad P(B|\bar{A}) = \frac{m}{n-1}, \quad \text{deci } P(B) = \frac{m}{n} \cdot \frac{m-1}{n-1} + \left(1 - \frac{m}{n}\right) \cdot \frac{m}{n-1} = \frac{m}{n}.$$

Prin urmare cei doi studenți au aceeași șansă să tragă cîte un bilet „favorabil”.

1.57. Într-o urnă se află n bile. Sînt posibile $(n+1)$ ipoteze asupra numărului de bile albe din urnă: H_0, H_1, \dots, H_n unde H_i este ipoteza care constă în aceea că în urnă i sînt i bile albe.

Presupunem că toate aceste ipoteze sînt egal probabile, deci $P(H_i) = \frac{1}{n+1}$ pentru orice i .

Din urnă se extrage la întîmplare o bilă care este albă. Să se calculeze probabilitățile condiționate $P(H_i|B)$, $0 \leq i \leq n$, unde B este evenimentul ca extrăgînd o bilă din urnă, aceasta este albă.

Soluție. Avem $P(B|H_i) = \frac{i}{n}$, deci prin aplicarea formulei lui Bayes (vezi punctul h_3 din rezumatul teoretic) obținem

$$P(H_i|B) = \frac{\frac{1}{n+1} \cdot \frac{i}{n}}{\sum_{k=0}^n \frac{1}{n+1} \cdot \frac{k}{n}} = \frac{2i}{n(n+1)}.$$

Prin urmare cea mai mare probabilitate a posteriori (adică în ipostaza că B s-a produs) o are ipoteza H_n .

Observație. Problema precedentă este un caz particular al problemei 1.56, unde $A_i = \{\text{studentul } i \text{ trage un bilet favorabil}\}$.

1.58. Sultanul îi spune lui Ali Baba: „Iată două urne identice, m bile albe și n bile negre. Repartizează cum vrei bilele în cele două urne așa încît să nu fie nici o urnă vidă și eu voi extrage o bilă. Dacă bila va fi albă, îți dăruiesc viața”. Cum a făcut Ali Baba să-și maximizeze șansele?

Soluție. Fie U_1, U_2 cele două urne și să presupunem că în U_1 se află $x \geq 1$ bile albe și y bile negre, deci în U_2 se află $m - x$ bile albe și $n - y$ bile negre.

Fie $E_1^{x,y}$ mulțimea bilelor din U_1 , $E_2^{x,y}$ mulțimea bilelor din U_2 și $P_i^{x,y}$ probabilitatea clasică pe $E_i^{x,y}$.

Deoarece cele două urne sînt alese cu aceeași probabilitate $\frac{1}{2}$ rezultă că experimentul aleator constituit din extragerea unei bile din cele două urne este modelat de suma directă

$$(E_1^{x,y}, \mathfrak{E}(E_1^{x,y}), P_1^{x,y}) \oplus (E_2^{x,y}, \mathfrak{E}(E_2^{x,y}), P_2^{x,y}).$$

Fie F evenimentul de a extrage o bilă albă. Atunci

$$\begin{aligned} P^{x,y}(F) &= \frac{1}{2} P_1^{x,y}(F \cap E_1^{x,y}) + \frac{1}{2} P_2^{x,y}(F \cap E_2^{x,y}) = \\ &= \frac{1}{2} \left(\frac{x}{x+y} + \frac{m-x}{m+n-x-y} \right). \end{aligned}$$

(am notat cu $(E_1^{x,y} \cup E_2^{x,y}, \mathfrak{A}(E_1^{x,y} \cup E_2^{x,y}), P^{x,y})$ suma directă precedentă).

Vom arăta că dacă $1 \leq x \leq m-x$ (fapt ce nu restrânge generalitatea?)

$$P^{x,y}(F) \leq P^{x,0}(F) \leq P^{1,0}(F),$$

fapt ce arată că Ali Baba își mărește șansele dacă pune o bilă albă într-o urnă, iar restul bilelor în a doua urnă. Avem

$$\begin{aligned} P^{x,0}(F) - P^{x,y}(F) &= \frac{1}{2} \left(1 + \frac{m-x}{m+n-x} \right) - \frac{1}{2} \left(\frac{x+y}{x} + \frac{m-x}{m+n-x-y} \right) = \\ &= \frac{1}{2} \frac{y}{m+n-x-y} \left(\frac{m+n}{x+y} - 1 - \frac{m-x}{m+n-x} \right). \end{aligned}$$

Suma din paranteze este minimă dacă $y = n$ când ia valoarea $\frac{(m-x)(m-2x)}{(n+x)(m+n-x)} \geq 0$, deci $P^{x,0}(F) - P^{x,y}(F) \geq 0$ pentru orice x, y . Apoi este ușor de văzut că $P^{1,0}(F) \geq P^{x,0}(F)$ pentru $x \geq 1$.

1.59. Dintr-o urnă ce conține m bile roșii și $n-m$ bile albe se extrag la întâmplare toate bilele, fără întoarcere. Să se afle probabilitățile următoarelor evenimente:

$$A_i = \{\text{bila } i \text{ este roșie}\};$$

$$A_{i,j} = \{\text{bilele } i, j \text{ sînt roșii}\};$$

$$B_{i,j} = \{\text{bila } i \text{ este albă și bila } j \text{ este roșie}\}.$$

Soluție. Pentru $L \subset \{1, \dots, n\}$ cu card $L = m$ fie $D_L = \bigcap_{i=1}^n B_i$ unde $B_i = A_i$ dacă $i \in L$ și $B_i = \bar{A}_i$ dacă $i \notin L$.

Arătăm că $P(D_L) = \frac{1}{C_n^m}$ pentru orice L ca mai sus.

Aplicînd punctul h_1 din rezumatul teoretic obținem

$$P(D_L) = P(B_1)P(B_2|B_1) \dots P(B_n|B_{n-1} \cap \dots \cap B_1)$$

Fie $L = \{i_1, \dots, i_m\}$ cu $i_1 < \dots < i_m$. Avem

$$\begin{aligned} P(B_1) &= \frac{n-m}{n}, \quad P(B_2|B_1) = \frac{n-m-1}{n-1} \dots, \\ \dots, \quad P(B_{i_1-1}|B_{i_1-2} \cap \dots \cap B_1) &= \frac{n-m-i_1+2}{n-i_1+2} \end{aligned}$$

$$\begin{aligned}
P(B_{i_1} | B_{i_1-1} \cap \dots \cap B_1) &= \frac{m}{n - i_1 + 1}, \quad P(B_{i_1+1} | B_{i_1} \cap \dots \cap B_1) = \\
&= \frac{n - m - i_1 + 1}{n - i_1}, \dots, P(B_{i_2-1} | B_{i_2-2} \cap \dots \cap B_1) = \frac{n - m - i_2 + 1}{n - i_2 + 2}, \\
P(B_{i_2} | B_{i_2-1} \cap \dots \cap B_1) &= \frac{m - 1}{n - i_2}, \dots, P(B_{i_m-1} | B_{i_m-2} \cap \dots \cap B_1) = \\
&= \frac{n - i_m + 1}{n - i_m + 2}, \\
P(B_{i_m} | B_{i_m-1} \cap \dots \cap B_1) &= \frac{1}{n - i_m + 1}, \\
P(B_{i_m+k} | B_{i_m+k-1} \cap \dots \cap B_1) &= 1
\end{aligned}$$

pentru $1 \leq k \leq m - i_m$.

Prin urmare

$$P(D_L) = m! \frac{(n-m) \dots (n-i_m+1)}{n(n-1) \dots (n-i_m+1)} = m! \frac{(n-m)!}{n!} = \frac{1}{C_n^m}$$

Mai departe

$$\begin{aligned}
A_i &= \bigcup_{i \in L} D_L, \text{ deci } P(A_i) = \sum_{\substack{L \subset \{1, \dots, n\} \\ i \in L, \text{ card } L = m}} P(D_L) = \\
&= \sum_{\substack{L \subset \{1, \dots, n\} \\ i \in L, \text{ card } L = m}} \frac{1}{C_n^m} = \frac{C_{n-1}^{m-1}}{C_n^m} = \frac{m}{n}
\end{aligned}$$

De asemenea $A_{i,j} = \bigcup_{i,j \in L} D_L$, deci

$$P(A_{i,j}) = \sum_{i,j \in L} \frac{1}{C_n^m} = \frac{C_{n-2}^{m-2}}{C_n^m} = \frac{m(m-1)}{n(n-1)}$$

În fine $B_{i,j} = \bigcup_{\substack{i \in L \\ j \in L}} D_L$, deci

$$P(B_{i,j}) = \sum_{i \in L, j \in L} P(D_L) = \frac{C_{n-2}^{m-1}}{C_n^m} = \frac{m(n-m)}{n(n-1)}$$

Altă soluție. Fie R mulțimea bilelor roșii și A cea a bilelor albe. Evenimentele elementare E sînt n -trupluri distincte și cu componente distincte $a = (a_1, \dots, a_n)$ formate cu elementele din $R \cup A$, deci $\text{card } E = n!$

A_i sînt n -truplurile din E care pe locul i au elemente din R .

Aplicația $\varphi: A_i \rightarrow A_1$ definită prin $\varphi((a_1, \dots, a_n)) = (a_1, a_2, \dots, a_{i-1}, a_1, a_{i+1}, \dots, a_n)$ este o bijecție, deci $\text{card } A_i = \text{card } A_1 = m(n-1)!$. Rezultă că

$$P(A_i) = P(A_1) = \frac{\text{card } A_1}{\text{card } E} = \frac{m(n-1)!}{n!} = \frac{m}{n},$$

$A_{i,j}$ sînt n -truplurile care pe locurile i, j au elemente din R .

Aplicația $\varphi: A_{i,j} \rightarrow A_{1,2}$ definită prin $\varphi(a_1, \dots, a_n) = (B_1, \dots, b_n)$, unde $b_1 = a_i, b_2 = a_j, b_l = a_l, l \neq i, j$, este o bijecție, deci

$$\text{card } A_{i,j} = \text{card } A_{1,2} = m(m-1)(n-2)!$$

Deci

$$P(A_{i,j}) = \frac{m(m-1)(n-2)!}{n!} = \frac{m(m-1)}{n(n-1)}.$$

Asemănător se arată că avem $\text{card } B_{i,j} = \text{card } B_{1,2} = m(n-m) \times (n-2)!$ și prin urmare $P(B_{i,j}) = \frac{m(n-m)(n-2)!}{n!} = \frac{m(n-m)}{n(n-1)}$.

1.60. Un punct A este ales la întîmplare din intervalul $[0, 1]$. Fie l_1 (resp. l_2) lungimea segmentului mai mare (resp. mai mic) determinat de A pe $[0, 1]$

Să se calculeze $P(l_1 \leq x), P(l_2 \leq x)$.

Soluție. Fie l coordonata punctului A . Au loc egalitățile

$$\{l_1 \leq x\} = \{l \leq x\} \cap \{1-l \leq x\} = [1-x, x]; \quad \frac{1}{2} \leq x \leq 1$$

$$\{l_2 > x\} = \{l > x\} \cap \{1-l > x\} = (x, 1-x); \quad 0 \leq x \leq \frac{1}{2}$$

Prin urmare (m fiind măsura Lebesgue):

$$P(l_1 \leq x) = m([1-x, x]) = 2x - 1 \quad \text{dacă } \frac{1}{2} \leq x \leq 1$$

$$P(l_2 \leq x) = 1 - P(l_2 > x) = 1 - m((x, 1-x)) = 1 - (1 - 2x) = 2x \quad \text{dacă } 0 \leq x \leq \frac{1}{2}.$$

1.61. Doi prieteni au hotărît să se întîlnească într-un anumit loc. Ambii sosesc aleator în acel loc între orele a și $a + T$. Fiecare așteaptă un sfert de oră și apoi pleacă dacă celălalt nu a venit.

Care este probabilitatea întîlnirii lor?

Soluție. Fie s, t timpii de sosire ai celor doi prieteni. Punctele (s, t) determină în plan pătratul $E = [0, a + T]^2$.

Condiția ca cei doi prieteni să se întâlnească determină în plan mulțimea

$$F = \left\{ (s, t) \in E; |s - t| \leq \frac{1}{4} \right\}$$

Observăm că dacă $T \leq \frac{1}{4}$,

atunci $F = E$, iar dacă $T > \frac{1}{4}$ mulțimea F este regiunea hașurată din figură.

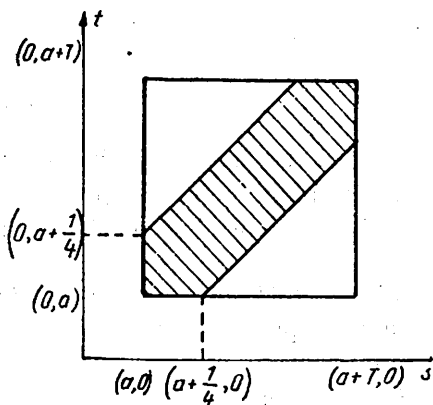


Fig. 2

Rezultă că

$$P(F) = \frac{\text{aria } F}{\text{aria } E} = \begin{cases} 1 & \text{dacă } T \leq \frac{1}{4} \\ \frac{T^2 - \left(T - \frac{1}{4}\right)^2}{T^2} = 1 - \left(1 - \frac{1}{4T}\right)^2 & \text{dacă } T > \frac{1}{4} \end{cases}$$

1.62. (Paradoxul lui Bertrand). Să se determine probabilitatea ca o coardă a unui cerc luată la întâmplare să fie mai mare decât latura triunghiului echilateral înscris în cerc.

Soluție. Considerăm ca fixă direcția coardei și ducem diametrul perpendicular pe această direcție.

Notăm cu A, B extremitățile diametrului și cu C, D mijloacele segmentelor OA, OB . Pentru ca lungimea coardei să fie mai mare decât latura triunghiului echilateral înscris în cerc trebuie ca mijlocul ei să fie situat între C și D . Dacă (ρ, θ) sînt coordonatele polare ale mijlocului M al coardei, atunci $-R \leq \rho \leq R$, $0 \leq \theta \leq 2\pi$ și M se află între C și D dacă $-\frac{R}{2} \leq \rho \leq \frac{R}{2}$, $0 \leq \theta \leq 2\pi$, R fiind raza cercului.

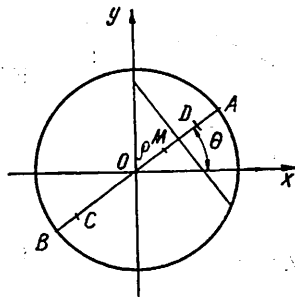


Fig. 3

Rezultă că probabilitatea căutată este (vezi punctul g din rezumatul teoretic)

$$p = \frac{2\pi R}{2\pi 2R} = \frac{1}{2}.$$

Bertrand a dat acestei probleme trei soluții (cea de față care corespunde cîmpului geometric și altele două ce folosesc alte cîmpuri de probabilitate) obținînd rezultate diferite.

1.63. (Problema acului lui Buffon). Un ac de lungime l este aruncat la întîmplare pe un parchet format din lame paralele de lățime a , unde $l < a$.

Să se afle probabilitatea ca acul să cadă pe una din liniile despărțitoare ale parchetului.

Soluție. Poziția acului este determinată de coordonatele (x, y) ale unei extremități ale sale (de exemplu A) și de unghiul θ pe care AB îl face cu axa Ox .

Presupunem că liniile despărțitoare sînt paralele cu axa Ox .

Nu restrîngem generalitatea dacă presupunem $0 \leq \theta \leq \frac{\pi}{2}$, $0 \leq y \leq a$. Pentru θ dat pozițiile acului AB care intersectează linia despărțitoare sînt date de banda hașurată din figura 4.

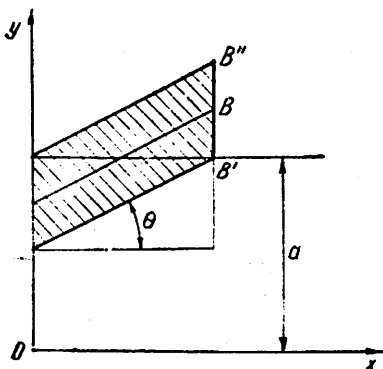


Fig. 4.

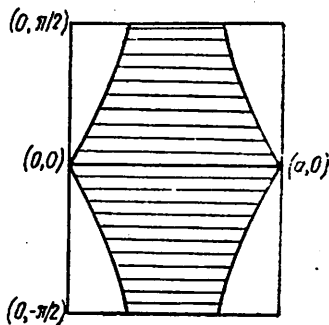


Fig. 5.

Rezultă că probabilitatea căutată este

$$p = \frac{\int_0^{\frac{\pi}{2}} \int_{a-l \sin \theta}^a d\theta dy}{\int_0^{\frac{\pi}{2}} \int_0^a d\theta dy} = \frac{\int_0^{\frac{\pi}{2}} l \sin \theta d\theta}{\frac{\pi}{2} a} = \frac{2l}{\pi a}.$$

Altă soluție. Determinăm poziția acului prin unghiul θ pe care îl face cu axa Ox și prin poziția mijlocului său M , dată de distanța d la Ox . Avem

$$0 \leq d \leq a, \quad -\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$$

Evenimentul ca acul să intersecteze linia despărțitoare este dat de domeniul hașurat din fig. 5, deci

$$F = \left\{ (d, \theta); d \leq \frac{l}{2} |\sin \theta| \text{ sau } d \geq a - \frac{l}{2} |\sin \theta| \right\}$$

Deci

$$P(F) = \frac{\text{aria } F}{\text{aria} \left([0, a] \times \left[-\frac{\pi}{2}, \frac{\pi}{2} \right] \right)} = \frac{2}{\pi a} \int_0^{\frac{\pi}{2}} \frac{l}{2} \sin \theta d\theta = \frac{2l}{\pi a}$$

CAPITOLUL II

REPARTIȚII ȘI VARIABLE ALEATOARE

Probabilitățile pe mulțimile boreliene din R^d se mai numesc *repartiții* sau *legi*. Din teorema de unicitate a probabilităților rezultă că o repartiție este unic determinată de valorile ei pe un sistem de generatori închis la intersecția finită, deci în particular de valorile ei pe orice tip de intervale.

Prin urmare dacă F este o repartiție, atunci aceasta este unic determinată de valorile funcției $\tilde{F}: R^d \rightarrow [0, 1]$, $\tilde{F}(x) = F((-\infty, x])$, Vom numi funcția \tilde{F} *funcția de repartiție* a lui F .

O clasă importantă de repartiții o formează cele care posedă densitate. Vom spune că repartiția F are densitatea ρ dacă:

a) $\rho \geq 0$ și ρ este boreliană,

$$b) \int_{R^d} \rho(x) dx = 1,$$

$$c) F(A) = \int_A \rho(x) dx,$$

pentru orice mulțime boreliană A sau este suficient ca A să parcurgă un sistem de generatori închis la intersecția finită (toate integralele vor fi integrale Lebesgue).

Observăm că densitatea unei repartiții este unic determinată în sensul că dacă ρ_1, ρ_2 satisfac a) - c), atunci $\rho_1 = \rho_2$ a.p.t. în raport cu măsura Lebesgue, deci în particular nu contează valorile densității pe o mulțime cel mult numărabilă.

Dintr-o teoremă a lui Lebesgue rezultă că dacă F are densitatea ρ atunci $\frac{\partial^n \tilde{F}}{\partial x_1 \dots \partial x_n}$ există și are loc egalitatea $\frac{\partial^n \tilde{F}}{\partial x_1 \dots \partial x_n} = \rho$ a.p.t. în raport cu măsura Lebesgue.

O repartiție F pe R^d este:

i) *discretă* dacă $F = \sum_{i \in I} a_i \varepsilon_{x_i}$, I cel mult numărabilă, $a_i > 0$, $\sum_{i \in I} a_i = 1$, $x_i \in R^d$, iar ε_x este repartiția definită prin $\varepsilon_x(A) = \lambda_A(x)$.

ii) *continuă* dacă $F(\{x\}) = 0$ pentru orice x .

Repartițiile discrete nu sînt continue, în schimb cele care posedă densitate sînt continue.

De asemenea o repartiție nu încarcă decît o mulțime cel mult numărabilă de puncte.

O funcție $f: (E, \mathcal{K}, R) \rightarrow (R^d, \mathfrak{B}_{R^d})$ care este măsurabilă (adică $f^{-1}(\mathfrak{B}_{R^d}) \subset \mathcal{K}$) se numește *variabilă aleatoare d -dimensională* ((E, \mathcal{K}, P) este un cîmp de probabilitate).

Deseori în loc de variabilă aleatoare 1-dimensională vom spune pe scurt variabile aleatoare și în loc de variabilă aleatoare d -dimensională, $d > 1$, vom spune vector aleator d -dimensional sau pe scurt vector aleator dacă d se subînțelege din context.

Probabilitatea $P \circ f^{-1}$ definită pe \mathfrak{B}_{R^d} prin egalitatea $(P \circ f^{-1})(A) = P(f \in A)$ se numește *repartiția variabilei aleatoare*.

Dacă f este discretă (adică $f(E)$ este cel mult numărabilă), atunci a cunoaște repartiția lui f este echivalent cu a cunoaște familia de numere $\{P(f = x)\}_{x \in f(E)}$. Pentru astfel de variabile aleatoare vom spune prin abuz de limbaj că sistemul de numere de mai sus este repartiția lui f . Funcția de repartiție asociată repartiției variabilei aleatoare f se notează cu F_f și se numește *funcția de repartiție a lui f* .

Dată fiind o repartiție F , există întotdeauna o variabilă aleatoare f avînd repartiția F .

Dacă variabilele aleatoare $(f_i)_{i \in I}$ au aceeași repartiție, vom mai spune că ele sînt *identic repartizate*.

O variabilă aleatoare are densitate dacă repartiția ei are densitate, deci variabila aleatoare f are densitatea ρ dacă:

a) $\rho \geq 0$ și ρ este boreliană;

$$b) \int_{R^d} \rho(x) dx = 1;$$

c) $P(f \in A) = \int_A \rho(x) dx$ pentru orice mulțime boreliană A sau este suficient

pentru A într-un sistem de generatori închis la intersecția finită.

Dacă (f_1, \dots, f_d) are densitatea ρ și $d' < d$, atunci $(f_1, \dots, f_{d'})$ are densitatea

$$q(y) = \int_{R^{d-d'}} \rho(x, y) dx.$$

O metodă de calcul a densităților este următoarea:

Fie w o transformare definită pe o mulțime deschisă $D \subset R^d$ și cu valori în $E \subset R^d$. Presupunem că w este bijectivă și cu jacobianul $|J_w(x)| \neq 0$ pentru orice $x \in D$.

Dacă f are densitatea ρ și $P(f \in D) = 1$, atunci $g = w(f)$ are densitatea

$$q(y) = \begin{cases} \rho(v(y)) |J_v(y)| & y \in E \\ 0 & y \notin E \end{cases} \quad (1)$$

(v este transformarea inversă a lui w).

În particular, dacă A este o matrice nesingulară de ordin d , $b \in R^d$, atunci variabila aleatoare $g = Af + b$ are densitatea

$$q(y) = \frac{1}{|\det A|} \rho(A^{-1}(y - b)) \quad (2)$$

Fie $f: (E, \mathcal{A}, P) \rightarrow R^d$ o variabilă aleatoare și $u: R^d \rightarrow R$ o funcție boreliană. Atunci are loc egalitatea următoare (cunoscută sub numele de *formula de transport* sau de *schimbare de variabilă*):

$$\int_E u(f) dP = \int_{R^d} u(x) d(P \circ f^{-1})(x) \quad (3)$$

în sensul că existența unui membru implică existența celuilalt și egalitatea lor.

În particular dacă f are densitatea ρ atunci (3) devine

$$\int_E u(f) dP = \int_{R^d} u(x) \rho(x) dx \quad (4)$$

Vom spune că variabilele aleatoare $(f_i)_{i \in I}$ sînt independente dacă corpurile boreliene $\{\mathfrak{B}(f_i)\}_{i \in I}$ sînt independente, cu alte cuvinte dacă pentru orice parte finită J a lui I și pentru orice $A_i \in \mathfrak{B}_{R^d}$ astfel încît $\{i; A_i \neq R^d\}$ este finită avem că:

$$P\left(\bigcap_{i \in I} \{f_i \in A_i\}\right) = \prod_{i \in I} P(f_i \in A_i).$$

Este vizibil că dacă I este finită atunci variabilele aleatoare $(f_i)_{i \in I}$ sînt independente dacă și numai dacă are loc egalitatea precedentă pentru orice $A_i \in \mathfrak{B}_{R^d}$.

Variabilele aleatoare discrete f_1, \dots, f_n sînt independente dacă și numai dacă

$$P(f_1 = x_1, \dots, f_n = x_n) = P(f_1 = x_1) \dots P(f_n = x_n) \quad (5)$$

pentru orice $x_i \in R^d$ sau numai pentru x_i în imaginea lui f_i .

Următoarele afirmații sînt echivalente:

i₁) f_1, \dots, f_n sînt independente;

i₂) $P(f_1 \in A_1, \dots, f_n \in A_n) = P(f_1 \in A_1) \dots P(f_n \in A_n)$

pentru orice $A_i \in \mathfrak{M}_i, \mathfrak{N}_i$ fiind un sistem de generatori închis la intersecția finită pentru \mathfrak{B}_{R^d} .

i₃) $P \circ (f_1, \dots, f_n)^{-1} = (P \circ f_1^{-1}) \otimes \dots \otimes (P \circ f_n^{-1})$ (cu alte cuvinte repartiția vectorului este egală cu produsul repartițiilor componentelor).

În termeni de funcții de repartiție și densități, independența se exprimă prin următoarele afirmații echivalente:

j₁) f_1, \dots, f_n sînt independente

j₂) $F_{(f_1, \dots, f_n)}(x_1, \dots, x_n) = F_{f_1}(x_1) \dots F_{f_n}(x_n)$ pentru orice x_i (cu alte cuvinte funcția de repartiție a vectorului este egală cu produsul funcțiilor de repartiție ale componentelor).

j₃) În ipoteza că f_i are densitatea ρ_i , funcția

$$(\rho_1 \otimes \dots \otimes \rho_n)(x_1, \dots, x_n) \stackrel{\text{def}}{=} \rho_1(x_1) \dots \rho_n(x_n)$$

este densitatea de repartiție a vectorului (f_1, \dots, f_n) .

În ipoteza de independență putem calcula densitatea și funcția de repartiție a sumei.

Dacă f_1, f_2 sînt variabile aleatoare independente atunci

$$P \circ (f_1 + f_2)^{-1} = (P \circ f_1^{-1}) * (P \circ f_2^{-1}) \quad (6)$$

deci repartiția sumei de variabile aleatoare independente este egală cu convoluția repartițiilor componentelor.

Dacă în plus f_i are densitatea ρ_i , atunci $f_1 + f_2$ are densitatea

$$(\rho_1 * \rho_2)(x_2) = \int \rho_1(x_1 - x_2) \rho_2(x_2) dx_2 \quad (7)$$

Un rezultat important este dat de *asociativitatea independenței* care afirmă că: dacă variabilele aleatoare $(f_i)_{i \in I}$ sînt independente și $(I_\alpha)_{\alpha \in A}$ este o partiție a lui I și pentru orice $\alpha \in A$ se dă $u_\alpha: R^{I_\alpha} \rightarrow R$ măsurabilă (pe R^{I_α} se consideră corpul borelian produs de \mathfrak{B}_R), atunci variabilele aleatoare $\{u_\alpha((f_i)_{i \in I_\alpha})\}_{\alpha \in A}$ sînt independente.

De asemenea să amintim și următorul rezultat care va fi folosit în mod tacit de multe ori. Fiind dată o familie $(F_i)_{i \in I}$ de repartiții există pe un cîmp de probabilitate (E, \mathcal{K}, P) o familie de variabile aleatoare independente $(f_i)_{i \in I}$ așa încît f_i are repartiția F_i pentru orice i .

În continuare vom asocia repartițiilor (deci și variabilelor aleatoare) o serie de parametri (caracteristici numerice).

Fie F o repartiție pe R și f o variabilă aleatoare. Atunci definim:

1) Momentul de ordin $r = 1, 2, \dots$, al lui F prin

$$M_r(F) = \int_R x^r dF(x)$$

cu condiția ca integrala să aibă sens.

Momentul de ordinul întâi se numește *media lui F* și se notează cu $M(F)$.

2) Momentul absolut de ordin r prin

$$\bar{M}_r(F) = \int_R |x|^r dF(x)$$

3) Momentul centrat de ordin r al lui F prin

$$\bar{M}_r^c(F) = \int_R [x - M(F)]^r dF(x)$$

cu condiția ca $\bar{M}_r(F) < \infty$.

Momentul centrat de ordinul doi al lui F se numește *dispersia (sau variația) lui F* și se notează cu $D^2(F)$ sau $\sigma^2(F)$,

Avem deci că

$$D^2(F) = \sigma^2(F) = \int_R [x - M(F)]^2 dF(x) = M_2(F) - [M(F)]^2$$

cu condiția ca $M_2(F) < \infty$.

4) Momentul de ordin r , momentul absolut de ordin r , momentul centrat de ordin r , media, dispersia (variația) lui f prin

$$M_r(f) = M_r(P \circ f^{-1}); \quad \bar{M}_r(f) = \bar{M}_r(P \circ f^{-1}); \quad M_r^c(f) = M_r^c(P \circ f^{-1})$$

$$M(f) = M(P \circ f^{-1}); \quad D^2(f) = \sigma^2(f) = D^2(P \circ f^{-1})$$

deci se definesc caracteristicile unei variabile aleatoare ca fiind caracteristicile repartiției sale.

5) Dacă A este un eveniment de probabilitate nenulă, definim media condiționată a lui f în raport cu A prin

$$M_A(f) = M(f|A) = \int f dP_A$$

Din formula de transport obținem valabilitatea egalităților

$$M_r(f) = \int f^r dP; \quad M_r(f) = \int |f|^r dP; \quad M_r^c(f) = \int (f - M(f))^r dP$$

$$D^2(f) = \int (f - M(f))^2 dP = \int f^2 dP - (M(f))^2$$

În particular dacă f este o variabilă aleatoare discretă (scrierea egalităților următoare pentru repartiții este imediată), avem

$$\begin{aligned} M_r(f) &= \sum_{x \in f(E)} x^r P(f=x); \quad D^2(f) = \sum_{x \in f(E)} (x - M(f))^2 P(f=x) = \\ &= \sum_{x \in f(E)} x^2 P(f=x) - \left\{ \sum_{x \in f(E)} x P(f=x) \right\}^2 \end{aligned}$$

iar dacă f are densitatea ρ , atunci

$$M_r(f) = \int_R x^r \rho(x) dx; \quad D^2(f) = \int_R (x - M(f))^2 \rho(x) dx = \int_R x^2 \rho(x) dx - \left(\int_R x \rho(x) dx \right)^2$$

Este vizibil că toate caracteristicile introduse sînt unic determinate de repartiție.

6) Dacă f și g sînt variabile cu momentul de ordinul doi finit definim covariația (sau corelația) dintre f și g prin

$$\text{cov}(f, g) = M(fg) - M(f)M(g) = M[(f - M(f))(g - M(g))].$$

7) Dacă $f = (f_1, \dots, f_d)$ este un vector aleator atunci definim media lui f prin $M(f) = (M(f_1), \dots, M(f_d))$ cu condiția ca $M(f_i)$ să existe pentru orice i , iar dacă $M_2(f_i) < \infty$ pentru orice i definim matricea de covariație (sau de corelație) a lui f prin $C_f = (\text{cov}(f_i, f_j))_{1 \leq i, j \leq d}$.

Dacă $f: (E, \mathcal{X}, P) \rightarrow R$ este o variabilă aleatoare de medie finită și \mathcal{F} este un corp borelian inclus în \mathcal{X} , atunci se definește media condiționată a lui f în raport cu \mathcal{F} ca fiind o variabilă aleatoare cu următoarele două proprietăți (ce o determină unic în sensul egalității a.s.):

k_1) $M[f|\mathcal{F}]$ este \mathcal{F} - măsurabilă și de medie finită

$$k_2) \int_A M[f|\mathcal{F}] dP = \int_A f dP \text{ pentru orice } A \in \mathcal{F}.$$

Dacă \mathcal{F} este corpul borelian generat de variabilele aleatoare $(f_i)_{i \in I}$ atunci vom spune $M[f|f_i; i \in I]$ în loc de $M[f|\mathcal{B}(f_i; i \in I)]$.

Date două variabile aleatoare f, g cu f avînd medie finită atunci prin $M[f|g = x]$ vom nota o funcție definită pe R și cu valori în R care este boreliană și satisfac egalitatea

$$M[f|g = x] \circ g = M[f|g] \quad P - a.s.$$

Astfel definită $M[f|g = x]$ este unică $P \circ g^{-1} - a.s.$

În conexiune cu caracteristicile variabilelor aleatoare în mod curent se folosesc următoarele proprietăți:

— *Inegalitatea lui Markov*: $P(|f| \geq c) < \frac{\overline{M_r(f)}}{c^r}$ pentru orice $c > 0$ și $r \geq 1$.

În particular pentru $r = 2$ și $f \rightarrow f - M(f)$ obținem *inegalitatea lui Cebîșev*

$$P(|f - M(f)| \geq c) < \frac{D^2(f)}{c^2} \quad (8)$$

— Dacă f_1, \dots, f_n sînt variabile aleatoare independente de medie finită atunci

$$M(f_1, \dots, f_n) = M(f_1) \dots M(f_n) \quad (\text{formula de multiplicare}) \quad (9)$$

Dacă f_1, \dots, f_n sînt variabile aleatoare independente de dispersie finită atunci

$$D^2(f_1 + \dots + f_n) = D^2(f_1) + \dots + D^2(f_n) \quad (10)$$

— Dacă f are media finită și $(A_i)_{i \in I}$ este o partiție cel mult numărabilă a lui E , (E, \mathcal{K}, P) este cîmpul de probabilitate pe care este definită f atunci

$$M(f) = \sum_{i \in I} P(A_i) M(f|A_i) \quad (\text{cu convenția că dacă } P(A_i) = 0 \text{ atunci } P(A_i) M(f|A_i) \stackrel{\text{def}}{=} 0) \quad (11)$$

— Dacă f, g sînt variabile aleatoare și f are media finită atunci

$$M(f) = \int_R M(f|g = x) d(P \circ g^{-1})(x). \quad (12)$$

— Dacă f, λ_A sînt independente, f, g sînt independente și f are medie finită, atunci

$$M(f|A) = M(f); \quad M(f|g) = M(f). \quad (13)$$

— Dacă $f = (f_1, \dots, f_d)$ este un vector aleator și A o matrice de ordin d (A^T înseamnă transpusa)

$$M(Af) = AM(f); \quad C_{Af} = AC_f A^T \quad (14)$$

Un instrument eficace în teoria probabilităților îl constituie funcția caracteristică și funcția generatoare.

Dată fiind F o repartiție pe R^d atunci aplicația $\varphi_F: R^d \rightarrow C$ definită prin $\varphi_F(t) = \int_{R^d} \exp(i \langle t, x \rangle) dF(x)$ se numește *funcția caracteristică a lui F*.

Pentru o variabilă aleatoare f funcția caracteristică asociată lui $P \circ f^{-1}$ se notează cu φ_f și se numește *funcția caracteristică a lui f*.

Prin urmare putem scrie

$$\varphi_f(t) = \int_{R^d} \exp(i \langle t, x \rangle) d(P \circ f^{-1})(x) = \int_E \exp(i \langle t, f \rangle) dP$$

Proprietățile mai importante ale funcțiilor caracteristice sînt (le scriem numai în termeni de variabile aleatoare):

— Dacă f_1, \dots, f_n sînt variabile aleatoare independente atunci

$$\varphi_{f_1 + \dots + f_n}(t) = \varphi_{f_1}(t) \varphi_{f_2}(t) \dots \varphi_{f_n}(t) \quad (15)$$

— Dacă variabila aleatoare f are $\overline{M}_r(f) < \infty$ atunci $\varphi_f^{(r)}(0)$ există și este dată de egalitatea

$$\varphi_f^{(r)}(0) = i^r \overline{M}_r(f) \quad (16)$$

— *Teorema de unicitate*: Dacă $\varphi_{f_1} = \varphi_{f_2}$ atunci $P \circ f_1^{-1} = P \circ f_2^{-1}$

— *Teorema de inversiune*: Dacă f este o variabilă aleatoare așa încît $\int_R |\varphi_f(t)| dt < \infty$

atunci f are densitatea $\rho(x)$ dată de egalitatea

$$\rho(x) = \frac{1}{(2\pi)^d} \int_{R^d} \exp(-i \langle t, x \rangle) \varphi_f(t) dt \quad (17)$$

— Variabilele aleatoare f_1, \dots, f_n sînt independente dacă și numai dacă

$$\varphi_{f_1, \dots, f_n}(t_1, \dots, t_n) = \varphi_{f_1}(t_1) \dots \varphi_{f_n}(t_n) \quad (18)$$

pentru orice t_i . Fie $F = \sum_{k=0}^{\infty} p_k \varepsilon_k$ o repartiție discretă, deci avem $p_k \geq 0$ și $\sum_{k=0}^{\infty} p_k = 1$.

Funcția $G_F: [-1, 1] \rightarrow R$ definită prin $G_F(z) = \sum_{k=0}^{\infty} p_k z^k$ se numește *funcția generatoare a lui F*.

Dacă f este o variabilă aleatoare cu valori în $\{0, 1, \dots\}$ atunci se definește funcția generatoare a lui f și se notează cu G_f ca fiind funcția generatoare a repartiției

$\{P(f = n)\}_n$ a lui f , deci $G_f(z) = \sum_{n=0}^{\infty} P(f = n) z^n$.

— Dacă două repartiții discrete (respectiv variabile aleatoare discrete) au aceeași funcție generatoare atunci repartițiile (respectiv repartițiile variabilelor aleatoare) coincid.

— Dacă f_1, \dots, f_n sînt variabile aleatoare discrete independente atunci

$$G_{f_1 + \dots + f_n}(z) = G_{f_1}(z) \dots G_{f_n}(z) \quad (19)$$

pentru orice $|z| < 1$.

— Dacă f este discretă și $\overline{M}(f) < \infty$ atunci $M(f) = g'(1)$ iar dacă $D^2(f) < \infty$ atunci $D^2(f) = G_f''(1) + G_f'(1) - [G_f'(1)]^2$

Încheiem considerațiile din acest paragraf prin listarea repartițiilor mai importante.

Tabela 1

Repartiții discrete 1-dimensionale

Denumirea (Notație)	Definiția	Media	Dispersia	Funcția generatoare
Dirac în punctul $a \in R$ (ε_a)	$\varepsilon_a(A) = \lambda_A(a)$	a	0	—
Uniformă pe mulțimea $\{1, \dots, n\}$	$\sum_{k=1}^n \frac{1}{n} \varepsilon_k$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$	$\frac{1}{n} \cdot \frac{1-z^n}{1-z}$
Geometrică de parametru $0 < p < 1$	$\sum_{k=0}^{\infty} (1-p)p^k \varepsilon_k$	$\frac{p}{1-p}$	$\frac{p}{(1-p)^2}$	$\frac{1-p}{1-pz}$
Binomială de parametri $n, 0 < p < 1$ ($B_n(p)$)	$\sum_{k=0}^n C_n^k p^k (1-p)^{n-k} \varepsilon_k$	np	$np(1-p)$	$(1+p(z-1))^n$
Poisson de parametru $\lambda > 0$ (π_λ)	$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} \varepsilon_k$	λ	λ	$e^{\lambda(z-1)}$

Repartiții continue 1-dimensionale

Denumirea (Notație)	Forma densității	Media	Dispersia	Funcția caract.
Uniformă pe (a, b) ($U_{a, b}$)	$\frac{1}{b-a}; a < x < b$	$\frac{a+b}{2}$	$\frac{1}{12}(b-a)^2$	$\frac{1}{b-a} \cdot \frac{e^{ib} - e^{ia}}{it}$
Exponențială de parametru $\lambda > 0$ (e_λ)	$\lambda e^{-\lambda x}; x \geq 0$	λ^{-1}	λ^{-2}	$\frac{\lambda}{\lambda - it}$
Normală de parametri m, σ^2 ($N(m, \sigma^2)$)	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}; x \in R$	m	σ^2	$e^{imt - \frac{\sigma^2 t^2}{2}}$
Cauchy de parametri α, β ($C_{\alpha, \beta}$)	$\frac{1}{\pi} \frac{\alpha}{\alpha^2 + (x-\beta)^2}; x \in R$	nu are	nu are	$e^{-\alpha t + i\beta t}$
Gama de parametri $\alpha, \beta(\gamma(\alpha, \beta))$	$\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}; x \geq 0$	$\alpha\beta^{-1}$	$\alpha\beta^{-2}$	$\beta^\alpha(\beta - it)^{-\alpha}$

Repartiții multidimensionale. 1) Repartiția Dirac în punctul $a \in R^d$ este definită prin $\varepsilon_a(A) = \lambda_A(a)$. Dacă vectorul aleator f are repartiția ε_a atunci

$$M(f) = a; C_f = 0, \varphi_f(t) = e^{i\langle t, a \rangle}.$$

1) Repartiția multidimensională de parametri $n, 0 < p_1, \dots, p_d < 1$ este definită prin

$$M_n(p_1, \dots, p_d) = \sum_{\substack{n_i \geq 0 \\ n_1 + \dots + n_d = n}} \frac{n!}{n_1! \dots n_d!} p_1^{n_1} \dots p_d^{n_d} \varepsilon_{(n_1, \dots, n_d)}.$$

Dacă vectorul aleator f are repartiția $M_n(p_1, \dots, p_d)$ atunci

$$M(f) = (np_1, \dots, np_d); C_f = \begin{pmatrix} np_1(1-p_1) & -np_1p_2 & \dots & -np_1p_d \\ -np_2p_1 & np_2(1-p_2) & \dots & -np_2p_d \\ \dots & \dots & \dots & \dots \\ -np_dp_1 & -np_dp_2 & \dots & np_d(1-p_d) \end{pmatrix}$$

$$\varphi_f(t_1, \dots, t_d) = (p_1 e^{it_1} + \dots + p_d e^{it_d})^n$$

1₂) Repartiția uniformă pe mulțimea $D = (a_1, b_1) \times \dots \times (a_d, b_d)$ este definită prin densitatea $\rho(x) = \frac{1}{(b_1 - a_1) \dots (b_d - a_d)} \lambda_D(x)$. Dacă vectorul aleator f are repartiția uniformă pe D atunci

$$M(f) = \left(\frac{a_1 + b_1}{2}, \dots, \frac{a_d + b_d}{2} \right); \quad C_f = \left\{ \frac{1}{12} (b_i - a_i)^2 \delta_{i,j} \right\}_{1 \leq i, j \leq d}$$

$$\varphi_f(t_1, \dots, t_d) = \frac{1}{\prod_{k=1}^d (b_k - a_k)} \prod_{k=1}^d \frac{e^{it_k b_k} - e^{it_k a_k}}{it_k}$$

1₃) Repartiția normală d -dimensională de parametri m, B (unde $m \in R^d$ și B este o matrice pătrată de ordin d , simetrică și pozitiv definită) este repartiția $N_d(m, B)$ de densitate

$$\rho(x) = \left[\frac{1}{(2\pi)^d \det B} \right]^{\frac{1}{2}} \exp \left(-\frac{1}{2} \langle B^{-1}(x - m), x - m \rangle \right).$$

Dacă f are repartiția $N_d(m, B)$ atunci

$$M(f) = m; \quad C_f = B; \quad \varphi_f(t) = e^{i \langle t, m \rangle - \frac{1}{2} \langle t, Bt \rangle}$$

§1. Repartiția variabilelor aleatoare

2.1. Fie g o variabilă aleatoare cu funcția de repartiție F . Să se calculeze funcția de repartiție a variabilei aleatoare $f = F(g)$. Să se deducă apoi că dacă F este continuă atunci f are repartiția uniformă pe $(0, 1)$.

Soluție. Pentru $y \in (0, 1)$ fie $s(y) = \sup \{x; F(x) < y\}$. Deoarece F este nedescrescătoare și continuă la stînga rezultă pentru $y \in (0, 1)$ că

$$\{x; F(x) < y\} = \begin{cases} (-\infty, s(y)) & \text{dacă } F(s(y)) = y \\ (-\infty, s(y)) & F(s(y)) < y \end{cases}$$

Prin urmare

$$F_f(y) = P(F(g) < y) = \begin{cases} P(g < s(y)) = F(s(y)) = y & \text{dacă } F(s(y)) = y \\ P(g \leq s(y)) = F(s(y) + 0) & \text{dacă } F(s(y)) < y \end{cases}$$

Este clar că $F_f(y) = 0$ dacă $y \leq 0$ și $F_f(y) = 1$ dacă $y > 1$.

Apoi $F_f(1) = \lim_{\substack{y \rightarrow 1 \\ y < 1}} F_f(y)$ (orice funcție de repartiție este continuă la stînga). Repartiția lui f este deci următoarea :

$$F_f(y) = \begin{cases} 0 & \text{dacă } y \leq 0 \\ y & \text{dacă } 0 < y < 1 \text{ și } F(s(y)) = y \\ F(s(y) + 0) & \text{dacă } 0 < y < 1 \text{ și } F(s(y)) < y \\ \lim_{\substack{x \rightarrow 1 \\ x < 1}} F_f(x) & \text{dacă } y = 1 \\ 1 & \text{dacă } y > 1 \end{cases}$$

Dacă F este continuă atunci $F(s(y)) = y$ pentru orice $y \in (0, 1)$, așa că în acest caz obținem egalitatea

$$F_f(y) = \begin{cases} 0 & y \leq 0 \\ y & 0 < y < 1 \\ 1 & y \geq 1 \end{cases}$$

cu alte cuvinte f are repartiția uniformă pe $(0, 1)$.

2.2. Să se arate că dacă variabilele aleatoare f, g sînt egale a.s atunci ele au aceeași repartiție.

Soluție. Fie $B = \{f \neq g\}$. Prin ipoteză avem că $P(B) = 0$. Apoi din incluziunea $\{f \in A\} \subset \{g \in A\} \cup B$ rezultă că

$P(f \in A) \leq P(g \in A) + P(B) = P(g \in A)$. Analog se obține că

$P(g \in A) \leq P(f \in A)$, așa că $P(f \in A) = P(g \in A)$.

2.3. Fie f, g variabile aleatoare independente.

a) Dacă f, g au repartiția discretă $P(f = x_k) = P(g = x_k) = p_k$; $k = 1, \dots, n$ să se afle $P(f = g)$.

b) Dacă f sau g are repartiția continuă să se arate că $P(f = g) = 0$.

Soluție. a) $P(f = g) = P\left(\bigcup_{k=1}^n \{f = x_k; g = x_k\}\right) = \sum_{k=1}^n P(f = x_k; g = x_k) = \sum_{k=1}^n p_k^2$

b) Fie $\Delta = \{(x, x); x \in R\}$. Atunci se știe că

$$P \circ (f, g)^{-1} = (P \circ f^{-1}) \otimes (P \circ g^{-1}).$$

Să presupunem de exemplu că f are repartiție continuă.

Rezultă că

$$P(f = g) = [P \circ (f, g)^{-1}] (\Delta) = [(P \circ f^{-1}) \otimes (P \circ g^{-1})] (\Delta) = \\ = \int (P \circ f^{-1})(\Delta_y) d(P \circ g^{-1})(y) = \int P(f = y) d(P \circ g^{-1})(y) = 0$$

deoarece $P(f = y) = 0$ pentru orice y .

2.4. Vectorul aleator (f_1, \dots, f_n) are densitatea de forma $p(x_1, \dots, x_n) = p_1(x_1) \dots p_n(x_n)$, unde $\int_R p_i(x) dx = 1$ și $p_i \geq 0$ pentru orice i .

Să se arate că variabilele aleatoare f_1, \dots, f_n sînt independente și că f_i are densitatea p_i .

Soluție. Dacă A este o mulțime boreliană atunci

$$P(f_i \in A) = P(f_1 \in R, \dots, f_{i-1} \in R, f_i \in A, f_{i+1} \in R, \dots, f_n \in R) =$$

$$= \int_R \dots \int_R \int_A \dots \int_R p_1(x_1) \dots p_n(x_n) dx_1 \dots dx_n =$$

(teorema lui Fubini)

$$= \int_R p_1(x_1) dx_1 \dots \int_R p_{i-1}(x_{i-1}) dx_{i-1} \int_A p_i(x_i) dx_i \int_R p_{i+1}(x_{i+1}) dx_{i+1} \dots \\ \dots \int_R p_n(x_n) dx_n = \int_A p_i(x_i) dx_i$$

cu alte cuvinte p_i este densitatea lui f_i .

Independența variabilelor aleatoare f_1, \dots, f_n este acum cunoscută (vezi afirmațiile $j_1 - j_3$ din rezumatul de la începutul capitolului).

2.5. Fie f_1, \dots, f_n, \dots un șir de variabile aleatoare independente, identic repartizate și fie $\tau_1 < \tau_2 < \dots < \tau_n < \dots$ un alt șir de variabile aleatoare cu valori în $\{1, \dots, n, \dots\}$ așa încît $\{\tau_j = n\} \in \mathfrak{B}(f_1, \dots, f_{n-1})$ pentru orice n, j .

Să se arate că pentru orice n variabila aleatoare f_{τ_n} are aceeași repartiție ca f_1 și că variabilele $f_{\tau_1}, \dots, f_{\tau_n}, \dots$ sînt independente.

Soluție. În primul rând din egalitatea

$$\{f_{\tau_n} \in A\} = \bigcup_{j=1}^{\infty} \{f_j \in A, \tau_n = j\}$$

rezultă că f_{τ_n} este variabilă aleatoare.

Din asociativitatea independenței rezultă că pentru $n_1 < \dots < n_k$ evenimentele

$$\{\tau_1 = n_1, f_{n_1} \in A_1, \dots, \tau_{k-1} = n_{n_{k-1}} \in A_{k-1}, \tau_k = n_k\} \in \mathfrak{B}(f_1, \dots, f_{n_{k-1}})$$

$$\{f_{n_k} \in A_k\} \in \mathfrak{B}(f_{n_k})$$

sînt independente. Avem

$$P(f_{\tau_n} \in A) = \sum_{j=1}^{\infty} P(f_{\tau_n} \in A, \tau_n = j) = \sum_{j=1}^{\infty} P(f_j \in A, \tau_n = j) =$$

(asociativitatea independenței: $\{f_j \in A\}$ și $\{\tau_n = j\} \in \mathfrak{B}(f_1, \dots, f_{j-1})$ sînt independente).

$$= \sum_{j=1}^{\infty} P(f_j \in A) P(\tau_n = j) = P(f_1 \in A) \sum_{j=1}^{\infty} P(\tau_n = j) =$$

$$= P(f_1 \in A) P\left(\bigcup_{j=1}^{\infty} \{\tau_n = j\}\right) = P(f_1 \in A)$$

cu alte cuvinte f_{τ_n} și f_1 au aceeași repartiție.

Mai departe

$$P(f_{\tau_1} \in A_1, \dots, f_{\tau_k} \in A_k) = \sum_{1 \leq n_1 < \dots < n_k < \infty} P(\tau_1 = n_1, f_{n_1} \in A_1, \dots, \tau_k =$$

$$= n_k, f_{n_k} \in A_k) = \sum_{1 \leq n_1 < \dots < n_k < \infty} P(\tau_1 = n_1, \dots, f_{n_{k-1}} \in A_{k-1},$$

$$\tau_k = n_k) P(f_{n_k} \in A_k) = \sum_{1 \leq n_1 < \dots < n_k < \infty} P(\tau_1 = n_1, \dots, f_{n_{k-1}} \in A_{k-1},$$

$$\tau_k = n_k) P(f_1 \in A_k) = \sum_{1 \leq n_1 < \dots < n_{k-1} < \infty} P(\tau_1 = n_1, \dots, \tau_{k-1} = n_{k-1}$$

$$f_{n_{k-1}} \in A_{k-1}) P(f_1 \in A_k) = P(f_{\tau_1} \in A_1, \dots, f_{\tau_{k-1}} \in A_{k-1}) P(f_1 \in A_k) =$$

$$= \dots = P(f_1 \in A_1) \dots P(f_1 \in A_k) = P(f_{\tau_1} \in A_1) \dots P(f_{\tau_k} \in A_k)$$

deci $f_{\tau_1}, \dots, f_{\tau_k}$ sînt independente.

2.6. Vectorul aleator (f, g) are densitatea $\rho(x, y)$. Să se arate că variabilele aleatoare $f + g$, $f - g$ au densitățile

$$\rho_1(x) = \int_R \rho(x - y, y) dy; \quad \rho_2(x) = \int_R \rho(x + y, y) dy$$

În plus dacă $P(g \neq 0) = 1$ să se arate că fg , $\frac{f}{g}$ sau densitățile

$$q_1(x) = \int_R |y|^{-1} \rho(xy^{-1}, y) dy; \quad q_2(x) = \int_R |y| \rho(xy, y) dy$$

Soluție. Fie w transformarea $y_1 = x_1 + x_2$, $y_2 = x_2$. Inversa v a lui w este $x_1 = y_1 - y_2$, $x_2 = y_2$ al cărei iacobian este $J_v(y_1, y_2) = 1$. Din formula (1) obținem că vectorul $w(f, g) = (f + g, g)$ are densitatea $\rho(y_1 - y_2, y_2)$ de unde rezultă că $f + g$ are densitatea $\rho_1(y_1) = \int_R \rho(y_1 - y_2, y_2) dy_2$.

Pentru $f - g$ se ia transformarea $y_1 = x_1 - x_2$, $y_2 = x_2$ și se procedează la fel ca mai sus.

Pentru $\frac{f}{g}$ considerăm transformarea $w: R \times [R \setminus \{0\}] \rightarrow R \times [R \setminus \{0\}]$ definită prin $y_1 = \frac{x_1}{x_2}$, $y_2 = x_2$ a cărei inversă este $x_1 = y_1 y_2$, $x_2 = y_2$. Iacobianul invers este $J_v(y_1, y_2) = y_2$ așa că vectorul $w(f, g) = \left(\frac{f}{g}, g\right)$ are densitatea $|y_2| \rho(y_1 y_2, y_2)$ de unde obținem că densitatea citului $\frac{f}{g}$ este dată de

$$\int_R |y_2| \rho(y_1 y_2, y_2) dy_2$$

Pentru produs se ia transformarea $y_1 = x_1 x_2$, $y_2 = x_2$ și se procedează în mod similar.

Altă soluție.

$$\begin{aligned} P(f + g < x) &= \int_{y+z < x} \rho(y, z) dy dz = \int_R \left(\int_{-\infty}^{x-z} \rho(y, z) dy \right) dz = \\ &= \int_R \left(\int_{-\infty}^x \rho(u - z, z) du \right) dz \end{aligned}$$

de unde derivând în raport cu x obținem expresia lui ρ din enunț. Se procedează analog pentru $f - g$.

$$\begin{aligned} \text{Dacă punem } h = \frac{f}{g}, \text{ atunci } F_h(x) &= P\left(\frac{f}{g} < x\right) = \int_0^{\infty} \int_{-\infty}^{zx} \rho(y, z) dy dz + \\ &+ \int_{-\infty}^0 \int_{zx}^{\infty} \rho(y, z) dy dz \text{ de unde derivând în raport cu } x \text{ obținem } \rho_h(x) = \\ &= \int_0^{\infty} z \rho(zx, z) dz - \int_{-\infty}^0 z \rho(zx, z) dz = \int_{-\infty}^{+\infty} |z| \rho(zx, z) dz. \end{aligned}$$

$$\begin{aligned} \text{În fine } P(fg < x) &= \int_0^{\infty} \int_{-\infty}^{\frac{x}{z}} \rho(y, z) dy dz + \int_{-\infty}^0 \int_{\frac{x}{z}}^{\infty} \rho(y, z) dy dz = \\ &= \int_0^{\infty} \frac{1}{z} \int_{-\infty}^{\frac{x}{z}} \rho\left(\frac{u}{z}, z\right) du dz + \int_{-\infty}^0 \int_{\frac{x}{z}}^{\infty} \frac{1}{z} \rho\left(\frac{u}{z}, z\right) du dz, \text{ de unde derivând în} \\ &\text{raport cu } x \text{ obținem expresia lui } q_1. \end{aligned}$$

2.7. Vectorul aleator (f, g) are densitatea $p(x, y)$. Să se afle densitățile variabilelor aleatoare $m = \min(f, g)$ și $M = \max(f, g)$.

Soluție. Avem

$$\begin{aligned} P(m \geq x) &= P(f \geq x, g \geq x) = \int_x^{\infty} \int_x^{\infty} p(y, z) dy dz \text{ deci, } P(m < x) = \\ &= 1 - \int_x^{\infty} \int_x^{\infty} p(y, z) dy dz, \text{ de unde prin derivare în raport cu } x \text{ obținem} \\ &\text{că } m \text{ are densitatea} \end{aligned}$$

$$\rho_m(x) = \int_x^{\infty} p(y, x) dy + \int_x^{\infty} p(x, z) dz.$$

De asemenea

$$P(M < x) = P(f < x, g < x) = \int_{-\infty}^x \int_{-\infty}^x p(y, z) dy dz,$$

de unde prin derivare în raport cu x obținem că M are densitatea

$$\rho_M(x) = \int_{-\infty}^x p(x, y) dy + \int_{-\infty}^x p(z, x) dz$$

2.8. Variabilele aleatoare f_1, \dots, f_n sînt independente și au repartițiile F_1, \dots, F_n . Să se afle funcția de repartiție a vectorului aleator $(\min_k f_k, \max_k f_k)$.

Dacă în plus f_k are pentru orice k , densitatea p să se calculeze densitatea variabilei $\max_k f_k - \min_k f_k$.

Soluție. Fie $m = \min_k f_k$, $M = \max_k f_k$. Funcția de repartiție a vectorului (m, M) este

$$\begin{aligned} F(x, y) &= P(m < x, M < y) = P(M < y) - P(M < y, m \geq x) = \\ &= P(f_1 < y, \dots, f_n < y) - P(M < y, m \geq x) = \prod_{k=1}^n P(f_k < y) - \\ &\quad - P(M < y, m \geq x) \end{aligned}$$

Dacă $y \leq x$ atunci $\{M < y \leq x \leq m\} = \emptyset$, deci $P(M < y, m \geq x) = 0$ iar dacă $y > x$ atunci $\{x \leq m \leq M < y\} = \bigcap_{k=1}^n \{x \leq f_k < y\}$, deci

$$P(M < y, m \geq x) = \prod_{k=1}^n P(x \leq f_k < y) = \prod_{k=1}^n [F_k(y) - F_k(x)]$$

Rezultă că

$$F(x, y) = \begin{cases} \prod_{k=1}^n F_k(y) & \text{dacă } y \leq x \\ \prod_{k=1}^n F_k(y) - \prod_{k=1}^n [F_k(y) - F_k(x)] & \text{dacă } y > x \end{cases}$$

În cazul cînd $F_k = F$ pentru orice k obținem formula

$$F(x, y) = \begin{cases} [F(y)]^n & \text{dacă } y \leq x \\ [F(y)]^n - [F(y) - F(x)]^n & \text{dacă } y > x \end{cases}$$

Să presupunem că f_k are densitatea p . Rezultă că (m, M) are densitatea

$$\rho(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} = \begin{cases} 0 & \text{dacă } x \geq y \\ n(n-1) [F(y) - F(x)]^{n-2} p(x) p(y) & \text{dacă } x < y \end{cases}$$

Fie $k = M - m$. Din problema 2.6 rezultă că densitatea lui k este dată de

$$q(x) = \int_K p(y, y+x) dy \text{ dacă } x > 0 \text{ și } q(x) = 0 \text{ dacă } x \leq 0.$$

2.9. Fie f, g variabile aleatoare independente. Să se arate că

a) Dacă f are repartiția $B_m(p)$ și g are repartiția $B_n(p)$ atunci $f + g$ are repartiția $B_{m+n}(p)$.

b) Dacă f are repartiția π_λ și g are repartiția π_μ atunci $f + g$ are repartiția $\pi_{\lambda+\mu}$.

c) Dacă f are repartiția $N(m_1, \sigma_1^2)$ și g are repartiția $N(m_2, \sigma_2^2)$ atunci $f + g$ are repartiția $N(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$.

d) Dacă f are repartiția gama de parametri (α_1, β) și g are repartiția gama de parametri (α_2, β) , atunci $f + g$ are repartiția gama de parametri $(\alpha_1 + \alpha_2, \beta)$.

Soluție. a) Avem că

$$\begin{aligned} P(f + g = k) &= \sum_{j=0}^k P(f = j) P(g = k - j) = \\ &= \sum_{j=0}^k C_m^j p^j (1-p)^{m-j} C_n^{k-j} p^{k-j} (1-p)^{n-k+j} = \\ &= p^k (1-p)^{m+n-k} \sum_{j=0}^k C_m^j C_n^{k-j} = C_{m+n}^k p^k (1-p)^{m+n-k} \end{aligned}$$

pentru $k = 0, 1, \dots, m+n$, deci $f + g$ are repartiția $B_{m+n}(p)$.

b) Analog avem

$$\begin{aligned} P(f + g = k) &= \sum_{j=0}^k \frac{1}{j!} \lambda^j e^{-\lambda} \frac{1}{(k-j)!} \mu^{k-j} e^{-\mu} = \frac{e^{-(\lambda+\mu)}}{k!} \sum_{j=0}^k C_k^j \lambda^j \mu^{k-j} = \\ &= \frac{1}{k!} (\lambda + \mu)^k e^{-(\lambda+\mu)}, \text{ deci } f + g \text{ are repartiția } \pi_{\lambda+\mu} \end{aligned}$$

c) Fie $f_1 = f - m_1, f_2 = f - m_2$. Observăm că f_1, f_2 sînt independente și au repartițiile $N(0, \sigma_1^2), N(0, \sigma_2^2)$. Cum $f + g = f_1 + f_2 + m_1 + m_2$, rezultă că avem de arătat că $f_1 + f_2$ are repartiția $N(0, \sigma_1^2 + \sigma_2^2)$.

Din (7) avem că $f_1 + f_2$ are densitatea

$$\frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{+\infty} \exp\left[-\frac{(x-y)^2}{2\sigma_1^2} - \frac{y^2}{2\sigma_2^2}\right] dy$$

Această densitate se mai scrie

$$\begin{aligned} & \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{x^2}{2\sigma_1^2}\right) \int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2}\left(\frac{\sigma_1^2 + \sigma_2^2}{\sigma_1^2\sigma_2^2}\right)y^2 - 2\frac{xy}{\sigma_1^2}\right] dy = \\ & = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left\{-\frac{1}{2}\left[\frac{x^2}{\sigma_1^2} - \frac{\sigma_2^2}{\sigma_1^2(\sigma_1^2 + \sigma_2^2)}x^2\right]\right\} \int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2}\left(\frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sigma_1\sigma_2}y - \right. \right. \\ & \left. \left. - \frac{\sigma_2}{\sigma_1\sqrt{\sigma_1^2 + \sigma_2^2}}x\right)^2\right] dy = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left\{-\frac{x^2}{2(\sigma_1^2 + \sigma_2^2)}\right\} \frac{\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2} \int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} du = \\ & = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} \exp\left[-\frac{x^2}{2(\sigma_1^2 + \sigma_2^2)}\right] \end{aligned}$$

fapt ce arată că $f_1 + f_2$ are repartiția $N(0, \sigma_1^2 + \sigma_2^2)$.

d) Din (7) rezultă că pentru $x \geq 0$ densitatea lui $f + g$ este dată de

$$\begin{aligned} & \int_0^x \frac{\beta^{\alpha_1}}{\Gamma(\alpha_1)} (x-y)^{\alpha_1-1} e^{-\beta(x-y)} \frac{\beta}{\Gamma(\alpha_2)} y^{\alpha_2-1} e^{-\beta y} dy = \\ & = \frac{\beta^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} e^{-\beta x} \int_0^x (x-y)^{\alpha_1-1} \cdot y^{\alpha_2-1} dy = \end{aligned}$$

(cu schimbarea de variabilă $\frac{x}{y} = t$).

$$= \frac{\beta^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} x^{\alpha_1+\alpha_2-1} e^{-\beta x} \int_0^1 t^{\alpha_2-1} (1-t)^{\alpha_1-1} dt$$

și cum este valabilă egalitatea

$$\int_0^1 t^{\alpha_2-1} (1-t)^{\alpha_1-1} dt = B(\alpha_1, \alpha_2) = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1 + \alpha_2)}$$

ultima expresie devine

$$\frac{\beta^{\alpha_1 + \alpha_2}}{\Gamma(\alpha_1 + \alpha_2)} x^{\alpha_1 + \alpha_2 - 1} e^{-\beta x},$$

fapt ce arată că $f + g$ are repartiția gama de parametri

$$(\alpha_1 + \alpha_2, \beta).$$

2.10. Variabilele aleatoare f, g, h sînt uniform repartizate pe mulțimea $\{-n, -n + 1, \dots, -1, 0, 1, \dots, n - 1, n\}$ și f, h sînt independente. Fie α, β, γ variabilele aleatoare astfel încît valorile trinomialului $\alpha x^2 + \beta x + \gamma$ în punctele 1, 2, 3 să fie egale cu f, g, h .

Să se calculeze probabilitatea p_n ca numerele α, β, γ să fie întregi.

Soluție. Prin rezolvarea sistemului

$$\begin{cases} \alpha + \beta + \gamma = f \\ 4\alpha + 2\beta + \gamma = g \\ 9\alpha + 3\beta + \gamma = h \end{cases}$$

se obține soluția $\alpha = \frac{f+h}{2} - g, \beta = 4g - f - \frac{3}{2}(f+h), \gamma = 3f - 3g + h$

Este vizibil că α, β, γ sînt întregi dacă și numai dacă $f + h$ este par. Rezultă că

$$\begin{aligned} p_n &= P(f + h = \text{par}) = P(f, h = \text{pare}) + P(f, h = \text{impare}) = \\ &= P(f = \text{par}) P(h = \text{par}) + P(f = \text{impar}) P(h = \text{impar}) = \\ &= [P(f = \text{par})]^2 + [P(f = \text{impar})]^2. \end{aligned}$$

Dacă n este par atunci $P(f = \text{par}) = \frac{n+1}{2n+1}, P(f = \text{impar}) = \frac{n}{2n+1}$, deci

$$p_n = \frac{2n^2 + 2n + 1}{(2n + 1)^2}.$$

Dacă n este impar, atunci $P(f = \text{par}) = \frac{n}{2n+1}, P(f = \text{impar}) = \frac{n+1}{2n+1}$, deci

$$p_n = \frac{2n^2 + 2n + 1}{(2n + 1)^2}.$$

2.11. Variabilele aleatoare f, g sînt independente și uniform repartizate pe mulțimea $\{0, 1, \dots, 9\}$, suma $f + g$ se poate sub forma $f + g = 10\alpha_1 + \alpha_2, 0 \leq \alpha_i \leq 9$.

Să se afle repartițiile variabilelor aleatoare α_1, α_2 și să se vadă dacă α_1, α_2 sînt independente.

Soluție. Observăm că α_1 ia numai valorile 0, 1. Avem că :

$$\begin{aligned} P(\alpha_1 = 0) &= P(f + g = k, \text{ pentru } k \text{ între } 0 \text{ și } 9) = \\ &= \sum_{k=0}^{10} P(f + g = k) = \sum_{k=0}^9 \sum_{l=0}^k P(f = l) P(g = k - l) = \\ &= \sum_{k=0}^9 \sum_{l=0}^k \left(\frac{1}{10}\right)^2 = \sum_{k=0}^9 (k + 1) \left(\frac{1}{10}\right)^2 = \frac{1}{10^2} \cdot \frac{10 \cdot 11}{2} = 0,55 \end{aligned}$$

$$P(\alpha_1 = 1) = 1 - P(\alpha_1 = 0) = 0,45.$$

Dacă $0 \leq k \leq 9$ avem :

$$\begin{aligned} P(\alpha_2 = k) &= P(f + g = k) + P(f + g = 10 + k) = (k + 1) \frac{1}{10^2} + \\ &+ P(f + g = 10 + k) = (k + 1) \frac{1}{10^2} + \sum_{l=1}^{10-k-1} P(f=l, g=10+k-l) = \\ &= (k + 1) \frac{1}{10^2} + \sum_{l=1}^{10-k-1} P(f=l) \cdot P(g=10+k-l) = (k + 1) \frac{1}{10^2} + \\ &+ (10 - k - 1) \cdot \frac{1}{10^2} = \frac{1}{10} = 0,1. \end{aligned}$$

iar dacă $k = 9$ avem

$$P(\alpha_2 = 9) = P(f + g = 9) = 1 - \sum_{k=0}^8 P(f + g = k) = 1 - 0,9 = 0,1.$$

Variabilele aleatoare α_1, α_2 sînt dependente pentru că de exemplu

$$P(\alpha_1 = 0, \alpha_2 = 9) = P(f + g = 9) = 0,1 \neq P(\alpha_1 = 0) P(\alpha_2 = 9) = 0,45 \cdot 0,1.$$

2.12. Variabilele aleatoare f, g sînt repartizate Poisson cu parametrii λ, μ . Dacă $\lambda > \mu$ să se arate că $P(f \leq n) < P(g \leq n)$ pentru orice întreg n .

Soluție. Fie $(\Omega, \mathcal{F}, \mathcal{Q})$ cîmpul de probabilitate pe care sînt definite f, g și fie cîmpul de probabilitate $(N, \mathcal{Q}(N), \pi_{\lambda-\mu})$

Formăm cîmpul produs $(\Omega \times N, \mathcal{F} \otimes \mathcal{Q}(N), \mathcal{Q} \otimes \pi_{\lambda-\mu})$ și definim variabilele aleatoare $\tilde{f}, \tilde{g}, \tilde{h}$ prin

$$\tilde{f}(\omega, n) = f(\omega), \quad \tilde{g}(\omega, n) = g(\omega), \quad \tilde{h}(\omega, n) = n.$$

Atunci $\tilde{f}, \tilde{g}, \tilde{h}$ au repartițiile $\pi_\lambda, \pi_\mu, \pi_{\lambda-\mu}$ și în plus \tilde{h} este independentă de \tilde{f}, \tilde{g} .

De asemenea $\tilde{h} + \tilde{g}$ are repartiția $\pi_{\lambda-\mu+\mu} = \pi_{\lambda}$.
Prin urmare

$$P(f \leq n) = (P \otimes \pi_{\lambda-\mu})(\tilde{h} + \tilde{g} \leq n) = \sum_{m=0}^n (P \otimes \pi_{\lambda-\mu})(\tilde{h} = m)P(g \leq n-m) \leq (P \otimes \pi_{\lambda-\mu})(\tilde{h} \leq n)P(g \leq n) < P(g \leq n).$$

2.13. Variabilele aleatoare A, B sînt uniform repartizate pe $(-1, 1)$. Să se calculeze probabilitatea că rădăcinile ecuației $x^2 + Ax + B = 0$ să fie reale.

Soluție. Fie $p(x)$ densitatea lui A, B deci $p(x) = \frac{1}{2} \lambda_{(-1, 1)}(x)$.

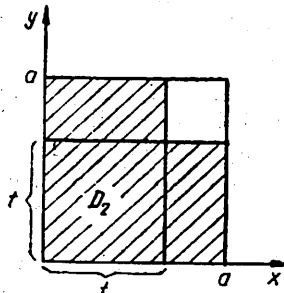
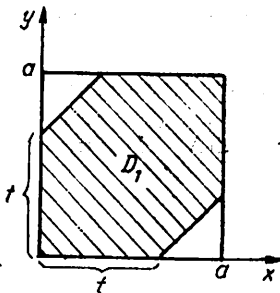
Probabilitatea căutată este egală cu

$$\begin{aligned} P(A^2 - 4B \geq 0) &= \int_{\substack{y^2 - 4x \geq 0 \\ -1 \leq x, y \leq 1}} p(x)p(y) dx dy = \frac{1}{4} \int_{\substack{y^2 - 4x \geq 0 \\ -1 \leq x, y \leq 1}} dx dy = \\ &= \frac{1}{4} \int_{-1}^0 dx \int_{-1}^1 dy + \frac{1}{4} \int_0^1 dx \left(\int_{2\sqrt{x}}^1 dy + \int_{-1}^{-2\sqrt{x}} dy \right) = \\ &= \frac{1}{2} + \frac{1}{2} \int_0^1 (1 - 2\sqrt{x}) dx = \frac{1}{3} \end{aligned}$$

2.14. Vectorul (f, g) are repartiția uniformă în pătratul $(0, a) \times (0, a)$. Să se arate că variabilele aleatoare $|f - g|$ și $\min(f, g)$ au aceeași repartiție.

Soluție. Este suficient să arătăm că $P(|f - g| \leq t) = P(\min(f, g) \leq t)$ pentru orice $0 \leq t \leq a$.

Fie $D_i, i = 1, 2$, domeniile hașurate din figurile alăturate.



Atunci $P(|f - g| \leq t) = \frac{1}{a^2}$ aria $D_1 = \frac{a^2 - (a - t)^2}{a^2}$,

$$P(\min(f, g) \leq t) = \frac{1}{a^2} \text{ aria } D_2 = \frac{a^2 - (a - t)^2}{a^2}.$$

Se poate proceda și astfel. Se observă că

$$P(|f - g| \leq t) = \frac{1}{a^2} \int_{\substack{|x-y| \leq t \\ 0 \leq x, y \leq a}} dx dy.$$

$$P(\min(f, g) \leq t) = \frac{1}{a^2} \int_{\substack{\min(x, y) \leq t \\ 0 \leq x, y \leq a}} dx dy.$$

Se calculează apoi integralele și se observă egalitatea lor.

2.15. Variabilele aleatoare f și g sînt independente. Să se calculeze densitatea cîtului dacă :

- 1) f și g au repartiția uniformă pe $(0, 1)$;
- 2) f și g au repartiția e_1 .

Soluție. Putem aplica rezultatul din problema 2.6.

a) Fie $p(x)$ densitatea lui f și g , deci $p(x) = \lambda_{(0, 1)}(x)$. Atunci densitatea vectorului (f, g) este $p(x)p(y)$ și deci densitatea lui $\frac{f}{g}$ este pentru $x > 0$:

$$p_{f/g}(x) = \int_0^1 zp(xz)p(z)dz = \begin{cases} \int_0^1 zdz & \text{dacă } x < 1 \\ \int_{1/x}^1 zdz & \text{dacă } x \geq 1 \end{cases}$$

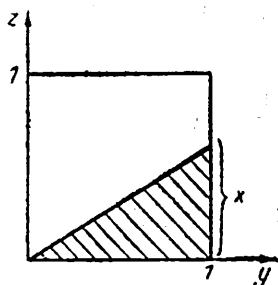
$$p_{f/g}(x) = \begin{cases} 1/2 & \text{dacă } x < 1 \\ 1/2x^2 & \text{dacă } x \geq 1 \end{cases}$$

Putem proceda și astfel. Vectorul (f, g) are ca repartiție produsul repartițiilor lui f și g , cu alte cuvinte tocmai măsura Lebesgue (aria

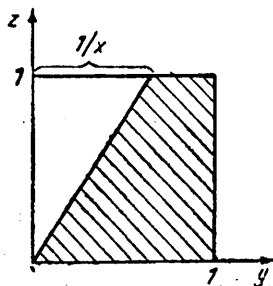
pe pătratul $(0,1) \times (0,1)$. Rezultă că funcția de repartiție a cîtului $\frac{f}{g}$ este

$$F_{f/g}(x) = \text{aria } A,$$

unde A este unul din domeniile hașurate din figurile următoare:



Pentru $x < 1$.



Pentru $x \geq 1$.

Deci:

$$F_{f/g}(x) = \begin{cases} \frac{x}{2} & \text{dacă } x < 1 \\ 1 - \frac{1}{2x} & \text{dacă } x \geq 1 \end{cases}$$

Prin derivare în raport cu x se obține expresia deja aflată mai sus.

b) Fie $q(x)$ densitatea lui f și g , deci $q(x) = e^{-x} \lambda_{R_+}(x)$. Densitatea vectorului (f, g) este $q(x)q(y)$, deci densitatea cîtului f/g este

$$\rho_{f/g}(x) = \int_0^{\infty} z e^{-xz} e^{-z} dz = \int_0^{\infty} z e^{-z(1+x)} dz = \frac{1}{(1+x)^2} \text{ dacă } x \geq 0.$$

2.16. Fie f_1, \dots, f_n variabile aleatoare independente cu repartiția uniformă pe $[0, a]$. Să se arate că densitatea de repartiție ρ_n și funcția de repartiție F_n a variabilei aleatoare $S_n = f_1 + \dots + f_n$ sînt date pentru $n \geq 2$ de egalitățile:

$$(1) \quad \rho_n(x) = \begin{cases} 0 & \text{dacă } x \leq 0 \\ \frac{1}{a^n(n-1)!} \sum_{k=0}^{n-1} (-1)^k C_n^k (x - ka)_+^{n-1} & \text{dacă } x > 0 \end{cases}$$

$$F_n(x) = \begin{cases} 0 & \text{dacă } x \leq 0 \\ \frac{1}{a^n n!} \sum_{k=0}^n (-1)^k C_n^k (x - ka)_+^n & \text{dacă } x > 0 \end{cases}$$

unde $x_+ = \max(x, 0)$.

Soluție. Vom folosi inducția după n . Presupunem $n = 2$ și să notăm cu $\rho(x)$ densitatea de repartiție a lui f_1 , deci $\rho(x) = \frac{1}{a}$ pentru $x \in (0, a)$ și $\rho(x) = 0$ pentru $x \notin (0, a)$. Avem

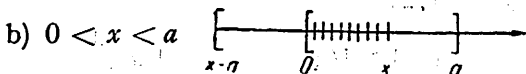
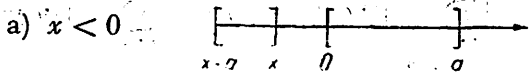
$$\rho_2(x) = \int_0^a \rho(x-y)\rho(y)dy = \frac{1}{a} \int_0^a \rho(x-y)dy = \frac{1}{a} \int_{x-a}^x \rho(z)dz.$$

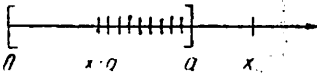
Figura alăturată arată că

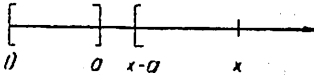
$$\rho_2(x) = \begin{cases} 0 & x \leq 0 \text{ sau } x > 2a \\ \frac{1}{a} \int_0^x \rho(z)dz & 0 < x \leq a \\ \frac{1}{a} \int_{x-a}^a \rho(z)dz & a \leq x \leq 2a \end{cases}$$

de unde rezultă că

$$\rho_2(x) = \begin{cases} 0 & x \leq 0 \text{ sau } x > 2a \\ \frac{x}{a^2} & 0 < x \leq a \\ \frac{2a-x}{a^2} & a < x \leq 2a \end{cases}$$



c) $a < x < 2a$ 

d) $x > 2a$ 

Nu este greu de văzut că $\rho_2(x) = \frac{1}{a^2} [x_+ - 2(x-a)_+]_+$, fapt ce justifică egalitatea (1) pentru $n = 2$.

Întrucît pentru $x > 0$

$$\int_0^x (z - ka)_+^{n-1} dz = \frac{1}{n} (x - ka)_+^n \quad (3)$$

$$F_n(x) = \int_0^x \rho_n(z) dz, \quad (4)$$

se observă că rezultă și egalitatea (2) pentru $n = 2$. De fapt din (3) și (4) rezultă valabilitatea egalității (2) pentru orice n de îndată ce am demonstrat egalitatea (1).

Presupunem acum că (1) este adevărată pentru n și să o arătăm pentru $n + 1$. Este vizibil că putem presupune $x > 0$. Avem

$$\begin{aligned} \rho_{n+1}(x) &= \int \rho_n(x-y)\rho(y)dy = \frac{1}{a} \int_0^a \rho_n(x-y)dy = \frac{1}{a} [F_n(x) - \\ &\quad - F_n(x-a)]. \end{aligned}$$

Folosind egalitatea (2) pentru F_n , obținem pentru ρ_{n+1} următoarea reprezentare:

$$\begin{aligned} \rho_{n+1}(x) &= \frac{1}{a^{n+1}n!} \left\{ \sum_{k=0}^n (-1)^k C_n^k (x - ka)_+^n + \sum_{k=0}^n (-1)^{k+1} C_n^k [x - (k + \right. \\ &\quad \left. + 1)a]_+^n \right\} = \frac{1}{a^{n+1}n!} \left\{ \sum_{k=0}^n (-1)^k C_n^k (x - ka)_+^n + \sum_{k=1}^{n+1} (-1)^k C_n^{k-1} (x - ka)_+^n \right\} = \\ &= \frac{1}{a^{n+1}n!} \left\{ \sum_{k=0}^n (-1)^k [C_n^k + C_n^{k-1}] (x - ka)_+^n + \right. \\ &\quad \left. + (-1)^{n+1} C_n^n [x - (n+1)a]_+^n \right\}, \end{aligned}$$

de unde, avînd în vedere că $C_n^k + C_n^{k-1} = C_{n+1}^k$, obținem egalitatea (1) pentru $n + 1$.

2.17. Fie F o repartiție pe R concentrată pe $(0, \infty)$. Să se arate că F este repartiție exponențială dacă și numai dacă este satisfăcută condiția

$$\tilde{F}(x + y) - \tilde{F}(y) = \tilde{F}(x) [1 - \tilde{F}(y)] \text{ pentru orice } x, y > 0. \quad (1)$$

Soluție. Să presupunem că $F = e_\lambda$. Este vizibil că e_λ este continuă și concentrată pe R_+ . Pe deasupra $e_\lambda(x) = 1 - \exp(-\lambda x)$ dacă $x > 0$.

Apoi

$$\begin{aligned} \tilde{e}_\lambda(x + y) - \tilde{e}_\lambda(y) &= 1 - \exp\{-\lambda(x + y)\} - 1 + \exp(-\lambda y) = \\ &= \exp(-\lambda y) - \exp\{-\lambda(x + y)\} = \exp(-\lambda y) [1 - \exp(-\lambda x)] = \\ &= [1 - \tilde{e}_\lambda(x)] \tilde{e}_\lambda(y), \end{aligned}$$

dacă $x, y > 0$.

Pentru a demonstra afirmația inversă, fie $G(x) = 1 - \tilde{F}(x)$. Relația (1) se mai scrie sub forma

$$G(x + y) = G(x)G(y) \text{ pentru orice } x, y > 0. \quad (2)$$

Observăm că $G(x) > 0$ pentru orice $x > 0$. Prin absurd fie $x_1 > 0$ așa încît $G(x_1) = 0$.

Din (2) rezultă că $G(x_1) = \left[G\left(\frac{x_1}{n}\right)\right]^n$ pentru orice n , deci $G\left(\frac{x_1}{n}\right) = 0$ pentru orice n , de unde făcînd $n \rightarrow \infty$ obținem $0 = \lim_{n \rightarrow \infty} G\left(\frac{x_1}{n}\right) = F(0, \infty) = 1$, ceea ce nu se poate.

Din (2) rezultă că pentru orice numere naturale m, n avem

$$G(m/n) = [G(1)]^{m/n}$$

și cum G este continuă la stînga, rezultă egalitatea

$$G(x) = [G(1)]^x \text{ pentru orice } x > 0. \quad (3)$$

În continuare, întrucît $\lim_{x \rightarrow \infty} G(x) = 1 - \lim_{x \rightarrow \infty} \tilde{F}(x) = 1 - 1 = 0$ rezultă că $G(1) < 1$.

Dacă punem $\lambda = -\ln G(1)$, atunci din $G(1) < 1$ rezultă $\lambda > 0$, așa că (3) se scrie

$$G(x) = \exp(-\lambda x) \quad \text{pentru orice } x > 0$$

sau ceea ce este totuna

$$\tilde{F}(x) = 1 - \exp(-\lambda x) \quad \text{pentru } x > 0.$$

Or această egalitate nu spune altceva decât că $F = e_\lambda$ (funcția de repartiție determină în mod unic repartiția).

2.18. Fie f și g variabile aleatoare nenegative cu densitatea p mărginită și continuă în sens topologic și $p(x) > 0$ pentru $x > 0$.

Dacă sînt îndeplinite condițiile :

a) f și g independente ;

b) $f - g$ și $\min(f, g)$ sînt independente,

să se arate că f și g au repartiție exponențială.

Soluție. Fie $U = f - g$ și $V = \min(f, g)$. Densitatea vectorului este în virtutea independenței egală cu $p(x)p(y)$.

Din problema 2.6 rezultă că U are densitatea

$$p_U(u) = \int_0^{\infty} p(u+x)p(x)dx$$

iar din problema 2.7 obținem că V are densitatea

$$p_V(v) = 2p(v) \int_v^{\infty} p(x) dx = 2p(v) [1 - F_f(v)].$$

Variabilele aleatoare U, V fiind independente rezultă că vectorul (U, V) are densitatea $p(u, v)$ dată de

$$p(u, v) = p_U(u)p_V(v) = 2 \int_0^{\infty} p(u+x)p(x)dx p(v) [1 - F_f(v)]. \quad (1)$$

Pe de altă parte funcția de repartiție a lui (U, V) este dacă $u < 0$

$$P(U < u, V < v) = \int_{\substack{x-y < u \\ \min(x, y) < v \\ 0 < x, y < \infty}} p(x)p(y)dx dy = \int_0^v p(x)dx \int_{x-u}^{\infty} p(y)dy.$$

Luind derivatele în raport cu u și v , obținem că densitatea lui (U, V) pentru $u < 0$ și $v > 0$ este

$$p(v)p(v - u) = p(v)p(v + |u|). \quad (2)$$

Analog se obține pentru $u \geq 0$

$$P(U < u, V < v) = \int_0^v p(y)dy \int_{-\infty}^{y+u} p(x)dx,$$

de unde prin derivare obținem că densitatea lui (U, V) este pentru $u < 0, v > 0$ dată tot de (2).

Egalind (1) și (2), obținem

$$p(v + |u|) = 2 \int_0^{\infty} p(u + x)p(x)dx [1 - F_f(v)]$$

(deoarece $p(v) > 0$ dacă $v > 0$).

Făcînd $u \rightarrow 0$ și avînd în vedere continuitatea lui p și faptul că integrandul este majorat de $\sup_y p(y) \cdot p(x)$, deci de o funcție integrabilă (prin urmare teorema de convergență dominată funcționează) obținem

$$F'_f(v) = p(v) = 2 \int_0^{\infty} p^2(x)dx [1 - F_f(v)] = \lambda [1 - F_f(v)], \lambda > 0.$$

Soluția acestei ecuații diferențiale este $F_f(v) = 1 - \exp(-\lambda v)$, deci f are repartiția exponențială de parametru λ .

2.19. Fie f_1, \dots, f_n, \dots variabile aleatoare independente cu repartiția e_λ . Să se arate că suma $f_1 + \dots + f_n$ are densitatea

$$p_n(x) = \frac{\lambda(\lambda x)^{n-1}}{(n-1)!} e^{-\lambda x} \lambda_{R^+}(x).$$

Să se deducă apoi că $f_1 + \dots + f_n + \dots = \infty$ a.s.

Soluție. Vom raționa prin inducție după n . Afirmația este adevărată pentru $n = 1$ și să o presupunem adevărată pentru n .

Fie $S_n = f_1 + \dots + f_n$. Cum S_n și f_{n+1} sînt independente (asociativitatea independenței), rezultă că S_{n+1} are densitatea

$$p_{n+1}(x) = \int_0^{\infty} p_n(y) p_1(x-y) dy$$

de unde

$$p_{n+1}(x) = \begin{cases} \frac{\lambda^{n+1} x^n}{n!} e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

Mai departe, dacă $t > 0$, avem

$$\begin{aligned} P(f_1 + \dots + f_n + \dots \leq t) &= \lim_{n \rightarrow \infty} P(f_1 + \dots + f_n \leq t) = \\ &= \lim_{n \rightarrow \infty} \int_0^t \frac{\lambda(\lambda x)^{n-1}}{(n-1)!} e^{-\lambda x} dx = \int_0^t \lim_{n \rightarrow \infty} \frac{\lambda(\lambda x)^{n-1}}{(n-1)!} e^{-\lambda x} dx = 0 \end{aligned}$$

șirul de sub integrală este majorat de 1, deci teorema de convergență dominată funcționează).

În fine

$$P(f_1 + f_n + \dots = \infty) = \lim_{t \rightarrow \infty} P(f_1 + \dots + f_n + \dots \leq t) = 0$$

cu alte cuvinte $f_1 + \dots + f_n + \dots = \infty$ a.s.

2.20. Variabilele aleatoare g_1, \dots, g_n sînt independente și au repartiția uniformă pe $(0, 1)$.

Să se afle densitatea de repartiție a produsului $P_n = g_1 \dots g_n$.

Soluție. Fie $f_i = -\ln g_i$. Atunci f_i sînt bine definite deoarece

$$P(g_i \in (0, 1)) = 1 \text{ și } P(f_i \in (0, \infty)) = 1.$$

Pentru $0 < x < \infty$ avem $P(f_i < x) = P(g_i < e^{-x}) = 1 - e^{-x}$, deci f_i are repartiția e_1 .

Cum f_1, \dots, f_n sînt independente (asociativitatea independenței) din problema precedentă, rezultă că $f_1 + \dots + f_n$ are densitatea p_n din enunțul aceleiași probleme.

Cum $P(f_n \in (0,1)) = 1$, rezultă că dacă $0 < x < 1$ avem

$$P(P_n < x) = P(f_1 + \dots + f_n > -\ln x) = 1 - \int_0^{-\ln x} \frac{t^{n-1} e^{-t}}{(n-1)!} dt =$$

$$= \frac{(-1)^{n-1}}{(n-1)!} \int_0^x (\ln y)^{n-1} dy.$$

Derivînd în raport cu x , obținem că P_n are densitatea

$$\frac{(-1)^{n-1}}{(n-1)!} (\ln x)^{n-1} \quad \text{dacă } 0 < x < 1.$$

2.21. Fie f_1, \dots, f_n, \dots un șir de variabile aleatoare independente, fiecare avînd repartiția exponențială de parametru λ și fie

$$\tau = \min(n \geq 1; f_1 + \dots + f_n > 1).$$

Să se afle repartițiile variabilelor aleatoare τ și f_τ .

Soluție. Din problema 2.19 se știe că $f_1 + \dots + f_n + \dots = \infty$ a.s. deci τ definită mai sus este finită a.s. și de asemenea $f_1 + \dots + f_n$ are densitatea

$$\frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x} \lambda_{R^+}(x) = f_{n-1}(x).$$

Are loc egalitatea: $\{\tau = n\} = \{f_1 + \dots + f_{n-1} \leq 1\} \setminus \{f_1 + \dots + f_n \leq 1\}$ și cum $\{f_1 + \dots + f_n \leq 1\} \subset \{f_1 + \dots + f_{n-1} \leq 1\}$ rezultă

$$P(\tau = n) = P(f_1 + \dots + f_{n-1} \leq 1) - P(f_1 + \dots + f_n \leq 1) =$$

$$= \frac{\lambda^{n-1}}{(n-2)!} \int_0^1 x^{n-2} e^{-\lambda x} dx - \frac{\lambda}{(n-1)!} \int_0^1 x^{n-1} e^{-\lambda x} dx = \frac{\lambda^{n-1}}{(n-1)!} e^{-\lambda x}.$$

Cu alte cuvinte τ are repartiție Poisson de parametru λ .

Mai departe

$$P(f_\tau < x) = \sum_{n=1}^{\infty} P(f_\tau < x, \tau = n) = P(1 < f_1 < x) +$$

$$+ \sum_{n=2}^{\infty} P\left(f_n < x, \sum_{i=1}^{n-1} f_i \leq 1 < \sum_{i=1}^n f_i\right).$$

Fie $\eta = f_1 + \dots + f_{n-1}$.

Cazul 1. $x \leq 1$. Avem $P(1 < f_1 < x) = 0$ și $P(f_n < x, \eta \leq 1, \eta + f_n > 1) = \int_{1-x}^1 \rho_{n-1}(u) \int_{1-u}^x \lambda e^{-\lambda v} dv du = \int_{1-x}^1 \rho_{n-1}(u) [e^{-\lambda(1-x)} - e^{-\lambda u}] du$ și cum

$$\sum_{n=2}^{\infty} \rho_{n-1}(x) = \lambda \text{ dacă } x \geq 0,$$

obținem

$$P(f_{\tau} < x) = \int_{1-x}^1 \lambda [e^{-\lambda(1-u)} - e^{-\lambda x}] du = 1 - e^{-\lambda x} - \lambda x e^{-\lambda x}.$$

Cazul 2. $x > 1$. Avem $P(1 < f_1 < x) = e^{-\lambda} - e^{-\lambda x}$ și $P(f_n < x, \eta \leq 1, \eta + f_n > 1) = \int_0^1 \rho_{n-1}(u) [e^{-\lambda(1-u)} - e^{-\lambda x}] dx$. Prin urmare

$$P(f_{\tau} < x) = e^{-\lambda} - e^{-\lambda x} + \int_0^1 \lambda [e^{-\lambda(1-u)} - e^{-\lambda x}] du = 1 - e^{-\lambda x} - \lambda e^{-\lambda x}.$$

În concluzie obținem că

$$P(f_{\tau} < x) = \begin{cases} 1 - e^{-\lambda x} - \lambda x e^{-\lambda x}, & x \leq 1, \\ 1 - e^{-\lambda x} - \lambda e^{-\lambda x}, & x > 1. \end{cases}$$

2.22. Dacă variabilele aleatoare f și g sînt independente și au repartiția $N(0, \sigma^2)$, să se arate că și $f + g, f - g$ sînt independente.

Soluție. Densitatea vectorului (f, g) este

$$\frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right) \exp\left(-\frac{y^2}{2\sigma^2}\right) = p(x, y).$$

Fie transformarea w definită de $y_1 = x_1 + x_2, y_2 = x_1 - x_2$. Inversa lui w este

$$x_1 = \frac{y_1 + y_2}{2}; \quad x_2 = \frac{y_1 - y_2}{2}$$

și are jacobianul $J = -\frac{1}{2}$.

$$\begin{aligned}
 & \text{Atunci vectorul } (f + g, f - g) \text{ are densitatea } \frac{1}{2} p\left(\frac{y_1 + y_2}{2}, \frac{y_1 - y_2}{2}\right) = \\
 & = \frac{1}{4\pi\sigma^2} \exp\left\{-\frac{\left(\frac{y_1 + y_2}{2}\right)^2}{2\sigma^2}\right\} \exp\left\{-\frac{\left(\frac{y_1 - y_2}{2}\right)^2}{2\sigma^2}\right\} = \frac{1}{4\pi\sigma^2} \exp\left(-\frac{y_1^2 + y_2^2}{2\sigma^2}\right) = \\
 & = \frac{1}{2\sigma\sqrt{\pi}} \exp\left(-\frac{y_1^2}{2\sigma^2}\right) \cdot \frac{1}{2\sigma\sqrt{\pi}} \exp\left(-\frac{y_2^2}{2\sigma^2}\right) = q(y_1)q(y_2), \quad \text{unde } q(x) = \\
 & = \frac{1}{2\sigma\sqrt{\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right) \text{ este densitatea repartiției } N(0, 2\sigma^2).
 \end{aligned}$$

Din problema 2.4 rezultă că $f + g$ și $f - g$ au repartiția $N(0, 2\sigma^2)$ și că sînt independente.

2.23. Fie f, g, h variabile aleatoare independente cu repartiția $N(0, 1)$. Să se arate că variabila aleatoare $\theta = \frac{f + gh}{\sqrt{1 + h^2}}$ are tot repartiția $N(0, 1)$.

Soluție. Din ipoteza de independență rezultă că vectorul (f, g, h) are densitatea

$$p(x, y, z) = \frac{1}{(\sqrt{2\pi})^3} \exp\left(-\frac{x^2 + y^2 + z^2}{2}\right).$$

Fie $D_t = \left\{ (x, y, z); \frac{x + yz}{\sqrt{1 + z^2}} < t \right\}$. Funcția de repartiție a lui θ va fi

$$\begin{aligned}
 F_\theta(t) & = P\left(\frac{f + gh}{\sqrt{1 + h^2}} < t\right) = P((f, g, h) \in D_t) = \\
 & = \frac{1}{(\sqrt{2\pi})^3} \int_{D_t} \exp\left(-\frac{x^2 + y^2 + z^2}{2}\right) dx dy dz = \\
 & = \frac{1}{(\sqrt{2\pi})^3} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{t\sqrt{1+z^2}-yz} \exp\left(-\frac{x^2 + y^2 + z^2}{2}\right) dx dy dz = \\
 & = \frac{1}{(\sqrt{2\pi})^3} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2}\left[\left(r\sqrt{1+z^2}-yz\right)^2 + y^2 + z^2\right]\right\} \sqrt{1+z^2} dy dz dr =
 \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{(\sqrt{2\pi})^3} \int_{-\infty}^t \int_{-\infty}^{+\infty} \sqrt{1+z^2} \exp\left(-\frac{r^2+z^2}{2}\right) dz dr \int_{-\infty}^{+\infty} \exp\left\{-\frac{(y\sqrt{1+z^2}-zr)^2}{2}\right\} dy = \\
&= \frac{1}{(\sqrt{2\pi})^3} \int_{-\infty}^t \int_{-\infty}^{+\infty} \exp\left(-\frac{r^2+z^2}{2}\right) dz dr \int_{-\infty}^{+\infty} \exp\left(-\frac{u^2}{2}\right) du = \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t \exp\left(-\frac{r^2}{2}\right) dr,
\end{aligned}$$

deci are repartiția $N(0, 1)$ (am folosit mai sus egalitatea

$$\int_{-\infty}^{+\infty} \exp\left(-\frac{x^2}{2}\right) dx = \sqrt{2\pi}.$$

2.24. Fie f_1, \dots, f_n variabilele aleatoare independente cu repartiția $N(0, 1)$. Să se arate că variabila aleatoare $g = f_1^2 + \dots + f_n^2$ are densitatea

$$p(x) = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} \lambda_{R_+(x)} \quad (1)$$

(repartiția cu densitatea de mai sus se numește repartiția χ^2 cu n grade de libertate).

Să se deducă volumul și aria sferei cu ajutorul acestui rezultat.

Soluția 1. Repartiția lui f_1^2 este $P(f_1^2 < x) = P(|f_1| < \sqrt{x}) = \frac{2}{\sqrt{2\pi}} \int_0^{\sqrt{x}} e^{-\frac{u^2}{2}} du$, $x > 0$, de unde prin derivare obținem că densitatea lui f_1^2 este

$$p(x) = \frac{1}{\sqrt{2\pi}} x^{\frac{1}{2}-1} e^{-\frac{x}{2}} \lambda_{R_+}(x),$$

cu alte cuvinte f_1^2 are repartiția gama de parametri $\left(\frac{1}{2}, \frac{1}{2}\right)$.

Problema 2.9 implică faptul că g are repartiția gama [de parametri $(\frac{n}{2}, \frac{1}{2})$, cu alte cuvinte g are densitatea dată de (1).

Soluția 2. Din independență rezultă că vectorul (f_1, \dots, f_n) are densitatea

$$q(x) = \frac{1}{(\sqrt{2\pi})^n} \exp\left(-\frac{x_1^2 + \dots + x_n^2}{2}\right).$$

Vom calcula mai întâi funcția de repartiție a variabilei $h = g/\sqrt{n}$. Aceasta este nulă pentru $y < 0$ iar pentru $y \geq 0$ este

$$F_h(y^2) = \int_{\sum_{i=1}^n x_i^2 \leq ny^2} \frac{1}{(\sqrt{2\pi})^n} \exp\left(-\frac{x_1^2 + \dots + x_n^2}{2}\right) dx_1 \dots dx_n.$$

Folosind schimbarea de variabilă

$$x_1 = r \cos \theta_1 \dots \cos \theta_{n-1}$$

$$x_2 = r \cos \theta_1 \dots \sin \theta_{n-1}$$

$$\dots \dots \dots$$

$$x_n = r \sin \theta_1$$

obținem

$$\begin{aligned} F_h(y^2) &= \int_{-\pi/2}^{\pi/2} \dots \int_{-\pi/2}^{\pi/2} \int_0^{y\sqrt{n}} \frac{1}{(\sqrt{2\pi})^n} \exp(-r^2/2) r^{n-1} J(\theta_1, \dots, \theta_{n-1}) d\theta_i = \\ &= C_n \int_0^{y\sqrt{n}} \exp(-r^2/2) r^{n-1} dr. \end{aligned}$$

Constanta C_n rezultă din condiția $\lim_{y \rightarrow \infty} F_h(y) = 1$, care se scrie

$$1 = C_n \int_0^{\infty} \exp(-r^2/2) r^{n-1} dr = C_n 2^{\frac{n}{2}-1} \Gamma(n/2), \text{ deci } C_n = \frac{1}{2^{\frac{n}{2}-1} \Gamma(n/2)}.$$

$$\text{Atunci } F_h(y^2) = \frac{1}{2^{\frac{n}{2}-1} \Gamma(n/2)} \int_0^{y\sqrt{n}} \exp(-r^2/2) r^{n-1} dr.$$

Funcția de repartiție a lui g se obține făcând în $F_h(y)$ substituția $y = x/\sqrt{n}$, deci

$$\begin{aligned} F_g(x) &= \frac{1}{2^{\frac{n}{2}-1} \Gamma(n/2)} \int_0^x r^{n-1} \exp(-r^2/2) dr = \\ &= \frac{1}{2^{n/2} \Gamma(n/2)} \int_0^{x^2} t^{\frac{n}{2}-1} \exp(-t/2) dt, \end{aligned}$$

de unde prin derivare se obține formula (1).

Să se determine acum volumul și aria sferei n -dimensionale. Fie $V(r)$ (respectiv $S(r)$) volumul (respectiv aria) sferei n -dimensionale cu centrul în origine și de rază r (să notăm cu $B(r)$ această sferă).

Au loc egalitățile

$$\begin{aligned} P(f_1^2 + \dots + f_n^2 \leq r^2) &= \int_0^{r^2} p(x) dx = \int_0^{r^2} \frac{x^{\frac{n}{2}-1} e^{-\frac{x}{2}}}{\Gamma\left(\frac{n}{2}\right) 2^{\frac{n}{2}}} dx = \\ &= \frac{r^n e^{-\frac{r^2}{2}}}{\Gamma\left(\frac{n}{2}\right) n 2^{\frac{n}{2}-1}} + \int_0^{r^2} \frac{x^{\frac{n}{2}-1} e^{-\frac{x}{2}}}{\Gamma\left(\frac{n}{2}\right) n 2^{\frac{n}{2}}} dx \end{aligned} \quad (2)$$

(integrând prin părți)

$$P(f_1^2 + \dots + f_n^2 \leq r^2) = P((f_1, \dots, f_n) \in B(r)) = \int_{B(r)} q(x) dx. \quad (3)$$

Avînd în vedere că $V(r) = r^n V(1)$ și teorema valorii medii vom obține

$$\begin{aligned} \lim_{r \rightarrow 0} \frac{P(f_1^2 + \dots + f_n^2 \leq r^2)}{V(r)} &= \lim_{r \rightarrow 0} \frac{P(f_1^2 + \dots + f_n^2 \leq r^2)}{r^n V(1)} = \\ &= \lim_{r \rightarrow 0} \frac{e^{-\frac{r^2}{2}}}{\Gamma\left(\frac{n}{2}\right) n 2^{\frac{n}{2}-1} V(1)} + \lim_{r \rightarrow 0} \frac{1}{r^n V(1)} \int_0^{r^2} \frac{x^{\frac{n}{2}-1} e^{-\frac{x}{2}}}{\Gamma\left(\frac{n}{2}\right) n 2^{\frac{n}{2}}} dx = \\ &= \left[\Gamma\left(\frac{n}{2}\right) n 2^{\frac{n}{2}-1} V(1) \right]^{-1}. \end{aligned}$$

Pe de altă parte

$$\lim_{r \rightarrow 0} \frac{P(f_1^2 + \dots + f_n^2 < r^2)}{V(r)} = \lim_{r \rightarrow 0} \frac{\int_{B(r)} q(x) dx}{V(r)} = q(0) = (\sqrt{2\pi})^{-n}.$$

Rezultă egalitatea

$$\left[\Gamma\left(\frac{n}{2}\right) n 2^{\frac{n}{2}-1} V(1) \right]^{-1} = (\sqrt{2\pi})^{-n},$$

$$\text{deci } V(1) = \frac{2\pi^{\frac{n}{2}}}{n \Gamma\left(\frac{n}{2}\right)} \text{ și prin urmare } V(r) = \frac{2\pi^{\frac{n}{2}}}{n \Gamma\left(\frac{n}{2}\right)} r^n = \frac{\pi^{\frac{n}{2}} r^n}{\Gamma\left(\frac{n}{2} + 1\right)}.$$

$$\text{În fine } S(r) = \frac{nV(r)}{r} = \frac{2\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} r^{n-1}.$$

2.25. Să se arate că dacă f, g sînt independente și cu repartiția $N(0, 1)$, atunci $f_1 = f_2 + g_2, f^2 = \frac{f}{g}$ sînt independente.

Soluție. Fie $\bar{p}(x, y)$ densitatea vectorului (f_1, f_2) . Vom arăta că $\bar{p}(x, y) = p_1(x)p_2(y)$, unde p_i este densitatea lui f_i .

Din ipoteza de independență rezultă că vectorul (f, g) are densitatea $\frac{1}{2\pi} \exp\left(-\frac{x^2 + y^2}{2}\right)$.

Din problema 2.24 rezultă că f_1 are repartiția χ^2 cu 2 grade de libertate, deci f_1 are densitatea $p_1(x) = \frac{1}{2} \exp\left(-\frac{x}{2}\right) \lambda_{\mathbb{R}_+}(x)$.

Din problema 2.6 rezultă că f_2 are densitatea

$$p_2(y) = \int_{\mathbb{R}} |x| q(x) q(x+y) dx \left(\text{unde } q(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \right).$$

Deci

$$\begin{aligned} p_2(y) &= -\frac{1}{2\pi} \int_{-\infty}^0 x \exp\left(-\frac{x^2}{2}\right) \exp\left(-\frac{x^2+y^2}{2}\right) dx + \\ &+ \frac{1}{2\pi} \int_0^{\infty} x \exp\left(-\frac{x^2}{2}\right) \exp\left(-\frac{x^2+y^2}{2}\right) dx = \frac{1}{\pi(1+y^2)}. \end{aligned}$$

Prin urmare

$$p_1(x)p_2(y) = \begin{cases} \frac{1}{2\pi(1+y^2)} \exp\left(-\frac{x}{2}\right), & x > 0, y \in \mathbb{R}, \\ 0 & \text{în rest.} \end{cases}$$

Mai departe avem

$$\begin{aligned} F_{(f_1, f_2)}(x, y) &= P(f^2 + g^2 < x, \frac{f}{g} < y) = \int_{\substack{u^2+v^2 < x \\ \frac{u}{v} < y}} q(u)q(v)du dv = \\ &= \frac{1}{2\pi} \int_{\substack{u^2+v^2 < x \\ \frac{u}{v} < y}} \exp\left(-\frac{u^2+v^2}{2}\right) du dv \quad \text{dacă } x > 0, y \in \mathbb{R}, \end{aligned}$$

de unde trecînd la coordonate polare $u = r \cos \theta$, $v = r \sin \theta$ se obține

$$\begin{aligned} F_{(f_1, f_2)}(x, y) &= \frac{1}{2\pi} \int_0^{\sqrt{x}} \exp\left(-\frac{r^2}{2}\right) r dr \left(\int_{\arctg y}^{\pi} d\theta + \int_{\arctg y + \pi}^{2\pi} d\theta \right) = \\ &= \frac{1}{2\pi} \left[1 - \exp\left(-\frac{x}{2}\right) \right] (2\pi - 2 \arctg y). \end{aligned}$$

Derivînd în raport cu x, y , obținem că densitatea lui (f_1, f_2) este

$$\bar{p}(x, y) = \frac{\partial^2 F_{(f_1, f_2)}(x, y)}{\partial x \partial y} = \frac{1}{2\pi} \exp\left(-\frac{x}{2}\right) \frac{1}{(1+y^2)} = p_1(x) p_2(y)$$

dacă $x > 0, y \in \mathbb{R}$.

Egalitatea de demonstrat este imediată dacă $x < 0$.

2.26. Pentru orice întreg $n \geq 1$, fie $r_n: ([0, 1], \mathfrak{B}_{[0, 1]}, m) \rightarrow \mathbb{R}$ (m este măsura Lebesgue) variabila aleatoare definită prin $r_n(x) = f(2^{n-1}x)$, unde

$$f(x) = \begin{cases} 1, & 0 \leq x < \frac{1}{2}, \\ -1, & \frac{1}{2} \leq x < 1, \\ f(x+1) & \text{în rest.} \end{cases}$$

Să se arate că r_1, \dots, r_n, \dots sînt variabile aleatoare independente.

Soluție. Fie $A_n = \{r_n = 1\}$ și \mathfrak{A}_n algebra generată de $\{A_1, \dots, A_n\}$. Arătăm că \mathfrak{A}_n este de asemenea generată de partiția $\left\{ \left[\frac{k}{2^n}, \frac{k+1}{2^n} \right]; 0 \leq k < 2^n \right\}$. Vom demonstra aceasta prin inducție după n . Afirmarea este imediată pentru $n = 1$. Să presupunem afirmația adevărată pentru n și să o probăm pentru $n + 1$.

Deoarece $A_{n+1} = \bigcup_{k=0}^{2^n-1} \left[\frac{2k}{2^{n+1}}, \frac{2k+1}{2^{n+1}} \right]$ rezultă că \mathfrak{A}_{n+1} este inclus în algebra generată de partiția $\left\{ \left[\frac{k}{2^{n+1}}, \frac{k+1}{2^{n+1}} \right]; 0 \leq k < 2^{n+1} \right\}$. Incluziunea inversă rezultă din egalitățile

$$\begin{aligned} \left[\frac{k}{2^n}, \frac{k+1}{2^n} \right) \cap A_{n+1} &= \left[\frac{2k}{2^{n+1}}, \frac{2k+1}{2^{n+1}} \right]; \left[\frac{k}{2^n}, \frac{k+1}{2^n} \right) \cap \complement A_{n+1} = \\ &= \left[\frac{2k+1}{2^{n+1}}, \frac{2k+2}{2^{n+1}} \right). \end{aligned}$$

Arătăm acum că A_{n+1} și \mathfrak{A}_n sînt independente. Pentru aceasta este suficient să arătăm că dacă $B = \left[\frac{k}{2^n}, \frac{k+1}{2^n} \right)$ atunci

$$P(B \cap A_{n+1}) = P(B)P(A_{n+1}).$$

Avem

$P(B) = \frac{1}{2^n}$, $P(A_{n+1}) = \frac{1}{2}$ și $P(B \cap A_{n+1}) = \frac{1}{2^{n+1}}$, deci (1) este adevărată.

r_1, \dots, r_n sînt independente dacă și numai dacă

$$P(A_1 \cap \dots \cap A_n) = P(A_1) \dots P(A_n). \quad (2)$$

Vom demonstra relația (2) prin inducție după n . Pentru $n = 2$ afirmația rezultă din faptul că A_2 și \mathfrak{A}_1 sînt independente.

Apoi deoarece $A_1 \cap \dots \cap A_n \in \mathfrak{A}_n$ rezultă că

$$\begin{aligned} P(A_1 \cap \dots \cap A_n \cap A_{n+1}) &= P(A_1 \cap \dots \cap A_n) P(A_{n+1}) = \\ &= P(A_1) \dots P(A_n) P(A_{n+1}) \end{aligned}$$

(în prima egalitate am folosit faptul că A_{n+1} și \mathfrak{A}_n sînt independente.

§ 2. Momentele variabilelor aleatoare

2.27. Să se deducă expresiile mediilor și dispersiilor pentru repartițiile unidimensionale listate în cele două tabele din rezultatul teoretic iar pentru repartițiile multidimensionale din rezumatul teoretic să se deducă expresiile mediilor și matricilor de corelație.

Soluție.

Repartiția Dirac. Folosind formula generală $\int f d\varepsilon_a = f(a)$ se obține imediat că $M(\varepsilon_a) = a$ și $D^2(\varepsilon_a) = 0$.

Repartiția uniformă pe mulțimea $\{1, \dots, n\}$.

$$M\left(\sum_{k=1}^n \frac{1}{n} \varepsilon_k\right) = \sum_{k=1}^n \frac{1}{n} M(\varepsilon_k) = \frac{1}{n} \sum_{k=1}^n k = \frac{n+1}{2}$$

$$M_2\left(\sum_{k=1}^n \frac{1}{n} \varepsilon_k\right) = \int x^2 d\left(\sum_{k=1}^n \frac{1}{n} \varepsilon_k\right)(x) = \frac{1}{n} \sum_{k=1}^n \int x^2 d\varepsilon_k(x) =$$

$$= \frac{1}{n} \sum_{k=1}^n k^2 = \frac{(n+1)(2n+1)}{6}$$

deci

$$D^2\left(\sum_{k=1}^n \frac{1}{n} \varepsilon_k\right) = \frac{(n+1)(2n+1)}{6} - \frac{(n-1)^2}{4} = \frac{n^2-1}{12}$$

Repartiția geometrică de parametru p .

$$M\left(\sum_{k=0}^{\infty} (1-p)p^k \varepsilon_k\right) = (1-p) \sum_{k=0}^{\infty} p^k M(\varepsilon_k) = (1-p)p \sum_{k=0}^{\infty} k p^{k-1} =$$

$$= (1-p)p \frac{1}{(1-p)^2} = \frac{p}{1-p},$$

$$M_2\left(\sum_{k=0}^{\infty} (1-p)p^k \varepsilon_k\right) = (1-p) \sum_{k=0}^{\infty} k^2 p^k = (1-p)p^2 \sum_{k=0}^{\infty} k^2 p^{k-2} =$$

$$= (1-p)p^2 \sum_{k=1}^{\infty} k(k-1)p^{k-2} + (1-p) \sum_{k=1}^{\infty} k p^k = (1-p)p^2 \left(\frac{1}{1-p}\right) +$$

$$+ \frac{p}{1-p} = (1-p)p^2 \frac{2}{(1-p)^2} + \frac{p}{1-p} = \frac{p(p+1)}{(1-p)^2}$$

deci

$$D^2 \left(\sum_{k=0}^{\infty} (1-p)p^k \varepsilon_k \right) = \frac{p(p+1)}{(1-p)^2} - \frac{p^2}{(1-p)^2} = \frac{p}{(1-p)^2}.$$

Repartiția binomială de parametri n, p .

$$M[B_n(p)] = \sum_{k=0}^n k C_n^k p^k (1-p)^{n-k} = np(p+1-p)^n = np,$$

$$\begin{aligned} M_2(B_n(p)) &= \sum_{k=0}^n k^2 C_n^k p^k (1-p)^{n-k} = np \sum_{k=1}^n k C_{n-1}^{k-1} p^{k-1} (1-p)^{n-k} = \\ &= np \sum_{k=1}^n (k-1) C_{n-1}^{k-1} p^{k-1} (1-p)^{n-k} + \sum_{k=1}^n C_{n-1}^{k-1} p^{k-1} (1-p)^{n-k} = \\ &= np [M(B_{n-1}(p))] + (p+1-p)^{n-1} = np [M(B_{n-1}(p)) + 1] = \\ &= np [(n-1)p + 1] \end{aligned}$$

de unde

$$D^2[B_n(p)] = np[(n-1)p + 1] - n^2 p^2 = np(1-p).$$

Repartiția Poisson de parametru λ

$$M(\pi_\lambda) = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = e^{-\lambda} e^{\lambda} \lambda = \lambda,$$

$$\begin{aligned} M_2(\pi_\lambda) &= \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} (k-1) \frac{\lambda^k}{(k-1)!} e^{-\lambda} + \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} = \\ &= \lambda e^{-\lambda} \sum_{k=1}^{\infty} (k-1) \frac{\lambda^{k-1}}{(k-1)!} + \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda M(\pi_\lambda) + \lambda e^{-\lambda} e^{\lambda} = \\ &= \lambda^2 + \lambda, \end{aligned}$$

de unde

$$D^2(\pi_\lambda) = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Repartiția uniformă pe (a, b)

$$M(U_{a,b}) = \frac{1}{b-a} \int_a^b x dx = \frac{a+b}{2},$$

$$M_2(U_{a,b}) = \frac{1}{b-a} \int_a^b x^2 dx = \frac{1}{3} \frac{b^3 - a^3}{b-a}$$

de unde

$$D^2(U_{a,b}) = \frac{1}{3}(a^2 + ab + b^2) - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}.$$

Repartiția exponențială de parametru λ

$$M(e_\lambda) = \int_0^\infty \lambda x e^{-\lambda x} dx = \lambda^{-1}; \quad M_2(e_\lambda) = \int_0^\infty \lambda x^2 e^{-\lambda x} dx = 2\lambda^{-2}$$

de unde

$$D^2(e_\lambda) = 2\lambda^{-2} - \lambda^{-2} = \lambda^{-2}.$$

Repartiția normală de parametri m, σ^2

Făcînd substituția $y = (x - m)/\sigma$ obținem

$$\begin{aligned} M[N(m, \sigma^2)] &= \frac{1}{\sigma\sqrt{2\pi}} \int x \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} dx = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (m + \sigma y) \exp\left(-\frac{y^2}{2}\right) dy = m - \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} d\left[\exp\left(-\frac{y^2}{2}\right)\right] = \\ &= m - \frac{\sigma}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) \Big|_{-\infty}^{+\infty} = m \end{aligned}$$

$$\begin{aligned} D^2[N(m, \sigma^2)] &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (x-m)^2 \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} dx = \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} y^2 \exp\left(-\frac{y^2}{2}\right) dy = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} y d\left(-\exp\left(-\frac{y^2}{2}\right)\right) = \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \left[\left(-y \exp\left(-\frac{y^2}{2}\right)\right) \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} \exp\left(-\frac{y^2}{2}\right) dy \right] = \sigma^2. \end{aligned}$$

Repartiția Cauchy de parametri α, β . Fie f o variabilă aleatoare cu repartiția $C_{\alpha, \beta}$. Atunci $g = (f - \beta)/\alpha$ are repartiția $C_{1,0}$ (se folosește formula (1) din rezumatul teoretic).

Avem

$$\begin{aligned} M(|g|) &= \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{|x|}{1+x^2} dx = \frac{2}{\pi} \int_0^\infty \frac{x}{1+x^2} dx = \frac{2}{\pi} \lim_{b \rightarrow \infty} \int_0^b \frac{x}{1+x^2} dx = \\ &= \frac{1}{\pi} \lim_{b \rightarrow \infty} \ln(1+b^2) = +\infty, \end{aligned}$$

fapt ce arată că g nu are medie (deci nici dispersie), deci și $f = \alpha g + \beta$ nu are medie.

Repartiția gama de parametri α, β .

$$M_r[\gamma(\alpha, \beta)] = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-1+r} e^{-\beta x} dx = \beta^{-r} \frac{\Gamma(\alpha+r)}{\Gamma(\alpha)} = \\ = \beta^{-r} \alpha(\alpha+1) \dots (\alpha+r-1).$$

În particular

$$M[\gamma(\alpha, \beta)] = \alpha\beta^{-1}, \quad M_2[\gamma(\alpha, \beta)] = \beta^{-2}\alpha(\alpha+1)$$

deci

$$D^2[\gamma(\alpha, \beta)] = \beta^{-2}\alpha(\alpha+1) - \alpha^2\beta^{-2} = \alpha\beta^{-2}.$$

Repartiția Dirac în punctul $a \in R^d$. Fie $f = (f_1, \dots, f_d)$ un vector aleator cu repartiția ε_a . Atunci $M(f_i) = \int_{R^d} x_i d\varepsilon_a(x) = a_i$, deci

$$M(f) = a.$$

$$\text{cov}(f_i, f_j) = M(f_i f_j) - a_i a_j = \int_{R^d} x_i x_j d\varepsilon_a(x) - a_i a_j = a_i a_j - a_i a_j = 0$$

deci $C_f = 0$.

Repartiția multinomială de parametri n, p_1, \dots, p_d . Fie $f = (f_1, \dots, f_d)$ un vector aleator cu repartiția $M_n(p_1, \dots, p_d)$.

Atunci

$$M(f_i) = \sum_{n_i=0}^n n_i P(f_i=n_i) = \sum_{n_i=0}^n n_i P\left(\frac{n!}{n_1+\dots+n_{i-1}+n_{i+1}+\dots+n_d=n-n_i} (f_1= \\ = n_1, \dots, f_d=n_d)\right) = \sum_{n_i=0}^n n_i \sum_{n_1+\dots+n_{i-1}+n_{i+1}+\dots+n_d=n-n_i} \frac{n!}{n_1! \dots n_d!} p_1^{n_1} \dots \\ \dots p_d^{n_d} = n p_i \sum_{n_1+\dots+n_d=n} \frac{(n-1)!}{n_1! \dots n_{i-1}! (n_i-1)! n_{i+1}! \dots n_d!} p_1^{n_1} \dots \\ \dots p_{i-1}^{n_{i-1}} p_{i+1}^{n_{i+1}} p_{i+1}^{n_{i+1}} \dots p_d^{n_d} = n p_i \sum_{m_1+\dots+m_d=n-1} \frac{(n-1)!}{m_1! \dots m_d!} p_1^{m_1} \dots \\ \dots p_d^{m_d} = n p_i (p_1 + \dots + p_d)^{n-1} = n p_i.$$

Aplicînd un raționament analog se obține că

$M[f_i(f_i - 1)] = n(n - 1)p_i^2$, deci $M(f_i^2) = n(n - 1)p_i^2 + np_i$ și prin urmare

$$D^2(f_i) = \text{cov}(f_i, f_i) = n(n - 1)p_i^2 + np_i - n^2p_i^2 = np_i(1 - p_i).$$

Dacă $i \neq j$ atunci

$$\begin{aligned} M(f_i, f_j) &= \sum_{n_i+n_j=0}^n n_i n_j P(f_i = n_i, f_j = n_j) = \\ &= \sum_{n_i+n_j=0}^n n_i n_j \sum_{\substack{n_k = n - n_i - n_j \\ k \neq i, j}} \frac{n!}{n_1! \dots n_d!} \cdot p_1^{n_1} \dots p_d^{n_d} = \\ &= n(n - 1)p_i p_j \sum_{m_1 + \dots + m_d = n - 2} \frac{(n - 2)!}{m_1! \dots m_d!} p_1^{m_1} \dots p_d^{m_d} = \\ &= n(n - 1)p_i p_j (p_1 + \dots + p_d)^{n - 2} = n(n - 1)p_i p_j, \end{aligned}$$

deci

$$\begin{aligned} \text{cov}(f_i, f_j) &= M(f_i, f_j) - M(f_i)M(f_j) = n(n - 1)p_i p_j - n^2 p_i p_j = \\ &= -n p_i p_j. \end{aligned}$$

Repartiția uniformă pe mulțimea $D = (a_1, b_1) \times \dots \times (a_d, b_d)$.

Fie $f = (f_1, \dots, f_d)$ un vector aleator cu repartiția uniformă pe D . Atunci întrucît f_i are densitatea $\rho_i(x) = \frac{1}{b_i - a_i} \lambda_{(a_i, b_i)}(x)$ (deci f_i are repartiția uniformă pe (a_i, b_i)) și $\rho(x_1, \dots, x_d) = \rho_1(x_1) \dots \rho_d(x_d)$ (unde ρ este densitatea lui f) rezultă că variabilele aleatoare f_1, \dots, f_d sînt independente.

Prin urmare $M(f_i) = \frac{a_i + b_i}{2}$, $D^2(f_i) = \text{cov}(f_i, f_j) = \frac{1}{12} (b_i - a_i)^2$, iar dacă

$$\begin{aligned} i \neq j, \quad \text{cov}(f_i, f_j) &= M[(f_i - M(f_i))(f_j - M(f_j))] = \\ &= M[f_i - M(f_i)]M[f_j - M(f_j)] = 0 \end{aligned}$$

(am aplicat formula de multiplicare întrucît $f_i - M(f_i)$ și $f_j - M(f_j)$ sînt independente).

Repartiția normală d-dimensională de parametri m, B . Presupunem pentru început că $m = 0$. Fie $\lambda_1, \dots, \lambda_d$ valorile proprii

ale matricei B (care sînt pozitive) și fie H o matrice ortogonală ($H^T = H^{-1}$) astfel încît

$$HBH^T = \Lambda \stackrel{\text{def}}{=} \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_d \end{pmatrix}.$$

Fie f un vector aleator cu repartiția $N_d(0, B)$ și $g = Hf$. Calculăm densitatea vectorului $g = (g_1, \dots, g_d)$.

Conform cu formula (1) din rezumatul teoretic avem că densitatea lui g este

$$\begin{aligned} \rho_g(y) &= \frac{1}{|\det H|} \rho_f(H^T y) = \frac{1}{[(2\pi)^d \det B]^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \langle B^{-1} H^T y, H^T y \rangle\right) = \\ &= \frac{1}{[(2\pi)^d \det B]^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \langle HB^{-1} H^T y, y \rangle\right) = \frac{1}{[(2\pi)^d \det B]^{\frac{1}{2}}} \times \\ &\times \exp\left(-\frac{1}{2} \langle \Lambda^{-1} y, y \rangle\right) = \frac{1}{(\sqrt{2\pi})^d \sqrt{\prod_{i=1}^d \lambda_i}} \exp\left(-\frac{1}{2} \sum_{i=1}^d \frac{1}{\lambda_i} y_i^2\right) = \\ &= \prod_{i=1}^d \frac{1}{\sqrt{2\pi\lambda_i}} \exp\left(-\frac{y_i^2}{2\lambda_i}\right) = \rho_1(y_1) \dots \rho_d(y_d), \end{aligned}$$

unde ρ_i este densitatea repartiției $N(0, \lambda_i)$.

Din § 1 (problema 2.4) rezultă că g_1, \dots, g_d sînt variabile aleatoare independente cu repartițiile $N(0, \lambda_1), \dots, N(0, \lambda_d)$.

Cum $f = H^T g$ rezultă din egalitățile (1.4) că

$M(f) = H^T M(g) = H^T 0 = 0$ și $C_f = C_{H^T g} = H^T C_g H = H^T \Lambda H = B$, $D^2(g_i) = \lambda_i$ și dacă $i \neq j$ din ipoteza de independență rezultă că avem $\text{cov}(g_i, g_j) = 0$, deci $C_g = \Lambda$. Dacă $m \neq 0$ luăm f un vector aleator cu repartiția $N_d(m, B)$. Atunci $g = f - m$ are repartiția $N(0, B)$ (se verifică imediat cu egalitatea (1)) și deci $M(g) = 0 = M(f - m)$ sau $M(f) = m$ și $C = B = C_{f-m} = C_f$ (este imediat faptul că există relația $\text{cov}(f+a, g+a) = \text{cov}(f, g)$ dacă a este o constantă).

2.28. Fie f și g variabile aleatoare cu dispersiile finite. Să se afle:

- 1) $\min_x M[(f - x)^2]$;
- 2) $\min_{x,y} M[(f - xg - y)^2]$ în ipoteza că $D^2(g) > 0$.

Soluție. Din egalitatea

1) $(f - x)^2 = (f - M(f))^2 + (x - M(f))^2 + 2(f - M(f))(M(f) - x)$
 rezultă relația

$$M[(f - x)^2] = D^2(f) + (x - M(f))^2$$

din care se vede că $M[(f - x)^2]$ este minim dacă $(x - M(f))^2 = 0$,
 deci dacă $x = M(f)$.

Prin urmare, $\min M[(f - x)^2] = D^2(f)$ și minimumul se atinge pentru $x = M(f)$.

2) Din punctul 1) rezultă că pentru x fixat expresia $M[(f - xg - y)^2]$ este minimă pentru $y = M(f - xg) = M(f) - xM(g)$.

Cu această alegere a lui y avem că

$$M[(f - xg - y)^2] = M[(f_1 - xg_1)^2]$$

unde $f_1 = f - M(f)$, $g_1 = g - M(g)$.

Mai departe

$$M[(f_1 - xg_1)^2] = M(f_1^2) - 2xM(f_1g_1) + x^2M(g_1^2) = \\ D^2(f) - 2xM(f_1g_1) + x^2D^2(g).$$

Derivând în raport cu x se obține că minimumul este atins pentru

$$x = \frac{M(f_1g_1)}{D^2(g)} = \frac{\text{cov}(f, g)}{D^2(g)} \text{ și este egal cu } D^2(f) - \frac{\text{cov}(f, g)}{D^2(g)}$$

2.29. a) Variabilele aleatoare f_1, \dots, f_n, \dots sînt independente și

$$P(f_n = 1) = \frac{1}{4}, P(f_n = -1) = \frac{1}{4}, P(f_n = 0) = \frac{1}{2}, n = 1, 2, \dots$$

Să se afle media și dispersia variabilei aleatoare $S_n = f_1 + \dots + f_n$.

b) Variabilele aleatoare $\theta_1, \dots, \theta_n, \dots$ sînt independente,

$$P(\theta_n = 0) = P(\theta_n = 1) = \frac{1}{2}, n = 1, 2, \dots, \text{ și fie } \alpha_n = \theta_n - \theta_{n+2}.$$

Pentru orice n să se arate că α_n și f_n au aceeași repartiție și că α_n, α_{n+1} sînt independente. Să se afle media și dispersia variabilei aleatoare $R_n = \alpha_1 + \dots + \alpha_n$.

Soluție. a) $M(f_n) = 1 \cdot \frac{1}{4} - 1 \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} = 0,$

$$M(f_n^2) = 1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} = \frac{1}{2}, D^2(f_n) = M(f_n^2) - M^2(f_n) = \frac{1}{2}, \text{ deci}$$

$$M(S_n) = \sum_{k=1}^n M(f_k) = 0, D^2(S_n) = \sum_{k=1}^n D^2(f_k) = \frac{n}{2}$$

$$b) P(\alpha_n = 0) = P(\theta_n = \theta_{n+2}) = P(\theta_n = 0, \theta_{n+2} = 0) + P(\theta_n = 1, \theta_{n+2} = 1) = P(\theta_n)P(\theta_{n+2} = 0) + P(\theta_n = 1)P(\theta_{n+2} = 1) = \frac{1}{2}$$

$$P(\alpha_n = 1) = P(\theta_n - \theta_{n+2} = 1) = P(\theta_n = 1, \theta_{n+2} = 0) = P(\theta_n = 1)P(\theta_{n+2} = 0) = \frac{1}{4}$$

$$P(\alpha_n = -1) = 1 - P(\alpha_n = 0) - P(\alpha_n = 1) = \frac{1}{4}$$

Independența dintre α_n, α_{n+1} rezultă din asociativitatea independenței sau direct astfel:

$$P(\alpha_n = 0, \alpha_{n+1} = 0) = P(\theta_n = \theta_{n+2}, \theta_{n+1} = \theta_{n+3}) = P(\theta_n = \theta_{n+1} = \theta_{n+2} = \theta_{n+3} = 0) + P(\theta_n = \theta_{n+1} = \theta_{n+2} = \theta_{n+3} = 1) = \\ = \prod_{i=n}^{n+2} P(\theta_i = 0) + \prod_{i=n}^{n+2} P(\theta_i = 1) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2} = P(\alpha_n = 0)P(\alpha_{n+1} = 0).$$

Similar se procedează pentru celelalte cazuri: $\alpha_n = 0, \alpha_{n+1} = 1$; $\alpha_n = 0, \alpha_{n+1} = -1$; $\alpha_n = 1, \alpha_{n+1} = 0$; $\alpha_n = 1, \alpha_{n+1} = 1$; $\alpha_n = 1, \alpha_{n+1} = -1$.

Mai departe

$$R_n = \theta_1 - \theta_3 + \theta_2 - \theta_4 + \theta_3 - \theta_5 + \dots + \theta_n - \theta_{n+2} = \\ = \theta_1 + \theta_2 - \theta_{n+1} - \theta_{n+2}.$$

$$M(R_n) = M(\theta_1) + M(\theta_2) - M(\theta_{n+1}) - M(\theta_{n+2}) = \frac{1}{2} + \frac{1}{2} - \\ - \frac{1}{2} - \frac{1}{2} = 0.$$

$$D^2(R_n) = D^2(\theta_1) + D^2(\theta_2) + D^2(\theta_{n+1}) + D^2(\theta_{n+2}) = \frac{1}{4} + \frac{1}{4} + \\ + \frac{1}{4} + \frac{1}{4} = 1.$$

2.30. Pentru orice întreg $p \geq 1$, fie f_1^p, \dots, f_n^p variabile aleatoare independente și uniform repartizate pe mulțimea $\{1, \dots, p\}$ și $f^p = \max_m f_m^p$.

$$\text{Să se arate că } \lim_{p \rightarrow \infty} \frac{M(f^p)}{p} = \frac{n}{n+1} \text{ și } \lim_{p \rightarrow \infty} \frac{D^2(f^p)}{p^2} = \frac{n}{(n+2)(n+1)^2}.$$

Soluție. Avem

$$\begin{aligned}
 P(f^p < x) &= P(f_1^p < x, \dots, f_n^p < x) = P(f_1^p < x) \dots P(f_n^p < x) = \\
 &= [P(f_1^p < x)]^n = \begin{cases} \left(\frac{1}{p}\right)^n & \text{dacă } 1 \leq x \leq 2, \\ \left(\frac{2}{p}\right)^n & \text{dacă } 2 < x \leq 3, \\ \left(\frac{p-1}{p}\right)^n & \text{dacă } p-1 < x \leq p, \\ 1 & \text{dacă } x > p, \end{cases}
 \end{aligned}$$

deci $P(f^p = k) = P(f^p < k+1) - P(f^p < k) = \left(\frac{k}{p}\right)^n - \left(\frac{k-1}{p}\right)^n$ dacă $1 \leq k \leq p$.

Atunci

$$M(f^p) = \sum_{k=1}^p k P(f^p = k) = p - \frac{1}{p^n} (1^n + \dots + (p-1)^n),$$

deci

$$\frac{M(f^p)}{p} = 1 - \frac{1^n + \dots + (p-1)^n}{p^{n+1}}.$$

Din criteriul lui Stoltz rezultă că $\frac{1^n + \dots + (p-1)^n}{p^{n+1}} \rightarrow \frac{1}{n+1}$ când

$$p \rightarrow \infty, \text{ de unde rezultă că } \frac{M(f^p)}{p} \rightarrow 1 - \frac{1}{n+1} = \frac{n}{n+1}.$$

În particular rezultă și că $\frac{M^2(f^p)}{p^2} \rightarrow \frac{n^2}{(n+1)^2}$.

Mai departe avem

$$\begin{aligned}
 M[(f^p)^2] &= \sum_{k=1}^p k^2 \left[\left(\frac{k}{p}\right)^n - \left(\frac{k-1}{p}\right)^n \right] = \frac{1}{p^n} \sum_{k=1}^p k^{n+2} - \frac{1}{p^n} \sum_{k=1}^p [(k-1)^2 + \\
 &+ 2(k-1) + 1](k-1)^n = \frac{1}{p^n} \sum_{k=1}^p k^{n+2} - \frac{1}{p^n} \sum_{k=1}^p (k-1)^{n+1} - \\
 &- \frac{2}{p^n} \sum_{k=1}^p (k-1)^{n+1} - \frac{1}{p^n} \sum_{k=1}^p (k-1)^n = p^2 - \frac{2}{p^n} \sum_{k=1}^p (k-1)^{n+1} - \\
 &- \frac{1}{p^n} \sum_{k=1}^p (k-1)^n
 \end{aligned}$$

de unde

$$\frac{M(f^2)}{p^2} = 1 - \frac{2}{p^{n+2}} \sum_{k=1}^p (k-1)^{n+1} - \frac{1}{p^{n+2}} \sum_{k=1}^p (k-1)^n \rightarrow 1 - \frac{2}{(n+2)} = \frac{n}{n+2}$$

și prin urmare

$$\frac{D^2(f^2)}{p^2} \rightarrow \frac{n}{n+2} - \frac{n^2}{(n+1)^2} = \frac{n}{(n+1)^2(n+2)}$$

2.31. Variabilele aleatoare f, g sînt uniform repartizate pe $[0, 1]$,

a) Să se arate inegalitatea $M(|f - g|) \leq \frac{1}{2}$.

b) Dacă în plus f, g sînt independente să se afle valoarea minimă și maximă a mărimumi $M(h)$ după toate variabilele aleatoare h cu proprietatea $P(f = h) = P(g = h) = \frac{1}{2}$.

Soluție. a) Avem că $M\left(\left|f - \frac{1}{2}\right|\right) = \int_0^1 \left|x - \frac{1}{2}\right| dx = \frac{1}{4}$, de unde

$$\begin{aligned} M(|f - g|) &= M\left(\left|f - \frac{1}{2} - \left(g - \frac{1}{2}\right)\right|\right) \leq M\left(\left|f - \frac{1}{2}\right|\right) + M\left(\left|g - \frac{1}{2}\right|\right) = \\ &= 2M\left(\left|f - \frac{1}{2}\right|\right) \leq 2 \cdot \frac{1}{4} = \frac{1}{2}; \end{aligned}$$

b) Observăm că $\min(f, g), \max(f, g)$ satisfac $P(f = h) = P(g = h) = \frac{1}{2}$. Într-adevăr este suficient să verificăm (din cauza simetriei) că $P(\min(f, g) = f) = \frac{1}{2}$. Or avem

$$P(\min(f, g) = f) = P(f \leq g) = \int_{x \leq y} dx dy = \int_0^1 \left(\int_0^y dx\right) dy = \int_0^1 y dy = \frac{1}{2}.$$

Dacă vom arăta că $P(\min(f, g) \leq h \leq \max(f, g)) = 1$ pentru h ca în enunț atunci va rezulta că $\min_h M(h) = M[\min(f, g)]$, $\max_h M(h) =$

$= M[\max(f, g)]$ și cum

$$\begin{aligned} M[\min(f, g)] &= \int_0^1 \int_0^1 \min(x, y) \, dx dy = \int_0^1 \left(\int_{x < y} x \, dx + \int_{x > y} y \, dx \right) dy = \\ &= \int_0^1 \left[\frac{y^2}{2} + y(1-y) \right] dy = \frac{1}{3} \end{aligned}$$

$$\begin{aligned} M[\max(f, g)] &= \int_0^1 \int_0^1 \max(x, y) \, dx dy = \int_0^1 \left(\int_{x < y} y \, dx + \int_{x > y} x \, dx \right) dy = \\ &= \int_0^1 \left(\frac{1}{2} + \frac{y^2}{2} \right) dy = \frac{2}{3}. \end{aligned}$$

Va rezulta că $\min M(h) = \frac{1}{3}$, $\max M(h) = \frac{2}{3}$.

Deoarece $P(f = g) = 0$ (vezi problema 2.3) rezultă că

$$P(f = h \text{ sau } g = h) = P(f = h) + P(g = h) = 1.$$

Apoi incluziunea $\{f = h\} \cup \{g = h\} \subset \{\min(f, g) \leq h \leq \max(f, g)\}$ implică faptul că $P(\min(f, g) \leq h \leq \max(f, g)) = 1$.

2.32. Fie f_1, \dots, f_n, \dots un șir de variabile aleatoare independente avînd ca repartiții $e_{\lambda_1}, \dots, e_{\lambda_n}, \dots$. Să se arate că:

a) Dacă $\sum_{n=1}^{\infty} \lambda_n^{-1} < \infty$ atunci $\sum_{n=1}^{\infty} f_n < \infty$ a.s.

b) Dacă $\sum_{n=1}^{\infty} \lambda_n^{-1} = \infty$ atunci $\sum_{n=1}^{\infty} f_n = \infty$ a.s.

Soluție. a) Se știe că $M(f_n) = \lambda_n^{-1}$, deci

$$M\left(\sum_{n=1}^{\infty} f_n\right) = \sum_{n=1}^{\infty} M(f_n) = \sum_{n=1}^{\infty} \lambda_n^{-1} < \infty,$$

de unde $\sum_{n=1}^{\infty} f_n < \infty$ a.s.

$$\begin{aligned} \text{b) Avem } M \left[\exp \left(- \sum_{n=1}^{\infty} f_n \right) \right] &= \lim_{n \rightarrow \infty} M \left[\exp \left(- \sum_{k=1}^n f_k \right) \right] = \\ &= \lim_{n \rightarrow \infty} \prod_{k=1}^n M(e^{-f_k}) \end{aligned}$$

(formula de multiplicare)

$$\text{Apoi } M(e^{-f_k}) = \int_0^{\infty} e^{-x} \lambda_k \exp(-\lambda_k x) dx = \lambda_k (1 + \lambda_k)^{-1} = (1 + \lambda_k^{-1})^{-1}$$

aşa că

$$M \left[\exp \left(- \sum_{n=1}^{\infty} f_n \right) \right] = \lim_{n \rightarrow \infty} \left[\prod_{k=1}^n (1 + \lambda_k^{-1}) \right]^{-1} \leq \lim_{n \rightarrow \infty} \left(\sum_{k=1}^n \lambda_k^{-1} \right)^{-1} = 0.$$

Prin urmare $\exp \left(- \sum_{n=1}^{\infty} f_n \right) = 0$ a.s. fapt ce atrage după sine că

$$\sum_{n=1}^{\infty} f_n = \infty \text{ a.s.}$$

2.33. O matrice $A = (f_{i,j})_{1 \leq i, j \leq n}$ are elementele variabile aleatoare independente cu $M(f_{i,j}) = 0$, $D^2(f_{i,j}) = \sigma^2$ pentru orice i, j .

Să se calculeze media și dispersia determinantului lui A . Să se deducă apoi că maximul determinantilor de ordin n cu elementele ± 1 este mai mare sau egal cu $\sqrt{n!}$

Soluție. Avem că $\det A = \sum_{\sigma} \prod_{i=1}^n f_{i, \sigma(i)}$ unde suma este luată după toate permutările mulțimii $(1, \dots, n)$ și $\varepsilon_{\sigma} = \pm 1$ este semnul lui σ .

Având în vedere formula de multiplicare, obținem:

$$M(\det A) = \sum_{\sigma} \varepsilon_{\sigma} \prod_{i=1}^n M(f_{i, \sigma(i)}) = 0;$$

$$D^2(\det A) = M[(\det A)^2] = \sum_{\sigma, \sigma'} \varepsilon_{\sigma} \varepsilon_{\sigma'} M \left(\prod_{i=1}^n f_{i, \sigma(i)} f_{i, \sigma'(i)} \right) =$$

$$= \sum_{\sigma} M \left(\prod_{i=1}^n f_{i, \sigma(i)}^2 \right) + \sum_{\sigma \neq \sigma'} \varepsilon_{\sigma} \varepsilon_{\sigma'} M \left(\prod_{i=1}^n f_{i, \sigma(i)} f_{i, \sigma'(i)} \right) =$$

$$= \sum_{\sigma} \prod_{i=1}^n M(f_{i, \sigma(i)}^2) + \sum_{\sigma \neq \sigma'} \varepsilon_{\sigma} \varepsilon_{\sigma'} M \left[\prod_{\sigma(i) = \sigma'(i)} f_{i, \sigma(i)}^2 \cdot \prod_{\sigma(i) \neq \sigma'(i)} f_{i, \sigma(i)} f_{i, \sigma'(i)} \right] = \sigma^{2n} n! + \sum_{\sigma \neq \sigma'} \varepsilon_{\sigma} \varepsilon_{\sigma'} \prod_{\sigma(i) = \sigma'(i)} M(f_{i, \sigma(i)}^2) \prod_{\sigma(i) \neq \sigma'(i)} M(f_{i, \sigma(i)}) M(f_{i, \sigma'(i)}) = \sigma^{2n} n!$$

Alegem elementele $f_{i,j}$ ale matricii A așa încît $P(f_{i,j} = \pm 1) = 1/2$ pentru orice i, j , deci în acest caz $D^2(f_{i,j}) = 1$.

Fie D_n maximul din enunț. Avem că $D_n^2 \geq (\det A)^2$ de unde $D_n^2 \geq M(\det A)^2 = n!$.

2.34. Fie f o variabilă aleatoare așa încît f și $-f$ au aceeași repartiție (f se zice că are repartiție simetrică). Să se arate că $M(f^{2k+1}) = 0$ pentru orice $k \geq 0$.

Să se deducă apoi că dacă matricea $A = (f_{ij})_{1 \leq i, j \leq n}$ este așa încît f_1, \dots, f_n sînt variabile aleatoare independente cu repartiție simetrică și cu momentul absolut de ordinul k finit atunci $M(A^k)$ este matrice diagonală (dacă B este o matrice cu elemente variabile aleatoare atunci $M(B)$ este matricea cu elementele mediile componentelor lui B și se numește media lui B).

Soluție. Avem $M(f^{2k+1}) = \int x^{2k+1} d(P \circ f^{-1})(x) = \int x^{2k+1} d(P \circ (-f)^{-1})(x) = \int (-f)^{2k+1} dP = -M(f^{2k+1})$, deci $M(f^{2k+1}) = 0$.

Reamintim că dacă $B = (b_{ij})_{1 \leq i, j \leq n}$, atunci

$$B^k = \left(\sum_{r_1, \dots, r_{k-1}=1}^n a_{ir_1} a_{r_1 r_2} \dots a_{r_{k-1} j} \right)_{1 \leq i, j \leq n}$$

În cazul nostru obținem expresia

$$A^k = \left(\sum_{r_1, \dots, r_{k-1}=1}^n f_i f_{r_1}^2 \dots f_{r_{k-1}}^2 f_j \right)_{1 \leq i, j \leq n}$$

Este suficient să arătăm că dacă $i \neq j$ și $1 \leq r_1, \dots, r_{k-1} \leq n$, atunci

$$M(f_i f_{r_1}^2 \dots f_{r_{k-1}}^2 f_j) = 0.$$

Dacă printre r_1, \dots, r_{k-1} nu se află i și j atunci din asociativitatea independenței și formula de multiplicare obținem

$$M(f_i f_{r_1}^2 \dots f_{r_{k-1}}^2 f_j) M(f_i) = M(f_{r_1}^2 \dots f_{r_{k-1}}^2) M(f_j) = 0$$

(din prima parte a problemei avem că $M(f_i) = M(f_j) = 0$).

Să presupunem acumă că cel puțin unul dintre i, j se află printre r_1, \dots, r_{k-1} .

Pentru a face o alegere să presupunem că i se află printre r_1, \dots, r_{k-1} . Fie I mulțimea indicilor din r_1, \dots, r_{k-1} ce sînt egali cu i .

Avem

$$M(f_i f_{r_1}^2 \dots f_{r_{k-1}}^2 f_j) = M(f_i^{2^{\text{card } I+1}} \prod_{r_1 \neq i} f_{r_1}^2 f_j) = M(f_i^{2^{\text{card } I+1}}) M(\prod_{r_1 \neq i} f_{r_1}^2 f_j) = 0$$

(asociativitatea independenței și formula de multiplicare),

deoarece $M(f_i^{2^{\text{card } I+1}}) = 0$, conform primei părți a problemei.

2.35. Fie f_1, \dots, f_n, \dots un șir de variabile aleatoare independente și de medie finită și fie τ o variabilă aleatoare cu valori în $\{1, \dots, n, \dots\}$ ce are media finită și este independentă de șirul $(f_n)_n$.

Să se deducă egalitatea

$$M(f_1 + \dots + f_\tau) = M(f_1) M(\tau) \quad (\text{identitatea lui Wald}). \quad (1)$$

Dacă în plus f_1 și τ au dispersii finite să se deducă egalitatea

$$D^2(f_1 + \dots + f_\tau) = M(\tau) D^2(f_1) + D^2(\tau) M^2(f_1) \quad (2)$$

Soluție. Fie $S_n = f_1 + \dots + f_n$. Avem

$$\begin{aligned} M(|S_\tau|) &= \sum_{k=1}^{\infty} \int_{(\tau=k)} |S_\tau| dP = \sum_{k=1}^{\infty} \int_{(\tau=k)} |S_k| dP \leq \\ &\leq \sum_{k=1}^{\infty} \int_{(\tau=k)} (|f_1| + \dots + |f_k|) dP = \sum_{k=1}^{\infty} P(\tau = k) M(|f_1| + \dots \\ &\dots + |f_k|) = M(|f_1|) \sum_{k=1}^{\infty} k P(\tau = k) = M(|f_1|) M(\tau), \end{aligned}$$

deci S_τ are medie finită.

Reluînd calculele fără modul (în acest caz în calculele de mai sus avem numai egalități) obținem egalitatea (1).

Presupunem acum că τ și f_1 au dispersii finite. Ținând cont că $D^2(S_k) = D^2(f_1) + \dots + D^2(f_k) = kD^2(f_1)$ și de formula de multiplicare, putem scrie

$$\begin{aligned} M(S_k^2) &= \sum_{k=1}^{\infty} \int_{(\tau=k)} S_k^2 dP = \sum_{k=1}^{\infty} M(S_k^2) P(\tau = k) = \sum_{k=1}^{\infty} [M(S_k^2) - \\ &- M^2(S_k)] P(\tau = k) + \sum_{k=1}^{\infty} M^2(S_k) P(\tau = k) = \sum_{k=1}^{\infty} D^2(S_k) P(\tau = k) + \\ &+ M^2(f_1) \sum_{k=1}^{\infty} k^2 P(\tau = k) = D^2(f_1) \sum_{k=1}^{\infty} k P(\tau = k) + M^2(f_1) M(\tau^2) = \\ &= D^2(f_1) M(\tau) + M^2(f_1) M(\tau^2) \end{aligned}$$

2.36. Variabilele aleatoare f_1, \dots, f_n sînt pozitive, independente și identic repartizate. Dacă punem $g_k = f_k(f_1 + \dots + f_n)$ se cere:

1) Să se arate că g_1, \dots, g_n sînt identic repartizate și să se calculeze $M(g_k)$.

2) Să se arate că vectorii aleatori $\{(g_k, g_l)\}_{k \neq l}$ sînt identic repartizați și să se calculeze $\frac{\text{cov}(g_k, g_l)}{D(g_k) D(g_l)}$ dacă f, g sînt variabile aleatoare atunci raportul $\rho(f, g) = \frac{\text{cov}(f, g)}{D(f) D(g)}$ se numește coeficientul de corelație dintre f și g .

Soluție. Deoarece $0 \leq g_k \leq 1$ rezultă că g_k are momente de orice ordin.

1) Fie $F = P \circ f_k^{-1}$ și fie $u: R^n \rightarrow R$ aplicația continuă (deci boreliană) definită prin $u(x_1, \dots, x_n) = x_k/(x_1 + \dots + x_n)$.

Avem

$$\begin{aligned} P \circ (f_1, \dots, f_n)^{-1} &= (P \circ f_1^{-1}) \otimes \dots \otimes (P \circ f_n^{-1}) = F \otimes \dots \otimes F \\ P \circ (f_k, f_2, \dots, f_{k-1}, f_1, f_{k+1}, \dots, f_n)^{-1} &= (P \circ f_k^{-1}) \otimes \dots \otimes (P \circ f_{k-1}^{-1}) \otimes \\ &\otimes (P \circ f_1^{-1}) \otimes (P \circ f_{k+1}^{-1}) \otimes \dots \otimes (P \circ f_n^{-1}) = F \otimes \dots \otimes F \end{aligned}$$

$$\begin{aligned} \text{Atunci } P \circ g_k^{-1} &= P \circ [u(f_1, \dots, f_n)]^{-1} = P \circ (f_1, \dots, f_n)^{-1} \circ u^{-1} = \\ &= P \circ (f_k, f_2, \dots, f_{k-1}, f_1, f_{k+1}, \dots, f_n)^{-1} \circ u^{-1} = \\ &= P \circ [u(f_k, f_2, \dots, f_{k-1}, f_1, f_{k+1}, \dots, f_n)]^{-1} = P \circ g_1^{-1} \end{aligned}$$

deci g_k are aceeași repartiție cu g_1 .

În particular rezultă că $M(g_k) = M(g_1)$ și cum $g_1 + \dots + g_k = 1$ avem că $1 = M(g_1) + \dots + M(g_n) = nM(g_1)$ ceea ce atrage după sine că $M(g_1) = M(g_k) = 1/n$.

2) Este vizibil că dacă $k = l$ atunci coeficientul de corelație care se cere este egal cu 1.

Fie acum $k \neq l$ și pentru a face o alegere să presupunem $k < l$ și fie funcția $v: R^n \rightarrow R^2$ continuă (deci boreliană) definită prin $v(x_1, \dots, x_n) = (x_k/(x_1 + \dots + x_n), x_l/(x_1 + \dots + x_n))$.

Avem

$v(f_1, \dots, f_n) = (g_k, g_l)$, $v(f_k, f_l, \dots, f_{k-1}, f_1, f_{k+1}, \dots, f_{l-1}, f_2, f_{l+1}, \dots, f_n) = (g_1, g_2)$. Ca mai înainte se arată că vectorii (f_1, \dots, f_n) și $(f_k, f_l, \dots, f_{k-1}, f_1, f_{k+1}, \dots, f_{l-1}, f_2, f_{l+1}, \dots, f_n)$ au aceeași repartiție, fapt ce atrage după sine că vectorii (g_1, g_2) și (g_k, g_l) au aceeași repartiție.

$$\text{Mai departe } D^2(g_k) = D^2(g_l) = D^2(g_1) = M(g_1^2) - M^2(g_1) = \\ = M(g_1^2) - 1/n^2$$

$$\text{cov}(g_k, g_l) = \text{cov}(g_1, g_2) = M(g_1, g_2) - M(g_1)M(g_2) = M(g_1g_2) - 1/n^2$$

$$1 = (g_1 + \dots + g_n)^2 = g_1^2 + \dots + g_n^2 - 2 \sum_{k < l} g_k g_l \text{ implică}$$

$$1 = M[(g_1 + \dots + g_n)^2] = nM(g_1^2) + n(n-1)M(g_1g_2), \text{ deci}$$

$$M(g_1g_2) = \frac{1 - nM(g_1^2)}{n(n-1)}$$

Rezultă că

$$\rho(g_k, g_l) = \frac{\text{cov}(g_k, g_l)}{D(g_k) D(g_l)} = \frac{M(g_1g_2) - 1/n^2}{M(g_1^2) - 1/n^2} = \\ = \frac{[1 - nM(g_1^2)]/n(n-1) - 1/n^2}{M(g_1^2) - 1/n^2} = -\frac{1}{n-1}$$

2.37. Să se arate că pentru orice variabilă aleatoare f au loc inegalitățile

$$\sum_{n=1}^{\infty} P(|f| \geq n) \leq M(|f|) \leq 1 + \sum_{n=1}^{\infty} P(|f| \geq n)$$

Observație. Din aceste inegalități rezultă în particular că f are medie finită dacă și numai dacă $\sum_{n=1}^{\infty} P(|f| \geq n) < \infty$

Soluție. Fie $A_n = \{n \leq |f| < n+1\}$. Atunci $M(|f|) = \sum_{n=0}^{\infty} \int_{A_n} |f| dP$

și cum $nP(A_n) \leq \int_{A_n} |f| dP \leq (n+1)P(A_n)$ obținem prin urmare

$$\begin{aligned} \sum_{n=0}^{\infty} nP(A_n) &\leq M(|f|) \leq \sum_{n=0}^{\infty} (n+1)P(A_n) = \sum_{n=0}^{\infty} P(A_n) + \\ &+ \sum_{n=0}^{\infty} nP(A_n) = 1 = 1 + \sum_{n=0}^{\infty} nP(A_n) \end{aligned}$$

fapt ce arată că totul se reduce la a arăta că

$$\sum_{n=0}^{\infty} nP(A_n) = \sum_{n=1}^{\infty} P(|f| \geq n).$$

Avem că

$$\sum_{n=0}^{\infty} nP(A_n) = \sum_{n=0}^{\infty} \sum_{i=1}^n P(A_i) = \sum_{i=1}^{\infty} \sum_{n=i}^{\infty} P(A_n) = \sum_{i=1}^{\infty} P(|f| \geq i)$$

(seriile cu termeni pozitivi comută).

2.38. Fie f o variabilă aleatoare cu valori în mulțimea a_1, \dots, a_n, \dots unde $a_1 < \dots < a_n < \dots$. Să se deducă egalitatea

$$M(f) = a_1 + \sum_{k=1}^{\infty} (a_{k+1} - a_k)P(f \geq a_{k+1}). \quad (1)$$

În particular dacă f ia valori în $\{0, 1, \dots, n, \dots\}$ egalitatea (1) se scrie

$$M(f) = \sum_{n=1}^{\infty} P(f \geq n). \quad (2)$$

Soluție. Avem că $M(f) = \sum_{k=1}^{\infty} a_k P(f = a_k) = \sum_{k=1}^{\infty} a_k [P(f \geq a_k) - P(f \geq a_{k+1})] = a_1 + \sum_{k=1}^{\infty} (a_{k+1} - a_k)P(f \geq a_{k+1})$ (prin grupare)

2.39. Fie f o variabilă aleatoare cu valori nenegative. Să se deducă egalitatea

$$M(f) = \int_0^{\infty} P(f > t) dt \quad (1)$$

Să se arate apoi că egalitatea (1) este echivalentă cu egalitatea

$$M(f) = \int_0^{\infty} P(f \geq t) dt \quad (2)$$

Soluție. Presupunem întâi că f ia valorile $a_1 < \dots < a_n < \dots$.
Avem că

$$\begin{aligned} \int_0^{\infty} P(f > t) dt &= \int_{[0, a_1)} P(f \geq t) dt + \sum_{k=1}^{\infty} \int_{[a_k, a_{k+1})} P(f > t) dt = \\ &= \int_{[0, a_1)} P(f \geq a_1) dt + \sum_{k=1}^{\infty} \int_{[a_k, a_{k+1})} P(f \geq a_{k+1}) dt = \\ &= a_1 + \sum_{k=1}^{\infty} (a_{k+1} - a_k) P(f \geq a_{k+1}) = M(f) \end{aligned}$$

(ultima egalitate rezultă din problema precedentă).

Dacă f este arbitrară (dar cu valori în R_+) alegem un șir f_n de variabile aleatoare simple nenegative așa încît $f_n \nearrow f$. Atunci avem că $\{f_n > t\} \nearrow \{f > t\}$ așa că $P(f_n > t) \nearrow P(f > t)$.

Variabilele f_n fiind simple deja știm că

$$M(f_n) = \int_0^{\infty} P(f_n > t) dt$$

Deoarece integrala comută cu limitele ascendente prin trecere la limită în ultima egalitate obținem (1).

Se știe că o repartiție încarcă numai o mulțime cel mult numărabilă de puncte, așa că $\{t; P(f > t) \neq P(f \geq t)\}$ are măsura Lebes-

gue nulă, deci

$$\begin{aligned} \int_0^{\infty} P(f \geq t) dt &= \int_{\{t; P(f > t) = P(f \geq t)\}} P(f > t) dt = \int_{\{t; P(f > t) = P(f \geq t)\}} P(f \geq t) dt = \\ &= \int_0^{\infty} P(f \geq t) dt, \text{ cu alte cuvinte și (2) are loc.} \end{aligned}$$

Soluția 2. Folosind teorema lui Fubini obținem:

$$\begin{aligned} \int_0^{\infty} P(f > t) dt &= \int_0^{\infty} \left[\int_t^{\infty} d(P \circ f^{-1})(x) \right] dt = \int_0^{\infty} \left(\int_0^x dt \right) d(P \circ f^{-1})(x) = \\ &= \int_0^{\infty} x d(P \circ f^{-1})(x) = \int f dP = M(f) \end{aligned}$$

2.40. Fie f_1, \dots, f_n, \dots un șir de variabile aleatoare independente așa încît $\sum_{n=1}^{\infty} M(f_n^2) < \infty$. Să se arate că $M(\sup_n |f_n|) < \infty$

Soluție. Fie $f^* = \sup_n f_n$, $f_n^* = \sup_{k \leq n} |f_k|$. Întrucît are loc egalitatea

$$M(f^*) = \int_0^{\infty} P(f^* > x) dx \text{ (vezi problema 2.39) rezultă că este sufi-$$

cient să arătăm că există $a > 0$ astfel încît $\int_0^{\infty} P(f^* > x) dx < \infty$.

Din $\{f_n^* > x\} \nearrow \{f^* > x\}$ rezultă că $P(f^* > x) = \lim_{n \rightarrow \infty} P(f_n^* > x) \leq$

$$\leq \lim_{n \rightarrow \infty} \sum_{k=1}^n P(|f_k| > x) \leq \frac{1}{x^2} \lim_{n \rightarrow \infty} \sum_{k=1}^n M(|f_k|^2) = \frac{1}{x^2} \sum_{k=1}^{\infty} M(|f_k|^2)$$

Deoarece pentru orice $a > 0$ avem $\int_a^{\infty} \frac{1}{x^2} dx < \infty$, rezultă că și

$$\int_a^{\infty} P(f^* > x) dx < \infty$$

2.41. Fie f_1, \dots, f_n, \dots un șir de variabile aleatoare independente, identic repartizate și cu repartiție continuă și fie $\tau = \min (n \geq 2, f_n > f_1)$:

- a) să se afle repartiția lui τ ;
 b) să se calculeze $M(\tau)$.

Soluție. Putem scrie

$$\tau = \begin{cases} 2 & f_2 > f_1 \\ k & f_k > f_1, f_j \leq f_1, j = 1, \dots, k-1 \\ \infty & f_n \leq f_1 \text{ pentru orice } n \geq 2 \end{cases}$$

asa că $P(\tau > n) = P(f_2 \leq f_1, f_3 \leq f_1, \dots, f_n \leq f_1) = P(f_1 \leq f_2, f_3 \leq f_2, \dots, f_n \leq f_2) = \frac{1}{n}$ și prin urmare $P(\tau = n) = P(\tau > n-1) -$

$P(\tau > n) = \frac{1}{(n-1)n}$ dacă $n \geq 2$. Deoarece $\sum_{n=2}^{\infty} P(\tau = n) = 1$ rezultă

că $P(\tau < \infty) = 1$.

$$\text{De asemenea } M(\tau) = \sum_{n=2}^{\infty} n P(\tau = n) = \sum_{n=2}^{\infty} \frac{1}{(n-1)} = \infty$$

2.42. În ipotezele problemei precedente fie $\tau_m = \min (n > m; f_n > \max_{k \leq m} f_k)$. Să se arate că $P(\tau_m > n) = \frac{m}{n}$.

Soluție. Aplicăm inducția după m . Pentru $m = 1$ obținem tocmai rezultatul din problema precedentă (vezi soluția). Presupunem afirmația adevărată pentru $m-1$ și să o arătăm pentru m . Avem pentru $n \geq m+1$:

$$\begin{aligned} P(\tau_m > n) &= P(f_{m+1} \leq \max(f_1, \dots, f_m), \dots, f_n \leq \max(f_1, \dots, f_m)) = \\ &= P(f_m \leq \max(f_1, \dots, f_{m-1}), f_{m+1} \leq \max(f_1, \dots, f_{m-1}), \dots, f_n \leq \\ &\leq \max(f_1, \dots, f_{m-1})) + P(\max(f_1, \dots, f_{m-1}) < f_m, f_{m+1} \leq f_m, \dots, \\ &f_n \leq f_m) = \frac{m-1}{n} + P(f_1 \leq f_m, \dots, f_{m-1} \leq f_m, f_{m+1} \leq f_m, \dots, f_n \leq f_m) = \\ &= \frac{m-1}{n} + \frac{1}{n} = \frac{m}{n}. \end{aligned}$$

Am folosit mai sus faptul că dacă f, g sînt variabile aleatoare independente și una are repartiția continuă atunci $P(f = g) = 0$ (vezi problema 2.3):

2.43. Dacă f_1, \dots, f_n, \dots este un șir de variabile aleatoare independente cu repartiția Cauchy de parametru 1 să se arate că

$$\lim_{n \rightarrow \infty} P\left(\frac{1}{n} \max_{k \leq n} f_k < x\right) = e^{-\frac{1}{\pi x}}.$$

Soluție. Avem că

$$\begin{aligned} P\left(\frac{1}{n} \max_{k \leq n} f_k < x\right) &= P(f_1 < nx, \dots, f_n < nx) = \\ &= P(f_1 < nx) \dots P(f_n < nx) = \left(\frac{1}{\pi} \int_{-\infty}^{nx} \frac{1}{1+y^2} dy\right)^n = \\ &= \left(1 - \frac{1}{\pi} \int_{nx}^{\infty} \frac{1}{1+y^2} dy\right)^n = \left[1 - \frac{1}{\pi} \left(\frac{\pi}{2} - \operatorname{arctg} nx\right)\right]^n = \\ &= \left\{ \left[1 - \frac{1}{\pi} \left(\frac{\pi}{2} - \operatorname{arctg} nx\right)\right]^{\frac{\pi}{2} - \operatorname{arctg} nx} \right\}^{-\frac{n}{\pi} \left(\frac{\pi}{2} - \operatorname{arctg} nx\right)} \end{aligned}$$

deci

$$\lim_{n \rightarrow \infty} P\left(\frac{1}{n} \max_{k \leq n} f_k < x\right) = e^{-\frac{1}{\pi} \lim_{n \rightarrow \infty} n \left(\frac{\pi}{2} - \operatorname{arctg} nx\right)} = e^{-\frac{1}{\pi x}}$$

2.44. Fie f_1, \dots, f_n, \dots un șir de variabile aleatoare independente, identic repartizate și de medie finită. Să se arate că

$$\lim_{n \rightarrow \infty} \frac{1}{n} M(\max_{k \leq n} |f_k|) = 0$$

Soluție. Fie F funcția de repartiție a lui $|f_1|$. Deoarece $M(|f_1|) < \infty$ rezultă că $\sum_{k=1}^{\infty} [1 - F(k)] < \infty$ (vezi problema 2.37).

Întrucît $\frac{1}{n} P\left(\max_{k \leq n} |f_k| \geq x\right) \leq \frac{1}{n} \sum_{k=1}^n P(|f_k| \geq x) = P(|f_1| \geq x) = 1 - F(x)$ uniform în n , rezultă că

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{l=1}^{\infty} P(\max_{k \leq n} |f_k| \geq l) = \sum_{l=1}^{\infty} \lim_{n \rightarrow \infty} \frac{1}{n} P(\max_{k \leq n} |f_k| \geq l) = 0$$

În fine din egalitatea $\frac{1}{n} M(\max_{k \leq n} |f_k|) = \frac{1}{n} \sum_{l=1}^{\infty} P(\max_{k \leq n} |f_k| \geq l)$

(vezi problema 2.38) obținem egalitatea din enunțul problemei.

2.45. Să se arate că dacă pentru două variabile aleatoare f, g este adevărată formula de multiplicare $M(fg) = M(f)M(g)$ nu rezultă că f și g sînt independente.

Soluție. Fie f, g așa încît vectorul (f, g) să aibă densitatea

$$p(x, y) = f(x, y) + \frac{1}{2\pi} \exp\{-(x^2 + y^2)\}$$

unde

$$f(x, y) = \begin{cases} \frac{e^{-4}}{4\pi} & 1 < |x| < 2 \text{ și } 1 < |y| < 2 \text{ sau} \\ & 0 < |x| < 1 \text{ și } 0 < |y| < 1 \\ -\frac{e^{-4}}{4\pi} & 1 < |x| < 2 \text{ și } 0 < |y| < 1 \text{ sau} \\ & 0 < |x| < 1 \text{ și } 1 < |y| < 2 \\ 0 & \text{în rest} \end{cases}$$

Rezultă că f și g au densitatea $\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$, cu alte cuvinte f și g au repartiția $N(0, 1)$ și deci în particular $M(f) = M(g) = 0$.

Pe de altă parte

$$M(fg) = \int xy p(x, y) dx dy = 0$$

așa că

$$M(fg) = M(f)M(g)$$

Însă f și g nu sînt independente pentru că densitatea vectorului (f, g) nu este egală cu produsul densităților lui f și g .

Un alt contraexemplu va fi dat în paragraful următor.

$$\mathbf{2.46.}$$
 Fie $f = \sum_{k=1}^m x_k \lambda_{\{f=x_k\}}$ și $g = \sum_{k=1}^n y_k \lambda_{\{g=y_k\}}$

variabile aleatoare simple.

Să se arate că f și g sînt independente dacă și numai dacă

$$M(f^r g^s) = M(f^r)M(g^s) \text{ pentru orice } 1 \leq r \leq m-1, 1 \leq s \leq n-1.$$

Soluție. Fie $p_k = P(f = x_k)$, $q_l = P(g = y_l)$, $p_{k,l} = P(f = x_k, g = y_l)$ $1 \leq k \leq m$, $1 \leq l \leq n$ și să presupunem că are loc egalitatea din enunț. Independența dintre f și g rezultă dacă vom arăta că $p_{k,l} = p_k q_l$.

Egalitatea din ipoteză se mai scrie sub forma

$$\sum_{k=1}^m \sum_{l=1}^n p_{k,l} x_k^r y_l^s = \left[\sum_{k=1}^m p_k x_k^r \right] \left[\sum_{l=1}^n q_l y_l^s \right] \quad \text{dacă } 1 \leq r \leq m-1, \\ 1 \leq s \leq n-1. \quad (1)$$

Se observă că egalitatea de mai sus este valabilă și pentru $r = 0$, $s = n-1$ și $r = m-1$, $s = 0$.

Cu notația $\delta_{k,l} = p_{k,l} - p_k q_l$ egalitatea (1) se scrie

$$\sum_{k=1}^m \sum_{l=1}^n \delta_{k,l} x_k^r y_l^s = 0 \quad \text{pentru } 0 \leq r \leq m-1, 0 \leq s \leq n-1. \quad (2)$$

Egalitatea (2) arată că pentru r fixat numerele $d_{r,l} = \sum_{k=1}^m \delta_{k,l} x_k^r$ verifică sistemul de n ecuații și n necunoscute

$$\sum_{l=1}^n d_{r,l} y_l^s = 0; \quad 0 \leq s \leq n-1. \quad (3)$$

Oricum acest sistem este omogen și are determinantul coeficienților un determinant Vandermonde (deci nenul) rezultă că el are numai soluția nulă, deci

$$\sum_{k=1}^m \delta_{k,l} x_k^r = 0; \quad 0 \leq r \leq m-1; 1 \leq l \leq n. \quad (4)$$

Noul sistem (4) fiind de genul lui (3) pentru l fixat și $0 \leq r \leq m-1$ atrage faptul că $\delta_{k,l} = 0$ pentru orice $1 \leq k \leq m$, $1 \leq l \leq n$.

Invers presupunem că f și g sînt independente. Atunci și f^r și g^s sînt independente și din formula de multiplicare obținem valabilitatea egalității din enunțul problemei.

2.47. Fie f și g variabile aleatoare mărginite. Să se arate că f și g sînt independente dacă și numai dacă $M(f^r g^s) = M(f^r) M(g^s)$ pentru orice întregi $r, s \geq 1$.

Soluție. Valabilitatea egalității din enunț în ipoteza de independență rezultă din independența lui f^r , g^s și formula de multiplicare.

Invers să presupunem că are loc egalitatea din enunț. Independența dintre f și g rezultă dacă vom arăta că

$$P(f \in D, g \in G) = P(f \in D)P(g \in G) \text{ pentru } D \text{ și } G \text{ deschiși în } R. \quad (1)$$

Fie $u_n(x) = \min(1, n \cdot \text{dist}(x, \complement D))$, $v(x) = \min(1, n \cdot \text{dist}(x, \complement G))$. Atunci u_n și v_n sînt continue, $0 \leq u_n, v_n \leq 1$, u_n converge către λ_D , v_n converge către λ_G .

Avînd în vedere teorema de convergență dominată rezultă că cei doi membri din (1) se scriu

$$\begin{aligned} P(f \in D, g \in G) &= \int \lambda_D(f) \lambda_G(g) dP = \lim_{n \rightarrow \infty} \int u_n(f) v_n(g) dP \\ P(f \in D)P(g \in G) &= \int_D (f) dP \cdot \int \lambda_G(g) dP = \lim_{n \rightarrow \infty} \int u_n(f) dP \cdot \\ &\cdot \lim \int v_n(g) dP = \lim_{n \rightarrow \infty} \int u_n(f) dP \int v_n(g) dP \end{aligned}$$

prin urmare este suficient să dovedim că

$$\int u(f)v(g)dP = \int u(f)dP \int v(g)dP. \quad (2)$$

pentru u, v continue și mărginite, sau este suficient (întrucît f și g sînt mărginite) ca u, v să fie continue și cu suport compact.

Teorema lui Weierstrass ne spune că există P_n, Q_n polinoame așa încît P_n converge uniform către u și Q_n converge uniform către v . Procedînd ca mai sus pentru obținerea lui (2) (teorema de convergență dominată funcționează) deducem că totul se reduce la a arăta că

$$\int P(f)Q(g) dP = \int P(f) dP \int Q(g) dP, \quad (3)$$

pentru P, Q polinoame.

În fine întrucît integrala este aditivă observăm că putem lua $P(x) = x^r, Q(x) = x^s$.

Or în acest caz (3) se scrie

$$M(f^r g^s) = M(f^r) M(g^s) \quad (4)$$

ceea ce nu este altceva decît ipoteza.

2.48. Dacă variabilele aleatoare f și g sînt independente și identic repartizate se cere să se calculeze $M(f/f + g)$.

Soluție. Fie $P \circ f^{-1} = P \circ g^{-1} = F$. Din independența rezultă că

$$P \circ (f, g)^{-1} = (P \circ f^{-1}) \otimes (P \circ g^{-1}) = F \otimes F; \quad P \circ (g, f)^{-1} =$$

$$= (P \circ g^{-1}) \otimes (P \circ f^{-1}) = F \otimes F$$

deci $P \circ (f, g)^{-1} = P \circ (g, f)^{-1}$.

Dacă A este o mulțime boreliană avem că

$$\int_{(f+g \in A)} f dP = \int f \lambda_A(f+g) dP = \int x \lambda_A(x+y) dP \circ (f, g)^{-1}(x, y) =$$

$$= \int x \lambda_A(x+y) dP \circ (g, f)^{-1}(x, y) = \int g \lambda_A(g+f) dP = \int_{(f+g \in A)} g dP$$

deci obținem că $M(f|f+g) = M(g|f+g)$. Pe de altă parte cum

$$M(f|f+g) + M(g|f+g) = M(f+g|f+g) = f+g$$

deducem că

$$M(f|f+g) = \frac{f+g}{2}$$

2.49. Fie f_1, \dots, f_n, \dots variabile aleatoare independente, identic repartizate și de medie finită și fie $S_n = f_1 + \dots + f_n$. Să se arate că

$$M(f_1/S_n, S_{n+1}, \dots) = \frac{1}{n} S_n$$

Soluție. Vom arăta pentru început că

$$M(f_1/S_n, S_{n+1}, \dots) = M(f_1/S_n)$$

Întrucît $\mathfrak{B}(S_n, S_{n+1}, \dots) = \mathfrak{B}(S_n, f_{n+1}, \dots)$ egalitatea precedentă se scrie

$$M[f_1|S_n, f_{n+1}, \dots] = M(f_1/S_n)$$

sau echivalent

$$\int_A f_1 dP = \int_A M(f_1/S_n) dP \text{ pentru orice } A \in \mathfrak{B}(S_n, f_{n+1}, \dots)$$

sau încă, dacă avem în vedere că $\mathfrak{B}(S_n, f_{n+1}, \dots) = \mathfrak{B}(\{S_n \in A\} \cap B; A \in \mathfrak{B}_R \text{ și } B \in \mathfrak{B}(f_{n+1}, \dots))$ și teorema de unicitate a probabilităților,

$$\int_{\{S_n \in A\} \cap B} f_1 dP = \int_{\{S_n \in A\} \cap B} M(f_1/S_n) dP \text{ pentru } A \in \mathfrak{B}_R \text{ și } B \in \mathfrak{B}(f_{n+1}, \dots).$$

Ori din asociativitatea independenței rezultă că (f_1, S_n) și $\mathfrak{B}(f_{n+1}, \dots)$ sînt independente, așa că ținînd cont și de formula de multiplicare obținem

$$\begin{aligned} \int_{\{S_n \in A\} \cap B} f_1 dP &= \int \lambda_A(S_n) f_1 \lambda_B dP = \int \lambda_A(S_n) f_1 dP \cdot P(B) = \\ &= \int \lambda_A(S_n) M(f_1/S_n) dP \cdot P(B) = \int \lambda_A(S_n) M(f_1/S_n) \lambda_B dP = \\ &= \int_{\{S_n \in A\} \cap B} M(f_1/S_n) dP \end{aligned}$$

ceea ce justifică egalitatea.

Asemănător ca în problema precedentă se arată că $M(f_i/S_n) = M(f_i/S_n)$. În fine putem scrie

$$M(f_1/S_n) = \frac{1}{n} \sum_{i=1}^n M[f_i/S_n] = \frac{1}{n} M\left[\sum_{i=1}^n f_i/S_n\right] = \frac{1}{n} M[S_n/S_n] = \frac{1}{n} S_n$$

2.50. Fie f_1, \dots, f_n variabile aleatoare independente și cu repartițiile $\pi_{\lambda_1}, \dots, \pi_{\lambda_n}$. Să se calculeze

$$P(f_n = l | f_1 + \dots + f_n = k); \quad M(f_n | f_1 + \dots + f_n)$$

Soluție. În primul rînd observăm că $0 \leq l \leq k$. Avem că

$$P(f_n = l | f_1 + \dots + f_n = k) = \frac{P(f_n = l, f_1 + \dots + f_n = k)}{P(f_1 + \dots + f_n = k)}$$

$$\frac{P(f_n = l, f_1 + \dots + f_{n-1} = k - l)}{P(f_1 + \dots + f_n = k)} = \frac{P(f_n = l)P(f_1 + \dots + f_{n-1} = k - l)}{P(f_1 + \dots + f_n = k)}$$

(sumă de variabilă aleatoare independente cu repartiții Poisson este încă o variabilă aleatoare cu repartiția Poisson de parametru suma parametrilor; vezi problema 2.9).

$$\begin{aligned} &\frac{\lambda_n^l}{l!} e^{-\lambda_n} \frac{(\lambda_1 + \dots + \lambda_{n-1})^{k-l}}{(k-l)!} e^{-(\lambda_1 + \dots + \lambda_{n-1})} \times \\ &\times \frac{k!}{(\lambda_1 + \dots + \lambda_n)^k} e^{-(\lambda_1 + \dots + \lambda_n)} = C_{k,l} \frac{\lambda_n^l (\lambda_1 + \dots + \lambda_{n-1})^{k-l}}{(\lambda_1 + \dots + \lambda_n)^k} \end{aligned}$$

Mai departe putem scrie

$$\begin{aligned} M(f_n/f_1 + \dots + f_n = k) &= \sum_{l=0}^k lP(f_n = l/f_1 + \dots + f_n = k) = \\ &= \sum_{l=0}^k l C_k^l \frac{\lambda_n^l (\lambda_1 + \dots + \lambda_{n-1})^{k-l}}{(\lambda_1 + \dots + \lambda_n)^k} = M\left(B_k\left(\frac{\lambda_n}{\lambda_1 + \dots + \lambda_n}\right)\right) = \\ &= k \frac{\lambda_n}{\lambda_1 + \dots + \lambda_n} \end{aligned}$$

de unde

$$\begin{aligned} M(f_n/f_1 + \dots + f_n) &= \sum_{k=0}^{\infty} P(f_1 + \dots + f_n = k) M(f_n/F_1 + \dots + f_n = k) = \\ &= \sum_{k=0}^{\infty} \frac{(\lambda_1 + \dots + \lambda_n)^k}{k!} \exp\{-(\lambda_1 + \dots + \lambda_n)\} \cdot k \frac{\lambda_n}{\lambda_1 + \dots + \lambda_n} = \\ &= \lambda_n \exp\{-(\lambda_1 + \dots + \lambda_n)\} \sum_{k=1}^{\infty} \frac{(\lambda_1 + \dots + \lambda_n)^{k-1}}{(k-1)!} = \\ &= \lambda_n \exp\{-(\lambda_1 + \dots + \lambda_n)\} \cdot \exp(\lambda_1 + \dots + \lambda_n) = \lambda_n \end{aligned}$$

2.51. Fie cîmpul de probabilitate $\left([-1, 1], \mathfrak{B}_{[-1, 1]}, \frac{1}{2} m\right)$, m măsura Lebesgue și fie $g: [-1, 1] \rightarrow [0, 1]$ variabila aleatoare definită prin $g(u) = u^2$.

Dacă f este o variabilă aleatoare de medie finită să se arate că $M(f/g)(u) = \frac{1}{2}f(u) + \frac{1}{2}f(-u)$ a.s. Să se găsească apoi o versiune a lui $M(f/g = u)$.

Soluție. Observăm că $\mathfrak{B}_{(0,1)} = \mathfrak{B}((0, t); 0 \leq t \leq 1)$ și deci $\mathfrak{B}(g) = g^{-1}(\mathfrak{B}_{(0,1)}) = \mathfrak{B}(-\sqrt{t}, \sqrt{t}); 0 \leq t \leq 1$ așa că teorema de unicitate a probabilităților ne spune că este suficient să probăm că

$$\int_{-\sqrt{t}}^{\sqrt{t}} f(u) d\left(\frac{1}{2} m\right)(u) = \int_{-\sqrt{t}}^{\sqrt{t}} \left[\frac{1}{2}f(u) + \frac{1}{2}f(-u)\right] d\left(\frac{1}{2} m\right)(u). \quad (1)$$

Avem că

$$\int_{-\sqrt{i}}^{\sqrt{i}} f(u) d\left(\frac{1}{2} m\right)(u) = \frac{1}{2} \int_{\sqrt{i}}^{\sqrt{i}} f(u) du$$

$$\int_{-\sqrt{i}}^{\sqrt{i}} \left[\frac{1}{2} f(u) + \frac{1}{2} f(-u) \right] d\left(\frac{1}{2} m\right)(u) = \frac{1}{4} \int_{-\sqrt{i}}^{\sqrt{i}} f(u) du + \frac{1}{4} \int_{-\sqrt{i}}^{\sqrt{i}} f(-u) du =$$

(prin schimbarea de variabilă $v = -u$ în integrala a doua)

$$= \frac{1}{4} \int_{-\sqrt{i}}^{\sqrt{i}} f(u) du + \frac{1}{4} \int_{-\sqrt{i}}^{\sqrt{i}} f(u) du = \frac{1}{2} \int_{-\sqrt{i}}^{\sqrt{i}} f(u) du$$

deci (1) este adevărată.

Prin definiție $M(f/g = u)$ are proprietatea că $M(f/g) \circ g = \frac{1}{2} f(\cdot) + \frac{1}{2} f(-\cdot)$ a.s. Se observă că $M(f/g = u) = \frac{1}{2} f(\sqrt{u}) + \frac{1}{2} f(-\sqrt{u})$ satisface această cerință.

2.52. Fie f o variabilă aleatoare cu densitatea $\frac{1}{B(r,s)} u^{r-1}(1-u)^{s-1} \lambda_{[0,1]}$ (u) ($B(r,s)$ este funcția Beta) și g o altă variabilă aleatoare cu valori în $\{0, 1, \dots, N\}$ așa încît pentru $k \in \{0, 1, \dots, N\}$ media condițională $M[\lambda_k(g) | f = p]$ este $C_N^k p^k (1-p)^{n-k}$.

Să se afle repartiția lui g . Cînd repartiția lui g este repartiția uniformă pe mulțimea $\{0, 1, \dots, N\}$?

Soluție. Din formula (12) (din rezumatul teoretic) rezultă că

$$P(g = k) = \int_0^1 M[\lambda_k(g) | f = p] \frac{1}{B(r,s)} p^{r-1} (1-p)^{s-1} dp =$$

$$= \int_0^1 C_N^k p^k (1-p)^{N-k} \frac{1}{B(r,s)} p^{r-1} (1-p)^{s-1} dp =$$

$$= \frac{C_N^k}{B(r,s)} \int_0^1 p^{k+r-1} (1-p)^{N+s+k-1} dp = \frac{C_N^k}{B(r,s)} B(k+r, N+s-k) =$$

$$= C_N^k \frac{\Gamma(k+r) \Gamma(N+s-k) \Gamma(s+r)}{\Gamma(r) \Gamma(s) \Gamma(N+s+r)}$$

Se observă că $P(g = k) = \frac{1}{N+1}$ (adică g are repartiția uniformă pe $\{0, 1, \dots, N\}$) dacă $r = s = 1$.

2.53. Fie $f_0 = 0, f_1, \dots, f_n, \dots$ un șir de variabile aleatoare așa încît pentru orice y , media condiționată

$$M[\lambda_{(-\infty, y)}(f_k) | f_{k-1} = x] = U_{x,1}((-\infty, y))$$

1) Să se arate că f_k are densitatea

$$p_k(x) = \frac{(-1)^{k-1}}{(k-1)!} [\ln(1-x)]^{k-1} \lambda_{(0,1)}(x)$$

2) Să se calculeze $M(f_k)$

Soluție. 1) Aplicăm inducția după k . Pentru $k = 1$ afirmația rezultă astfel

$P(f_1 < y) = \int M[\lambda_{(-\infty, y)}(f_1) | f_0 = x] dP \circ f^{-1}(x) = \int U_{x,1}((-\infty, y)) d\varepsilon_0(x) = U_{0,1}((-\infty, y)) = y$, deci f_1 are repartiția $U_{0,1}$, iar p_1 este densitatea de repartiție a lui $U_{0,1}$. Presupunem afirmația adevărată pentru k și să o arătăm pentru $k + 1$. Avem că

$$\begin{aligned} P(f_{k+1} < y) &= \int_0^y M[\lambda_{(-\infty, y)}(f_{k+1}) | f_k = x] \frac{(-1)^{k-1}}{(k-1)!} [\ln(1-x)]^{k-1} dx = \\ &= \frac{(-1)^{k-1}}{(k-1)!} \int_0^y \frac{y-x}{1-x} [\ln(1-x)]^{k-1} dx, \text{ de unde derivînd în raport cu} \end{aligned}$$

y obținem că f_{k+1} are densitatea

$$\begin{aligned} q_{k+1}(y) &= \frac{(-1)^{k-1}}{(k-1)!} \int_0^y \frac{[\ln(1-x)]^{k-1}}{1-x} dx = \frac{(-1)^k}{(k-1)!} \int_0^y \frac{[\ln(1-x)]^{k-1}}{1-x} dx = \\ &= \frac{(-1)^k}{(k-1)!} \frac{[\ln(1-x)]^k}{k} \Big|_0^y = \frac{(-1)^k}{k!} [\ln(1-y)]^k \end{aligned}$$

2) $M(f_k) = \frac{(-1)^{k-1}}{(k-1)!} \int_0^1 x [\ln(1-x)]^{k-1} dx = \frac{(-1)^{k-1}}{(k-1)!} I_k$. Cu substituția $\ln(1-x) = u$ obținem:

$$I_k = \left(1 - \frac{1}{2^k}\right) \int_{-\infty}^0 e^{uu^{k-1}} du = \left(1 - \frac{1}{2^k}\right) (-1)^{k-1} \Gamma(k) =$$

$$= (-1)^{k-1} \left(1 - \frac{1}{2^k}\right) (k-1)!$$

deci

$$M(f_k) = \frac{(-1)^{k-1}}{(k-1)!} I_k = 1 - \frac{1}{2^k}.$$

§ 3. Funcția caracteristică și funcția generatoare

2.54. Să se deducă următoarele proprietăți ale funcțiilor caracteristice:

i₁) $\varphi_f(0) = 1, |\varphi_f(t)| \leq 1$

i₂) $\varphi_f(-t) = \overline{\varphi_f(t)}$

i₃) Dacă $f = (f_1, \dots, f_d)$ este o variabilă aleatoare, A este o matrice $d \times d$, $b \in \mathbb{R}^d$, atunci

$$\varphi_{Af+b}(t) = e^{i\langle t, b \rangle} \varphi_f(A^T t) \quad (A^T \text{ este transpusă lui } A)$$

i₄) φ_f este uniform continuă.

Soluție. i₁) $\varphi_f(0) = M(e^{i\langle 0, f \rangle}) = M(1) = 1$;

$$|\varphi_f(t)| = |M(e^{i\langle t, f \rangle})| \leq M(|e^{i\langle t, f \rangle}|) = M(1) = 1.$$

i₂) $\varphi_f(-t) = \overline{M(e^{i\langle t, f \rangle})} = \overline{M(e^{i\langle t, f \rangle})} = \overline{\varphi_f(t)}$

i₃) $\varphi_{Af+b}(t) = M(e^{i\langle t, Af+b \rangle}) = [M(e^{i\langle Af, t \rangle} e^{i\langle t, b \rangle})] = e^{i\langle t, b \rangle} M(e^{i\langle t, A^T t \rangle}) =$
 $= e^{i\langle t, b \rangle} \varphi_f(A^T t)$

i₄) Avem

$$|\varphi_f(t+h) - \varphi_f(t)| = |M(e^{i\langle t+h, f \rangle}) - M(e^{i\langle t, f \rangle})| \leq M(|e^{i\langle h, f \rangle} - 1|) \leq$$

$$\leq M(\min(2, |h| |f|))$$

deci

$$\limsup_{h \rightarrow 0} \limsup_t |\varphi_f(t+h) - \varphi_f(t)| \leq \lim_{h \rightarrow 0} M(\min(2, |h| |f|)) = 0$$

(ultima egalitate este o consecință a teoremei de convergență dominată), cu alte cuvinte φ_f este uniform continuă.

2.55. Să se determine funcțiile generatoare ale repartițiilor discrete din tabelul 1 și funcțiile caracteristice ale repartițiilor continue din tabelul 2 (vezi rezumatul teoretic de la începutul capitolului).

Soluție.

Repartiția uniformă pe mulțimea $\{1, \dots, n\}$. Funcția generatoare a acestei repartiții este

$$G(z) = \sum_{k=1}^n \frac{1}{n} z^k = \frac{1}{n} \sum_{k=1}^n z^k = \frac{1}{n} z \frac{1-z^n}{1-z},$$

Repartiția geometrică de parametru p . Funcția generatoare este

$$G(z) = \sum_{k=0}^{\infty} (1-p) p^k z^k = (1-p) \sum_{k=0}^{\infty} (pz)^k = (1-p) \frac{1}{1-pz}.$$

Repartiția binomială de parametri n, p . Funcția generatoare este

$$G(z) = \sum_{k=0}^n C_n^k p^k (1-p)^{n-k} z^k = \sum_{k=0}^n C_n^k (pz)^k (1-p)^{n-k} = (pz + 1-p)^n.$$

Repartiția Poisson de parametru λ . Funcția generatoare este

$$G(z) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} z^k = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda z)^k}{k!} = e^{-\lambda} e^{\lambda z} = e^{\lambda(z-1)}.$$

Repartiția uniformă pe (a, b) . Funcția caracteristică este

$$\varphi(t) = \frac{1}{b-a} \int_a^b e^{itx} dx = \frac{1}{b-a} \frac{e^{itb} - e^{ita}}{it}.$$

Repartiția exponențială de parametru λ . Funcția caracteristică este

$$\varphi(t) = \int_0^{\infty} e^{itx} \lambda e^{-\lambda x} dx = \lambda \int_0^{\infty} e^{-\lambda x} \cos tx dx + i\lambda \int_0^{\infty} e^{-\lambda x} \sin tx dx.$$

Integrând prin părți obținem

$$\int_0^{\infty} \lambda e^{-\lambda x} \cos tx dx = 1 - t \int_0^{\infty} e^{-\lambda x} \sin tx dx.$$

de unde integrând încă odată prin părți deducem

$$\int_0^{\infty} \lambda e^{-\lambda x} \cos tx \, dx = 1 - \frac{t^2}{\lambda^2} \int_0^{\infty} \lambda e^{-\lambda x} \cos tx \, dx$$

deci

$$\int_0^{\infty} \lambda e^{-\lambda x} \cos tx \, dx = \frac{\lambda^2}{\lambda^2 + t^2} \text{ și atunci}$$

$$\lambda \int_0^{\infty} e^{-\lambda x} \sin tx \, dx = \lambda \cdot \frac{1 - \frac{\lambda^2}{\lambda^2 + t^2}}{t} = \frac{\lambda t}{\lambda^2 + t^2}$$

Rezultă că

$$\varphi(t) = \frac{\lambda^2}{\lambda^2 + t^2} + \frac{i\lambda t}{\lambda^2 + t^2} = \lambda \frac{\lambda + it}{\lambda^2 + t^2} = \frac{\lambda}{\lambda - it}.$$

Repartiția normală de parametri m, σ^2 . Considerăm întâi cazul $m = 0, \sigma^2 = 1$. Fie f o variabilă aleatoare cu repartiția $N(0,1)$.

Soluție. 1. Observăm întâi că pentru orice $\alpha \in R$ avem că

$$\begin{aligned} M(e^{\alpha|f|}) &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{\alpha|x|} e^{-\frac{x^2}{2}} dx = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{\alpha x - \frac{x^2}{2}} dx = \\ &= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{1}{2}(x-\alpha)^2} dx \cdot e^{\frac{\alpha^2}{2}} = \frac{2}{\sqrt{2\pi}} e^{\frac{\alpha^2}{2}} \int_{-\alpha}^{\infty} e^{-\frac{y^2}{2}} dy \leq \\ &\leq \frac{2}{\sqrt{2\pi}} e^{\frac{\alpha^2}{2}} \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}} dy = 2e^{\frac{\alpha^2}{2}} < \infty. \end{aligned}$$

Mai departe putem scrie

$$M_{2r+1}(f) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^{2r+1} e^{-\frac{x^2}{2}} dx = 0 \text{ (deoarece funcția de integrat este impară)}$$

$$\begin{aligned}
 M_{2r}(f) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^{2r} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^{2r-1} d\left(-e^{-\frac{x^2}{2}}\right) = \\
 &= \frac{1}{\sqrt{2\pi}} \left(-x^{2r-1} e^{-\frac{x^2}{2}}\right) \Big|_{-\infty}^{+\infty} + (2r-1) \int_{-\infty}^{+\infty} x^{2r-2} e^{-\frac{x^2}{2}} dx = (2r-1) M_{2r-2} = \\
 &= (2r-1) \dots 3 \cdot 1 = \frac{(2r)!}{2^r r!},
 \end{aligned}$$

deci f are funcția caracteristică

$$\varphi(t) = M(e^{itf}) = M\left(\sum_{r=0}^{\infty} \frac{(itf)^r}{r!}\right) =$$

(deoarece $\left|\sum_{r=0}^n \frac{(itf)^r}{r!}\right| \leq \sum_{r=1}^{\infty} \frac{|tf|^r}{r!} = e^{|t||f|}$ și $M(e^{|t||f|}) < \infty$, putem aplica teorema de convergență dominată și deci putem permuta seria cu integrala)

rema de convergență dominată și deci putem permuta seria cu integrala)

$$\begin{aligned}
 &= \sum_{r=0}^{\infty} \frac{i^r t^r}{r!} M_r(f) = \sum_{r=0}^{\infty} \frac{i^{2r} t^{2r}}{(2r)!} M_{2r}(f) = \sum_{r=0}^{\infty} (-1)^r \frac{t^{2r}}{(2r)!} \frac{(2r)!}{2^r r!} = \\
 &= \sum_{r=0}^{\infty} (-1)^r \frac{\left(\frac{t^2}{2}\right)^r}{r!} = e^{-\frac{t^2}{2}}
 \end{aligned}$$

Soluția 2. În egalitatea

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{itx - \frac{x^2}{2}} dx$$

derivăm în raport cu t . (teorema de derivare sub integrală funcționează deoarece $\left|e^{itx - \frac{x^2}{2}} ix\right| \leq |x| e^{-\frac{x^2}{2}}$ uniform în t și $|x| e^{-\frac{x^2}{2}}$ este integrabilă).

Obținem

$$\varphi'(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} ix e^{itx - \frac{x^2}{2}} dx$$

de unde integrând prin părți deducem că

$$\begin{aligned} \varphi'(t) &= -\frac{i}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{itx} d\left(e^{-\frac{x^2}{2}}\right) = -\frac{i}{\sqrt{2\pi}} e^{itx} e^{-\frac{x^2}{2}} \Big|_{-\infty}^{+\infty} - \\ &\quad - \frac{t}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{itx - \frac{x^2}{2}} dx = -t\varphi(t) \end{aligned}$$

Prin urmare φ satisface ecuația diferențială

$$\varphi'(t) = -t\varphi(t)$$

a cărei soluție este $\varphi(t) = C e^{-\frac{t^2}{2}}$. Întrucît $\varphi(0) = 1$ rezultă $C = 1$ și prin urmare $\varphi(t) = e^{-\frac{t^2}{2}}$.

Soluția 3. Pentru α real avem

$$M(e^{\alpha x}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{\alpha x - \frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} e^{\frac{\alpha^2}{2}} \int_{-\infty}^{+\infty} e^{-\frac{(x-\alpha)^2}{2}} dx = e^{\frac{\alpha^2}{2}}$$

Fie funcțiile u, v de variabilă complexă z definite prin

$$u(z) = M(e^{z^2}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{zx - \frac{x^2}{2}} dx; \quad v(z) = e^{\frac{z^2}{2}}$$

Evident v este olomorfă. Arătăm pentru început că u este

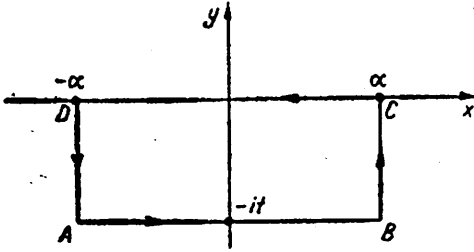
continuă. Fie $z_n \rightarrow z$; atunci $e^{\frac{z_n x - x^2}{2}} \rightarrow e^{\frac{z x - x^2}{2}}$ și cum $\left| e^{\frac{z_n x - x^2}{2}} \right| \leq e^{\frac{|z_n| |x| - x^2}{2}} \leq e^{\frac{\sup |z_n| |x| - x^2}{2}} \leq e^{\frac{\sup |z_n| |x| - x^2}{2}}$ este integrabilă, rezultă din teorema de convergență dominată că $u(z_n) \rightarrow u(z)$, deci u este continuă.

Dacă Γ este un contur închis atunci avem

$$\int_{\Gamma} u(z) dz = \int_{\Gamma} \left(\int_{\Gamma} e^{z^2} dP \right) dz = \int_{\Gamma} \left(\int_{\Gamma} e^{z^2} dz \right) dP = 0 \text{ deoarece } e^{z^2} \text{ fiind olomorfă, rezultă din teorema lui Fubini)}$$

morfă, rezultă din teorema lui Cauchy că $\int_{\Gamma} e^{z^2} dz = 0$.

Funcția u fiind continuă și integrala sa pe orice contur închis fiind nulă rezultă din teorema lui Morera că u este olomorfă. Cum am arătat că u și v coincid pe R rezultă din teorema de unicitate a funcțiilor olomorfe că $u = v$ pe C . În particular luând $z = it$ obținem că $u(it) = \varphi(t) = e^{-\frac{t^2}{2}}$



nem că $u(it) = \varphi(t) = e^{-\frac{t^2}{2}}$

Soluția 4. Avem că

$$\begin{aligned} \varphi(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{ix - \frac{x^2}{2}} dx = \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \int_{-\infty}^{+\infty} e^{-\frac{(x-it)^2}{2}} dx \end{aligned}$$

Fie conturul închis din figura alăturată

Avem deci $\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \int_{\text{Im } z = -it} e^{-\frac{z^2}{2}} dz$. Cum funcția $e^{-\frac{z^2}{2}}$ este olomorfă rezultă din teorema lui Cauchy că $\int_{ABCD} e^{-\frac{z^2}{2}} dz = 0$

$$\int_{ABCD} e^{-\frac{z^2}{2}} dz = 0$$

În continuare avem

$$\begin{aligned} \left| \frac{1}{\sqrt{2\pi}} \int_{BC} e^{-\frac{z^2}{2}} dz \right| &= \left| \frac{1}{\sqrt{2\pi}} \int_{-t}^0 e^{-\frac{(\alpha+iu)^2}{2}} du \right| \leq \frac{e^{-\frac{\alpha^2}{2}}}{\sqrt{2\pi}} \int_0^{|t|} \left| e^{\frac{u^2}{2} - i\alpha u} \right| du = \\ &= \frac{e^{-\frac{\alpha^2}{2}}}{\sqrt{2\pi}} \int_0^{|t|} e^{-\frac{u^2}{2}} du \rightarrow 0 \text{ cînd } \alpha \rightarrow \infty. \end{aligned}$$

Se procedează analog pentru segmentul DA . Rezultă

$$0 = \lim_{\alpha \rightarrow \infty} \int_{ABCD} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \int_{\text{Im } z = -it} e^{-\frac{z^2}{2}} dz + \frac{1}{\sqrt{2\pi}} \int_{+\infty}^{-\infty} e^{-\frac{z^2}{2}} dz$$

de unde

$$\frac{1}{\sqrt{2\pi}} \int_{\text{Im } z = -it} e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} dz = 1$$

și prin urmare

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \int_{\text{Im } z = -it} e^{-\frac{z^2}{2}} dz = e^{-\frac{t^2}{2}}$$

Fie acum g o variabilă aleatoare cu repartiția $N(m, \sigma^2)$. Atunci $f = \frac{g-m}{\sigma}$ are repartiția $N(0, 1)$ și $g = \sigma f + m$. Rezultă că repartiția $N(m, \sigma)$ are funcția caracteristică

$$\varphi_g(t) = e^{itm} \varphi_f(\sigma t) = e^{itm} e^{-\frac{\sigma^2 t^2}{2}} = e^{itm - \frac{\sigma^2 t^2}{2}}$$

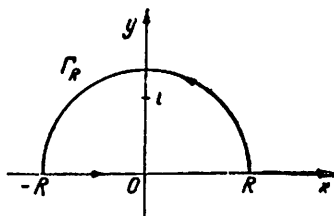
Repartiția Cauchy de parametri α, β . Presupunem întâi că $\alpha = 1, \beta = 0$.

Soluția 1. Funcția caracteristică a lui $C_{1,0}$ este dată de

$$\varphi(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} e^{itx} \frac{1}{1+x^2} dx$$

Presupunem întâi că $t > 0$. Fie funcția $u(z) = \frac{e^{itz}}{1+z^2}$ și conturul închis din figura alăturată. Din teorema rezidurilor rezultă că

$$(1) \int_{-R}^R u(z) dz + \int_{\Gamma_R} u(z) dz = \\ = \text{Rez}(u, i) \cdot 2\pi i$$



Pentru $z \in \Gamma_R$ avem că $\left| \frac{1}{1+z^2} \right| \leq$

$\leq \frac{1}{R^2-1} \rightarrow 0$ când $R \rightarrow \infty$, deci conform lemei lui Jordan avem că

$$\lim_{R \rightarrow \infty} \int_{\Gamma} u(z) dz = 0$$

Făcând $R \rightarrow \infty$ în (1) obținem

$$\int_{-\infty}^{+\infty} u(z) dz = \text{Rez}(u, i) \cdot 2\pi i = 2\pi i \cdot \frac{1}{2i} e^{-t} = \pi e^{-t} = \pi e^{-|t|}$$

și prin urmare

$$\varphi(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} u(z) dz = e^{-|t|} \quad \text{dacă } t > 0$$

Dacă $t < 0$ atunci $\varphi(t) = \varphi(-t) = e^{-|t|}$ și deci $\varphi(t) = e^{-|t|}$ pentru orice $t \in \mathbb{R}$.

Soluția 2. Avem

$$\varphi(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{e^{itx}}{1+x^2} dx = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{\cos tx}{1+x^2} dx$$

de unde luând $t > 0$, derivând în raport cu t și având în vedere egalitatea

tatea $\int_{-\infty}^{+\infty} \frac{\sin tx}{x} dx = \pi$ (teorema de derivare sub integrală funcțio-

nează deoarece din calculul ce urmează avem că $\left| \frac{\sin tx}{x(1+x^2)} \right| \leq \frac{|t|}{1+x^2}$

și $\frac{1}{1+x^2}$ este integrabilă), vom obține

$$\begin{aligned} \varphi'(t) + 1 &= \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{-x \sin tx}{1+x^2} dx + 1 = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{-x \sin tx}{1+x^2} dx + \\ &+ \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{\sin tx}{x} dx = \int_{-\infty}^{+\infty} \frac{1}{1+x^2} \frac{\sin tx}{x} dx \end{aligned}$$

Reșinând primul și ultimul dintre termenii acestor egalități și derivând încă odată în raport cu t obținem următoarea ecuație diferențială

$$\varphi''(t) = \varphi(t)$$

Soluția acestei ecuații este $\varphi(t) = C_1 e^t + C_2 e^{-t}$, $t > 0$. Cum φ este mărginită rezultă că $C_1 = 0$ și din $\varphi(0) = 1$ rezultă că $C_2 = 1$, așa că pentru $t > 0$ obținem $\varphi(t) = e^{-t} = e^{-|t|}$.

Fie acum f o variabilă aleatoare cu repartiția $C_{\alpha,\beta}$. Atunci variabila aleatoare $g = \frac{f-\beta}{\alpha}$ are repartiția $C_{1,0}$ și $f = g\alpha + \beta$

Rezultă că

$$\varphi_f(t) = e^{-it\beta} \varphi_g(\alpha t) = e^{it\beta} e^{-\alpha|t|} = e^{it\beta - \alpha|t|}$$

Repartiția Gamma de parametri α, β . Funcția caracteristică este

$$\varphi(t) = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{itx} x^{\alpha-1} e^{-\beta x} dx$$

Fie $u: \{\text{Re}z < \beta\} \rightarrow \mathbb{C}$ definită prin $u(z) = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{zx} x^{\alpha-1} e^{-\beta x} dx$. Dacă

$$\begin{aligned} \lambda < \beta \text{ rezultă că } u(\lambda) &= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{-(\beta-\lambda)x} x^{\alpha-1} dx = \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{-y} \frac{y^{\alpha-1}}{(\beta-\lambda)^{\alpha-1}} \frac{dy}{\beta-\lambda} = \frac{\beta^\alpha}{\Gamma(\alpha)(\beta-\lambda)^\alpha} \int_0^\infty y^{\alpha-1} e^{-y} dy = \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)(\beta-\lambda)^\alpha} \Gamma(\alpha) = \frac{\beta^\alpha}{(\beta-\lambda)^\alpha} \end{aligned}$$

Deoarece integrantul, din definiția lui u este continuu în z și $|e^{zx} x^{\alpha-1} e^{-\beta x}| \leq e^{\text{Re}z x} x^{\alpha-1} e^{-\beta x} \leq e^{kx} x^{\alpha-1} e^{-\beta x} = e^{-(\beta-k)x} x^{\alpha-1}$ dacă $\text{Re}z < k < \beta$, unde $e^{-(\beta-k)x} x^{\alpha-1}$ este integrabilă, rezultă din teorema de convergență dominată că u este continuă.

Dacă Γ este un contur închis în $\text{Re}z < \beta$ atunci prin utilizarea teoremei lui Fubini și a teoremei lui Cauchy (e^{zx} este olomorvă în z) obținem că

$$\int_{\Gamma} u(z) dz = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-1} e^{-\beta x} dx \left(\int_{\Gamma} e^{zx} dz \right) = 0$$

cu alte cuvinte u este olomorvă conform teoremei lui Morera.

Cum am arătat mai sus că u coincide cu funcția olomorvă pe domeniul $(\alpha > \beta) \frac{\beta^\alpha}{(\beta-z)^\alpha}$ rezultă din teorema de unicitate a funcțiilor olomorfe că cele două funcții coincid pe $\text{Re}z < \beta$.

În particular pentru $z = it$ obținem că $\varphi(t) = \beta^\alpha (\beta - it)^{-\alpha}$

2.56. Să se determine funcțiile caracteristice pentru repartițiile multidimensionale listate în rezumatul teoretic de la începutul capitolului.

Soluție.

Repartiția Dirac în punctul $a \in R^d$. Funcția caracteristică este

$$\varphi(t) = \int_{R^d} e^{i\langle t, x \rangle} d\varepsilon_a(x) = e^{i\langle t, a \rangle}.$$

Repartiția multidimensională de parametri n, p_1, \dots, p_d . Funcția caracteristică este

$$\begin{aligned} \varphi(t) &= \sum_{\substack{n_k \geq 0 \\ n_1 + \dots + n_d = n}} \frac{n!}{n_1! \dots n_d!} p_1^{n_1} \dots p_d^{n_d} e^{i(t_1 n_1 + \dots + t_d n_d)} = \\ &= \sum_{\substack{n_k \geq 0 \\ n_1 + \dots + n_d = n}} \frac{n!}{n_1! \dots n_d!} [p_1 e^{it_1}]^{n_1} \dots [p_d e^{it_d}]^{n_d} = \left[\sum_{k=1}^d p_k e^{it_k} \right]^n. \end{aligned}$$

Repartiția uniformă pe mulțimea $D = (a_1, b_1) \times \dots \times (a_d, b_d)$. Funcția caracteristică este

$$\varphi(t) = \frac{1}{(b_1 - a_1) \dots (b_d - a_d)} \int_{a_1}^{b_1} \dots \int_{a_d}^{b_d} e^{i(t_1 x_1 + \dots + t_d x_d)} dx_1 \dots dx_d =$$

(teorema lui Fubini)

$$\begin{aligned} &= \left[\frac{1}{(b_1 - a_1)} \int_{a_1}^{b_1} e^{it_1 x_1} dx_1 \right] \dots \left[\frac{1}{(b_d - a_d)} \int_{a_d}^{b_d} e^{it_d x_d} dx_d \right] = \\ &= \left[\frac{1}{(b_1 - a_1)} \frac{e^{it_1 b_1} - e^{it_1 a_1}}{it_1} \right] \dots \left[\frac{1}{(b_d - a_d)} \frac{e^{it_d b_d} - e^{it_d a_d}}{it_d} \right]. \end{aligned}$$

Repartiția normală d-dimensională de parametri m, B . Presupunem pentru început că $m = 0$. Fie $\lambda_1, \dots, \lambda_d$ valorile proprii ale lui B (sînt pozitive deoarece B este pozitiv definită) și fie H o

matrice ortogonală ($H^T = H^{-1}$) așa încît $HBH^T = \begin{pmatrix} \lambda_1 & & 0 \\ & \dots & \\ 0 & & \lambda_d \end{pmatrix} =$

Din problema 2.27 rezultă că vectorul $g = Hf$ are componentele g_1, \dots, g_d independente și cu repartițiile $N(0, \lambda_1) \dots N(0, \lambda_d)$. Putem deci scrie că $\varphi_g(t) = \varphi_{g_1}(t_1) \dots \varphi_{g_d}(t_d) = e^{-\frac{1}{2}t_1^2 \lambda_1} \dots e^{-\frac{1}{2}t_d^2 \lambda_d} = e^{-\frac{1}{2}\langle \Lambda t, t \rangle}$

Rezultă că

$$\varphi_f(t) = \varphi_{H^T g}(t) = \varphi_g(Ht) = e^{-\frac{1}{2}\langle \Lambda Ht, Ht \rangle} = e^{-\frac{1}{2}\langle H^T \Lambda H, t \rangle} = e^{-\frac{1}{2}\langle Bt, t \rangle}$$

deoarece $H^T \Lambda H = B$.

Dacă $m = 0$ atunci dacă f este o variabilă aleatoare cu repartiția $N(m, B)$ rezultă că variabila aleatoare $g = f - m$ are repartiția $N(0, B)$ și deci avem că

$$\varphi_f(t) = \varphi_{g+m}(t) = e^{i\langle t, m \rangle} \varphi_g(t) = e^{i\langle t, m \rangle - \frac{1}{2}\langle Bt, t \rangle}$$

2.57. Fie f_1 și f_2 variabile aleatoare independente. Folosindu-se funcțiile caracteristice să se arate că:

- dacă f_i are repartiția ε_{a_i} , atunci $f_1 + f_2$ are repartiția $\varepsilon_{a_1 + a_2}$;
- dacă f_i are repartiția $B_{n_i}(\rho)$, atunci $f_1 + f_2$ are repartiția $B_{n_1 + n_2}(\rho)$;
- dacă f_i are repartiția π_{λ_i} , atunci $f_1 + f_2$ are repartiția $\pi_{\lambda_1 + \lambda_2}$;
- dacă f_i are repartiția $N(m_i, \sigma_i^2)$ atunci $f_1 + f_2$ are repartiția $N(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$;
- dacă f_i are repartiția C_{α_i, β_i} , atunci $f_1 + f_2$ are repartiția $C_{\alpha_1 + \alpha_2, \beta_1 + \beta_2}$;
- dacă f_i are repartiția $\gamma(\alpha_i, \beta)$, atunci $f_1 + f_2$ are repartiția $\gamma(\alpha_1 + \alpha_2, \beta)$;
- dacă f_i are repartiția $M_{n_i}(\rho_1, \dots, \rho_n)$, atunci $f_1 + f_2$ are repartiția $M_{n_1 + n_2}(\rho_1, \dots, \rho_n)$.
- Dacă f_i are repartiția $N_d(m_i, B_i)$, atunci $f_1 + f_2$ are repartiția $N_d(m_1 + m_2, B_1 + B_2)$.

Soluție. Din independență rezultă că $\varphi_{f_1 + f_2}(t) = \varphi_{f_1}(t) \varphi_{f_2}(t)$. Deoarece funcția caracteristică determină în mod unic repartiția, ob-

servăm că este suficient să verificăm că $\varphi_{f_1+f_2}$ coincide cu funcția caracteristică a repartițiilor care se specifică în enunț.

$$a) \varphi_{f_1+f_2}(t) = e^{i\langle t, a_1 \rangle} e^{i\langle t, a_2 \rangle} = e^{i\langle t, a_1+a_2 \rangle}$$

Deoarece avem egalitatea $\varphi_f(t) = G_f(e^{it})$ rezultă că pentru repartiția $B_n(p)$ și π_λ funcțiile caracteristice sînt

$$[1 + p(e^{it} - 1)]^n \text{ și } e^{\lambda(e^{it} - 1)}$$

$$b) \varphi_{f_1+f_2}(t) = [1 + p(e^{it} - 1)]^{n_1} [1 + p(e^{it} - 1)]^{n_2} = [1 + p(e^{it} - 1)]^{n_1+n_2}.$$

$$c) \varphi_{f_1+f_2}(t) = e^{\lambda_1(e^{it} - 1)} e^{\lambda_2(e^{it} - 1)} = e^{(\lambda_1+\lambda_2)(e^{it} - 1)}.$$

$$d) \varphi_{f_1+f_2}(t) = e^{im_1t - \frac{\sigma_1^2 t^2}{2}} e^{im_2t - \frac{\sigma_2^2 t^2}{2}} = e^{i(m_1+m_2)t - (\sigma_1^2 + \sigma_2^2) \frac{t^2}{2}}.$$

$$e) \varphi_{f_1+f_2}(t) = e^{-\alpha_1|t| + i\beta_1 t} e^{-\alpha_2|t| + i\beta_2 t} = e^{-(\alpha_1+\alpha_2)|t| + i(\beta_1+\beta_2)t}.$$

$$f) \varphi_{f_1+f_2}(t) = \beta^{\alpha_1} (\beta - it)^{-\alpha_1} \cdot \beta^{\alpha_2} (\beta - it)^{-\alpha_2} = \beta^{\alpha_1+\alpha_2} (\beta - it)^{-(\alpha_1+\alpha_2)}.$$

$$g) \varphi_{f_1+f_2}(t) = (p_1 e^{it_1} + \dots + p_d e^{it_d})^{n_1} (p_1 e^{it_1} + \dots + p_d e^{it_d})^{n_2} = (p_1 e^{it_1} + \dots + p_d e^{it_d})^{n_1+n_2}.$$

$$h) \varphi_{f_1+f_2}(t) = e^{i\langle t, m_1 \rangle - \frac{1}{2} \langle B_1 t, t \rangle} \cdot e^{i\langle t, m_2 \rangle - \frac{1}{2} \langle B_2 t, t \rangle} = e^{i\langle t, m_1+m_2 \rangle - \frac{1}{2} \langle (B_1+B_2)t, t \rangle}.$$

2.58. Să se deducă următoarele proprietăți ale repartiției normale d -dimensionale :

i₁) Dacă $f = (f_1, \dots, f_d)$ are repartiția $N_d(m, A)$ și $s < d$ atunci $g = (f_1, \dots, f_s)$ are repartiția $N_s(n, B)$, unde $n = (m_1, \dots, m_s)$, $B = (a_{i,j})_{1 \leq i, j \leq s}$.

i₂) Dacă $f = (f_1, \dots, f_d)$ are repartiția $N_d(m, A)$, $g = (g_1, \dots, g_s)$ are repartiția $N_s(n, B)$ și f, g sînt independente atunci (f, g) are repartiția $N_{d+s}(\bar{m}, \bar{A})$ unde

$$\bar{m} = (m, n); \quad \bar{A} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

i₃) Dacă $(f, g) = (f_1, \dots, f_d, g_1, \dots, g_s)$ are repartiția normală și pentru orice i, j variabilele aleatoare f_i, g_j verifică formula de multiplicare, să se arate că f și g sînt independente.

i₄) Dacă f are repartiția $N_d(m, A)$ și $u: R^d \rightarrow R^{d'}$ este aplicația definită prin $u(x) = a + Cx$ să se arate că $u(f) \in N_{d'}(a + Cm, CBC^T)$, C^T fiind transpusa lui C .

Soluție. $i_1)$ Avem $\varphi_g(t') = \varphi_g(t_1, \dots, t_s) = \varphi_f(t_1, \dots, t_s, 0, \dots, 0) \equiv$
 $= \exp\left(i \sum_{k=1}^s m_k t_k - 1/2 \sum_{k,j=1}^s a_{k,j} t_k t_j\right) = \exp(i \langle n, t' \rangle - 1/2 \langle B t', t' \rangle)$

deci $g \in N_s(n, B)$ (funcția caracteristică determină în mod unic repartiția).

$i_2)$ Fie $\bar{t} = (t_1, \dots, t_d, t_{d+1}, \dots, t_{d+s})$, $t' = (t_1, \dots, t_d)$, $t'' = (t_{d+1}, \dots, t_{d+s})$. Din independența lui f și g rezultă că $\varphi_{(f,g)}(\bar{t}) = \varphi_f(t') \varphi_g(t'')$ (vezi egalitatea (18) din rezumatul teoretic), așa că $\varphi_{(f,g)}(\bar{t}) = \exp(i \langle t', m \rangle - 1/2 \langle A t', t' \rangle) \exp(i \langle t'', n \rangle - 1/2 \langle B t'', t'' \rangle) = \exp(i \langle \bar{t}, \bar{m} \rangle - 1/2 \langle \bar{A} \bar{t}, \bar{t} \rangle)$.

deci (f, g) are repartiția $N_{d+s}(\bar{m}, \bar{A})$.

$i_3)$ Vom arăta că $\varphi_{(f,g)}(t', t'') = \varphi_f(t') \varphi_g(t'')$. Fie $N_{d+s}(\bar{m}, \bar{A})$ repartiția vectorului (f, g) , $\bar{m} = (m, n)$, $m = (m_1, \dots, m_d)$, $n = (m_{d+1}, \dots, m_{d+s})$, $A = (\bar{a}_{i,j})_{1 \leq i,j \leq d}$, $B = (\bar{a}_{i,j})_{d+1 \leq i,j \leq d+s}$. Din independența lui f_i, g_j rezultă că $M(f_i g_j) = M(f_i) M(g_j)$ cu alte cuvinte $\bar{a}_{i,j} = 0$.

Forma matricii \bar{A} este următoarea $\bar{A} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$. Pe de altă parte din punctul $i_2)$ rezultă că f are repartiția $N_s(m, A)$, g are repartiția $N_s(n, B)$. De asemenea putem scrie

$$\begin{aligned} \varphi_{(f,g)}(t', t'') &= \exp(i \langle m, t' \rangle + i \langle n, t'' \rangle - 1/2 \langle A t', t' \rangle - \\ &- 1/2 \langle B t'', t'' \rangle) = \exp(i \langle m, t' \rangle - \frac{1}{2} \langle A t', t' \rangle) \exp(i \langle n, t'' \rangle - \\ &- \frac{1}{2} \langle B t'', t'' \rangle) = \varphi_f(t') \varphi_g(t'') \end{aligned}$$

$i_4)$ $\varphi_{u(f)}(v) = \exp(i \langle v, a \rangle) \varphi_f(C^T v) = \exp(i \langle v, a \rangle) \exp(i \langle C^T v, m \rangle - \frac{1}{2} \langle A C^T v, C^T v \rangle) = \exp(i \langle v, a + C m \rangle - \frac{1}{2} \langle C A C^T v, v \rangle)$, deci $u(f)$ are repartiția $N_d(a + C m, C A C^T)$.

2.59. Să se deducă faptul că următoarele funcții

$\cos^2 t, \sum_{n=0}^{\infty} a_n \cos nt, \left((a_n > 0) \sum_{n=0}^{\infty} a_n = 1 \right), e^{\lambda(e^t - 1)} (\lambda > 0), e^{-|t|} \sin t$ sunt funcții caracteristice.

Soluție. Avem

$$\cos^2 t = \left(\frac{e^{it} + e^{-it}}{2} \right)^2 = \frac{1}{4} e^{-2it} + \frac{1}{2} + \frac{1}{4} e^{2it}$$

deci $\cos^2 t$ este funcția caracteristică a repartiției discrete

$$F = \frac{1}{4} \varepsilon_{-2} + \frac{1}{2} \varepsilon_0 + \frac{1}{4} \varepsilon_2.$$

Mai departe: $\sum_{n=0}^{\infty} a_n \cos nt = \sum_{n=0}^{\infty} \left(\frac{a_n}{2} e^{int} + \frac{a_n}{2} e^{-int} \right)$, deci $\sum_{n=0}^{\infty} a_n \cos nt$ este funcția caracteristică a repartiției discrete

$$F = a_0 \varepsilon_0 + \sum_{n=1}^{\infty} \frac{a_n}{2} (\varepsilon_n + \varepsilon_{-n})$$

În continuare putem scrie

$$e^{\lambda(e^{it}-1)} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} e^{int} = \sum_{n=0}^{\infty} e^{int} \left(\frac{\lambda^n}{n!} e^{-\lambda} \right)$$

fapt ce ne arată că $\varepsilon^{\lambda(e^{it}-1)}$ este funcția caracteristică a repartiției discrete $F = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \varepsilon_n$, deci a repartiției π_λ .

Pentru ultima funcție vom folosi teorema de inversiune. Deoarece $\int_{\mathbb{R}} e^{-|t|} dt < \infty$ rezultă că putem defini

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} e^{-|t|} dt$$

Avem

$$p(x) = \frac{1}{2\pi} \int_0^{\infty} e^{-t}(e^{itx} + e^{-itx}) dt = \frac{1}{\pi} \int_0^{\infty} e^{-t} \cos tx dt = -\frac{1}{\pi} x^2 p(x)$$

deci $p(x) = \frac{1}{\pi(1+x^2)}$, cu alte cuvinte p este densitatea repartiției $C_{1,0}$ și deci $e^{-|t|}$ este funcția caracteristică a repartiției $C_{1,0}$

2.60. Variabilele aleatoare f, g, h sînt independente și $P(h=0) = p$, $P(h+1) = 1-p$. $0 < p < 1$.

Să se afle funcția caracteristică a variabilei $u = hf + (1 - h)g$.

Soluție. Putem scrie că

$$\begin{aligned}\varphi_u(t) &= M(e^{itu}) \stackrel{(11)}{=} M(e^{ihu}/h = 0) P(h = 0) + M(e^{itu}/h = 1) P(h = 1) = \\ &= M(e^{iut}/h = 0) p + M(e^{iut}/h = 1)(1 - p) \stackrel{(13)}{=} M(e^{iug}) p + M(e^{iuf})(1 - p) = \\ &= p\varphi_g(t) + (1 - p)\varphi_f(t)\end{aligned}$$

2.61. Să se determine $\alpha, \beta, \gamma, \delta$ așa încît funcția $G(z) = \frac{\alpha + \beta z}{\gamma + \delta z}$ să fie funcția generatoare a unei variabile aleatoare.

Soluție. Constantele $\alpha, \beta, \gamma, \delta$ le vom determina din următoarele condiții :

a) $G(1) = 1$

b) G este funcție analitică în $|z| \leq 1$ și în dezvoltarea sa în serie de puteri coeficienții sînt nenegativi.

Cazul 1. $\gamma = 0$ Rezultă că $\delta \neq 0$ și pentru ca G să fie analitică trebuie ca $\alpha = 0$ cu alte cuvinte obținem că $G(z) = \frac{\beta}{\delta}$ și cum $G(1) = 1$ rezultă că $\frac{\beta}{\delta} = 1$, deci $G(z) \equiv 1$, situație care corespunde unei variabile aleatoare f egală cu 0 a.s.

Cazul 2. $\gamma \neq 0$. În acest caz condiția $G(1) = 1$ dă relația

$$1 = \frac{\alpha + \beta}{\gamma + \delta} = \frac{\frac{\alpha}{\gamma} + \frac{\beta}{\gamma}}{1 + \frac{\delta}{\gamma}}, \quad (1)$$

Apoi putem scrie

$$G(z) = \frac{\frac{\alpha}{\gamma} + \frac{\beta}{\gamma}}{1 + \frac{\delta}{\gamma}} = \left(\frac{\alpha}{\gamma} + \frac{\beta}{\gamma} z \right) \cdot \frac{1}{1 + \frac{\delta}{\gamma}}$$

Or $\frac{1}{1 - \frac{\delta}{\gamma}}$ este analitică pentru $\left| \frac{\delta}{\gamma} z \right| < 1$ și cum $|z| \leq 1$ rezultă

că $\left| \frac{\delta}{\gamma} \right| < 1$ și în acest caz $\frac{1}{1 + \frac{\delta}{\gamma}} = \sum_{n=0}^{\infty} (-1)^n \left(\frac{\delta}{\gamma} \right)^n z^n$. Rezultă că

$$G(z) = \left(\frac{\alpha}{\gamma} + \frac{\beta}{\gamma} z \right) \sum_{n=0}^{\infty} (-1)^n \left(\frac{\delta}{\gamma} \right)^n z^n = \frac{\alpha}{\gamma} - \left(\frac{\alpha\delta}{\gamma^2} - \frac{\beta}{\gamma} \right) z + \frac{\delta}{\gamma} \left(\frac{\alpha\delta}{\gamma^2} - \frac{\beta}{\gamma} \right) z^2 - \\ - \frac{\delta^2}{\gamma^2} \left(\frac{\alpha\delta}{\gamma^2} - \frac{\beta}{\gamma} \right) z^3 + \dots$$

Coeficienții dezvoltării trebuind să fie nenegativi obținem următoarele condiții:

$$\frac{\alpha}{\gamma} = a, \quad 0 \leq a \leq 1 \quad (2)$$

$$\frac{\alpha\delta}{\gamma^2} - \frac{\beta}{\gamma} \leq 0 \quad (3)$$

$$\frac{\delta}{\gamma} \leq 0 \text{ și cum } \left| \frac{\delta}{\gamma} \right| < 1 \text{ rezultă că } \frac{\delta}{\gamma} = -d, \quad 0 \leq d < 1. \quad (4)$$

Din condiția (1) rezultă că

$$\frac{\beta}{\gamma} = 1 - d - a \quad (5)$$

În fine se observă că (3) rezultă de îndată din (2), (4), (5), deci dacă $\gamma \neq 0$ atunci G este funcție generatoare a unei variabile aleatoare dacă sînt îndeplinite condițiile (2), (4), (5).

2.62. Variabila aleatoare f cu valori întregi nenegative are funcția generatoare $G(z)$. Pentru $b > a > 0$ să se deducă egalitățile

$$M\left(\frac{1}{f+a}\right) = \int_0^1 z^{a-1} G(z) dz \quad (1)$$

$$M\left(\frac{1}{(f+a)(f+b)}\right) = \int_0^1 z^{b-a-1} \int_0^z u^{a-1} G(u) du dz \quad (2)$$

Soluție. Prin utilizarea egalității $\int_0^1 z^\alpha dz = \frac{1}{\alpha+1}$, $\alpha > -1$ obținem

$$M\left(\frac{1}{f+a}\right) = M\left(\frac{1}{a} \int_0^1 z^{\frac{f}{a}} dz\right) = \frac{1}{a} \int_0^1 M\left(z^{\frac{f}{a}}\right) dz = \frac{1}{a} \int_0^1 G_f\left(z^{\frac{1}{a}}\right) dz =$$

(substituția $z = u^a$)

$$= \int_0^1 u^{a-1} G(u) du, \text{ deci tocmai (1)}$$

Mai departe avem

$$\frac{1}{(f+a)(f+b)} = \frac{1}{(b-a)(f+a)} - \frac{1}{(b-a)(f+b)}$$

$$M\left[\frac{1}{(f+a)(f+b)}\right] = \frac{1}{b-a} M\left(\frac{1}{f+a}\right) - \frac{1}{b-a} M\left(\frac{1}{f+b}\right) =$$

$$= \frac{1}{b-a} \int_0^1 z^{a-1} G(z) dz - \frac{1}{b-a} \int_0^1 z^{b-1} G(z) dz$$

și prin aplicarea teoremei lui Fubini obținem

$$\int_0^1 z^{b-a-1} \int_0^z u^{a-1} G(u) du dz = \int_0^1 u^{a-1} G(u) du \int_u^1 z^{b-a-1} dz =$$

$$= \int_0^1 u^{a-1} G(u) \left(\frac{1}{b-a} - \frac{u^{b-a}}{b-a}\right) du = \frac{1}{b-a} \int_0^1 u^{a-1} G(u) du -$$

$$- \frac{1}{b-a} \int_0^1 u^{b-1} G(u) du,$$

deci (2) este adevărată.

2.63. Fie f o variabilă aleatoare cu valori în $\{\dots, -1, 0, 1, \dots\}$ și cu funcția caracteristică $\varphi(t)$. Să se calculeze

- 1) $P(f = 0 \pmod{k})$
- 2) $P(f = m \pmod{k})$

Soluție. 1) Pornind de la egalitatea $\sum_{j=0}^{k-1} e^{2\pi n j i} = k$ putem scrie

$$P(f = 0 \pmod{k}) = \sum_{n=-\infty}^{+\infty} P(f = nk) = \sum_{n=-\infty}^{+\infty} \left(\frac{1}{k} \sum_{j=0}^{k-1} e^{2\pi n j i} \right) P(f = nk) = \\ = \frac{1}{k} \sum_{j=0}^{k-1} \sum_{n=-\infty}^{+\infty} e^{2\pi n j i} P(f = nk) = \frac{1}{k} \sum_{j=0}^{k-1} \varphi \left(\frac{2\pi j}{k} \right)$$

2) Fie $g = f - m$. Atunci $\varphi_g(t) = e^{-m i} \varphi_f(t)$ și!

$$P(f = m \pmod{k}) = P(g = 0 \pmod{k}) = \frac{1}{k} \sum_{j=0}^{k-1} \varphi_g \left(\frac{2\pi j}{k} \right) = \\ = \frac{1}{k} \sum_{j=0}^{k-1} \varphi_g \left(\frac{2\pi j}{k} \right) e^{-\frac{2\pi j m i}{k}}$$

2.64. Fie f_1, \dots, f_n un șir de variabile aleatoare independente cu repartiția uniformă pe mulțimea $\{1, 2, \dots, 3k\}$. Să se afle

$$\max_{n \geq 1} P(f_1^2 + \dots + f_n^2 \equiv 0 \pmod{3}); \quad \min_{n \geq 1} P(f_1^2 + \dots + f_n^2 \equiv 0 \pmod{3})$$

Soluție. Fie r_i restul împărțirii lui f_i^2 prin 3. Observăm că f_i^2 nu poate fi decît de forma $3m, 3m + 1$, deci putem scrie că

$$r_i = \begin{cases} 0 & f_i^2 = 3m \\ 1 & f_i^2 = 3m + 1 \end{cases}$$

De asemenea $P(r_i = 0) = \frac{1}{3}$, $P(r_i = 1) = \frac{2}{3}$. Rezultă că funcția caracteristică a lui r_i este $\varphi_{r_i}(t) = \frac{1}{3} + \frac{1}{3} e^{it}$ și cum r_1, \dots, r_n sînt independente ($r_i = u(f_i)$ și se aplică asociativitatea independenței) rezultă că $\varphi_{r_1 + \dots + r_n}(t) = \frac{1}{3^n} (1 + 2 e^{it})^n$. Folosind rezultatul din problema precedentă putem scrie

$$P(f^2 + \dots + f_n^2 \equiv 0 \pmod{3}) = P(r_1 + \dots + r_n \equiv 0 \pmod{3}) = \\ = \frac{1}{3} \left[\varphi_{r_1 + \dots + r_n}(0) + \varphi_{r_1 + \dots + r_n} \left(\frac{2\pi}{3} \right) + \varphi_{r_1 + \dots + r_n} \left(\frac{4\pi}{3} \right) \right] = \\ = \frac{1}{3} \left[1 + \frac{1}{3^n} \left(1 + 2e^{\frac{2\pi i}{3}} \right)^n + \frac{1}{3^n} \left(1 + 2e^{\frac{3\pi i}{3}} \right)^n \right] = \frac{1}{3} \left[1 + \frac{i^n}{3^n} (\sqrt{3})^n + \right. \\ \left. + \frac{(-i)^n}{3^n} (\sqrt{3})^n \right]$$

În fine se obține ușor că

$$\max_{n \geq 1} P(f_1^2 + \dots + f_n^2 \equiv 0 \pmod{3}) = \frac{11}{27}$$

$$\min_{n \geq 1} P(f_1^2 + \dots + f_n^2 \equiv 0 \pmod{3}) = \frac{1}{9}$$

2.65. Fie f o variabilă aleatoare. Să se arate că:

1) Dacă există doi întregi nenegativi $m < p$ așa încît

$$P(f \equiv m \pmod{p}) = 1 \text{ atunci } \left| \varphi_f \left(\frac{2\pi n}{p} \right) \right| = 1 \text{ pentru orice întreg } n.$$

2) Dacă $|\varphi_f(t_0)| = 1$ pentru $t_0 \neq 0$ atunci există un real a așa încît $P\left(f = a + \frac{2\pi}{t_0} k; k = 0, \pm 1, \dots\right) = 1$

Soluție. 1) Avem că

$$\varphi_f(t) = \sum_{k=-\infty}^{+\infty} e^{i(kp+m)t} P(f = kp + m) = e^{itm} \sum_{k=-\infty}^{+\infty} P(f = km + p) e^{ikp}$$

$$\varphi_f \left(\frac{2\pi n}{p} \right) = e^{\frac{i 2\pi nm}{p}} \sum_{k=-\infty}^{+\infty} P(f = km + p) e^{i 2\pi nk} = e^{\frac{i 2\pi nm}{p}} \sum_{k=-\infty}^{+\infty} P(f = km + p) =$$

$$= e^{\frac{i 2\pi nm}{p}}, \text{ deci } \left| \varphi_f \left(\frac{2\pi n}{p} \right) \right| = 1$$

2) Din $|\varphi_f(t_0)| = 1$ rezultă că există un număr real a așa încît $\varphi_f(t_0) = e^{it_0 a}$

Fie $g = f - a$; atunci $\varphi_g(t_0) = e^{-it_0 a} \varphi_f(t_0) = e^{-it_0 a} e^{it_0 a} = 1$, adică

$$\int \cos t_0 g dP = 1 \text{ sau } \int (1 - \cos t_0 g) dP = 0, \text{ de unde } 1 - \cos t_0 g = 0$$

a.s. și prin urmare $g = \frac{2\pi}{t_0} n; n = 0, \pm 1, \dots$, a.s. sau echivalen

$f = a + \frac{2\pi}{t_0} n, n = 0, \pm 1, \dots$ a. s. sau încă

$$P\left(f = a + \frac{2\pi}{t_0} n; n = 0, \pm 1, \dots\right) = 1.$$

2.66. Să se arate că o variabilă aleatoare f are repartiție simetrică (adică f și $-f$ au aceeași repartiție) dacă și numai dacă φ_f este reală.

Soluție. Din teorema de unicitate a funcțiilor caracteristice rezultă că f are repartiție simetrică dacă și numai dacă $\varphi_f(t) = \overline{\varphi_{-f}(t)}$. Pe de altă parte cum $\varphi_{-f}(t) = \varphi_f(-t) = \overline{\varphi_f(t)}$ rezultă că f are repartiție simetrică dacă și numai dacă $\varphi_f(t) = \overline{\varphi_f(t)}$, cu alte cuvinte dacă și numai dacă φ_f este reală.

2.67. Fie f o variabilă aleatoare cu repartiție simetrică așa încît $|f| \geq 1$.

Să se arate că φ_f se anulează în intervalul $(0, \pi)$.

Soluție. Este suficient să arătăm că $\int_0^\pi \varphi_f(t) \sin t \, dt \leq 0$. Deoarece f are repartiție simetrică rezultă din problema precedentă că φ_f este reală, deci $\varphi_f(t) = \int_{|f|>1} \cos tf \, dP$

Putem scrie

$$\begin{aligned} \int_0^\pi \varphi_f(t) \sin t \, dt &= \int_0^\pi \left(\int_{|f|>1} \cos tf \, dP \right) \sin t \, dt = \int_{|f|>1} \left(\int_0^\pi \cos tf \sin t \, dt \right) dP = \\ & \text{(teorema lui Fubini)} \\ &= \int_{|f|>1} \left(\int_0^\pi \frac{\sin [t(f+1)] + \sin [t(1-f)]}{2} \, dt \right) dP = \\ &= \frac{1}{2} \int_{|f|>1} \left\{ -\frac{1}{f+1} \cos [t(f+1)] - \frac{1}{1-f} \cos [t(1-f)] \right\} \Big|_0^\pi dP = \\ &= \frac{1}{2} \int_{|f|>1} \left\{ -\frac{1}{f+1} \cos [\pi(f+1)] - \frac{1}{1-f} \cos [\pi(1-f)] + \frac{1}{f+1} + \right. \\ & \left. + \frac{1}{f-1} \right\} dP = \frac{1}{2} \int_{|f|>1} \frac{(-1+f) \cos [\pi(f+1)] - (1+f) \cos [\pi(1-f)] + 2}{1-f^2} dP = \\ &= \frac{1}{2} \int_{|f|>1} \frac{-(-1+f) \cos \pi f + (1+f) \cos \pi f + 2}{1-f^2} dP = \int_{|f|>1} \frac{1 + \cos \pi f}{1-f^2} dP \leq 0 \end{aligned}$$

2.68. Fie f_1, \dots, f_n variabile aleatoare independente cu repartiția e_λ . Să se determine

1) Repartiția sumei $f_1 + \dots + f_n$

2) $M[(f_1 + \dots + f_n)^k]$

Soluție. 1) Funcția caracteristică a lui f_1 este $\frac{\lambda}{\lambda - it}$. Din independență obținem egalitatea

$$\varphi_{f_1 + \dots + f_n}(t) = \left(\frac{\lambda}{\lambda - it} \right)^n$$

Prin derivare în raport cu t de $n - 1$ ori în egalitatea

$$\frac{\lambda}{\lambda - it} = \int_0^{\infty} \lambda e^{itx} e^{-\lambda x} dx \text{ obținem}$$

$$i^{n-1}(n-1)! \frac{\lambda}{(\lambda - it)^n} = \lambda \int_0^{\infty} i^{n-1} x^{n-1} e^{itx} e^{-\lambda x} dx$$

sau echivalent

$$\left(\frac{\lambda}{\lambda - it} \right)^n = \frac{\lambda}{(n-1)!} \int_0^{\infty} e^{itx} (\lambda x)^{n-1} e^{-\lambda x} dx$$

din care rezultă că $P(f_1 + \dots + f_n \in A) = \int_A \frac{1}{(n-1)!} (\lambda x)^{n-1} \lambda e^{-\lambda x} dx$

2) Din egalitatea (16) (din rezumatul teoretic) rezultă că

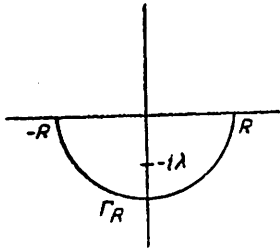
$$\begin{aligned} M[(f_1 + \dots + f_n)^k] &= \frac{1}{i^k} \varphi_{f_1 + \dots + f_n}^{(k)}(0) = \frac{1}{i^k} \left[\left(\frac{\lambda}{\lambda - it} \right)^n \right]^{(k)} \Big|_{t=0} = \\ &= \lambda^{-k} n(n+1) \dots (n+k-1) \end{aligned}$$

Se poate aplica (pentru o altă soluție) și teorema de inversiune. Deoarece

$$\int |\varphi_{f_1 + \dots + f_n}(t)| dt = \lambda^n \int_{-\infty}^{+\infty} \frac{1}{(\lambda^2 + t^2)^{\frac{n}{2}}} dt < \infty$$

rezultă din teorema de inversiune că $f_1 + \dots + f_n$ are densitatea

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \varphi_{f_1 + \dots + f_n}(t) dt = \frac{\lambda^n}{2\pi} \int_{-\infty}^{+\infty} \frac{e^{-itx}}{(\lambda - it)^n} dt$$



Fie pentru $x > 0$ funcția $g(z) = \frac{e^{-izx}}{(\lambda - iz)^n}$ și conturul alăturat.

Funcția g are în conturul alăturat polul $-i\lambda$, așa că din teorema rezidurilor rezultă că

$$(1) \int_R^{-R} g(z) dz + \int_{\Gamma_R} g(z) dz = 2\pi i \operatorname{Rez}(g, -i\lambda)$$

Avem

$$\begin{aligned} \operatorname{Rez}(g, -i\lambda) &= \frac{1}{(n-1)!} \lim_{z \rightarrow -\lambda i} [(z + \lambda i)^n g(z)]^{(n-1)} = \\ &= \frac{(-\lambda)^n}{i^n (n-1)!} \lim_{z \rightarrow -\lambda i} (e^{-izx})^{(n-1)} = -\frac{1}{i(n-1)!} x^{n-1} e^{-\lambda x} \end{aligned}$$

Pe de altă parte pe Γ_R

$$\left| \frac{1}{(\lambda - iz)^n} \right| \leq \frac{1}{|\lambda - R|^n} \rightarrow 0 \text{ cînd } R \rightarrow \infty$$

Din lema lui Jordan rezultă că $\lim_{R \rightarrow \infty} \int_{\Gamma_R} g(z) dz = 0$. Făcînd $R \rightarrow \infty$

în egalitatea (1) obținem că

$$\begin{aligned} \int_{-\infty}^{+\infty} g(z) dz &= -2\pi i \operatorname{Rez}(g, -i\lambda) = 2\pi \frac{1}{(n-1)!} x^{n-1} e^{-\lambda x}, \text{ deci} \\ \rho(x) &= \frac{\lambda^n}{2\pi} \cdot 2\pi \frac{1}{(n-1)!} x^{n-1} e^{-\lambda x} = \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x}, \quad x > 0 \end{aligned}$$

Dacă $x < 0$ se obține conturul simetric celui de mai sus. Funcția g este olomorfă în domeniul delimitat de contur, așa că integrala sa pe

contur este nulă (teorema lui Cauchy), deci $\int_{-\infty}^{+\infty} g(z) dz = 0$.

În concluzie am obținut

$$\rho(x) = \begin{cases} 0 & x \leq 0 \\ \frac{\lambda^n e^{-\lambda x}}{(n-1)!} & x > 0 \end{cases}$$

În continuare dăm o soluție directă pentru punctul 1). Calculăm funcția de repartiție a lui $f_1 + \dots + f_n$. Pentru $b \geq 0$ avem

$$\begin{aligned}
 P(f_1 + \dots + f_n < b) &= \lambda^n \int_{0 \leq x_1 + \dots + x_n < b} e^{-\lambda(x_1 + \dots + x_n)} dx_1 \dots dx_n = \\
 &= \lambda^n \int_0^b \int_0^{b-x_1} \dots \int_0^{b-x_1-\dots-x_{n-1}} e^{-\lambda(x_1 + \dots + x_n)} dx_1 \dots dx_{n-1} = I_n \\
 &\int_0^{b-x_1-\dots-x_{n-1}} e^{-\lambda(x_1 + \dots + x_n)} dx_n = -\frac{1}{\lambda} [e^{-\lambda b} - e^{-\lambda(x_1 + \dots + x_{n-1})}]
 \end{aligned}$$

Prin urmare are loc următoarea relație de recurență:

$$I_n = -\lambda^{n-1} \int_0^b \int_0^{b-x_1} \dots \int_0^{b-x_1-\dots-x_{n-2}} e^{-\lambda b} dx_1 \dots dx_{n-1} + I_{n-1}$$

Mai departe

$$\begin{aligned}
 &\int_0^b \int_0^{b-x_1} \dots \int_0^{b-x_1-\dots-x_{n-2}} dx_1 \dots dx_{n-1} = \\
 &= \int_0^b \dots \int_0^{b-x_1-\dots-x_{n-3}} (b - x_1 - \dots - x_{n-2}) dx_1 \dots dx_{n-2} = \\
 &= \int_0^b \dots \int_0^{b-x_1-\dots-x_{n-4}} \frac{(b - x_1 - \dots - x_{n-3})^2}{2!} dx_1 \dots dx_{n-3} = \\
 &= \int_0^b \dots \int_0^{b-x_1-\dots-x_{n-5}} \frac{(b - x_1 - \dots - x_{n-4})^3}{3!} dx_1 \dots dx_{n-4} = \\
 &= \dots = b^{n-1} / (n-1)!
 \end{aligned}$$

Rezultă egalitatea

$$I_n = I_{n-1} - [(\lambda b)^{n-1} / (n-1)!] e^{-\lambda b}$$

de unde deducem

$$P(f_1 + \dots + f_n < b) = I_n - 1 - c^{-\lambda b} [(\lambda b)^{n-1}/(n-1)! + \dots + \lambda b + 1]$$

Prin urmare obținem că $f_1 + \dots + f_n$ are densitatea

$$\rho(b) = I'_n = \lambda e^{-\lambda b} [(\lambda b)^{n-1}/(n-1)! + \dots + \lambda b + 1] - \lambda e^{-\lambda b} [(\lambda b)^{n-2}/(n-2)! + \dots + \lambda b + 1] = e^{-\lambda b} \lambda^n b^{n-1}/(n-1)!$$

2.69. Fie f, g variabile aleatoare independente, identic repartizate, de medie zero și de dispersie finită.

Să se demonstreze că f și g au repartiție normală dacă și numai dacă $f + g$ și $f - g$ sînt independente.

Soluție. Presupunem că f, g au repartiția $N(0, \sigma^2)$ și să arătăm că $f + g, f - g$ sînt independente.

Pentru aceasta este suficient să dovedim că

$$\varphi_{(f+g, f-g)}(u, v) = \varphi_{f+g}(u) \varphi_{f-g}(v)$$

Avem

$$\begin{aligned} \varphi_{(f+g, f-g)}(u, v) &= \int e^{i[u(f+g)+v(f-g)]} dP = \int e^{i[(u+v)f+(u-v)g]} dP = \\ &= \varphi_{(f, g)}(u+v, u-v) = \varphi_f(u+v) \varphi_g(u-v) = e^{-\frac{\sigma^2(u+v)^2}{2}} e^{-\frac{\sigma^2(u-v)^2}{2}} = \\ &= e^{-\sigma^2(u^2+v^2)} \end{aligned}$$

$$\varphi_{f+g}(u) = \varphi_f(u) \varphi_g(u) = e^{-\frac{\sigma^2 u^2}{2}} e^{-\frac{\sigma^2 u^2}{2}} = e^{-\sigma^2 u^2}$$

$$\varphi_{f-g}(v) = \varphi_f(v) \varphi_g(-v) = \varphi_f(v) \varphi_g(-v) = e^{-\frac{\sigma^2 v^2}{2}} e^{-\frac{\sigma^2 v^2}{2}} = e^{-\sigma^2 v^2}$$

deci

$$\varphi_{f+g}(u) \varphi_{f-g}(v) = e^{-\sigma^2(u^2+v^2)} = \varphi_{(f+g, f-g)}(u, v)$$

Presupunem acum că $f + g, f - g$ sînt independente și să arătăm că f și g au repartiție normală. Fără a restringe generalitatea putem presupune că $D^2(f) = D^2(g) = 1$. Fie φ funcția caracteristică a lui f și g . Va trebui să dovedim că $\varphi(t) = e^{-\frac{t^2}{2}}$. Deoarece $f + g, f - g$ sînt independente rezultă că

$$\varphi_{f+g+f-g}(t) = \varphi_{f+g}(t) \varphi_{f-g}(t) = \varphi_f(t) \varphi_g(t) \varphi_f(t) \varphi_g(-t) = \varphi^2(t) \varphi(-t),$$

dar

$$\varphi_{f+s+f-s}(t) = \varphi_{2f}(t) = \varphi_f(2t) = \varphi(2t).$$

Este valabilă deci egalitatea

$$\varphi(2t) = \varphi^3(t) \varphi(-t), \quad (1)$$

Afirmăm că $\varphi(t) \neq 0$ pentru orice t . Prin absurd dacă $\varphi(t_0) = 0$ atunci din (1) rezultă că $\varphi\left(\frac{t_0}{2}\right) = 0$, și în general $\varphi\left(\frac{t_0}{2^n}\right) = 0$, de unde

$$0 = \lim_{n \rightarrow \infty} \varphi\left(\frac{t_0}{2^n}\right) = \varphi(0) = 1, \text{ ceea ce nu se poate.}$$

Putem deci defini $\psi(t) = \ln \varphi(t)$. Din (1) rezultă că

$$\psi(2t) = 3\psi(t) + \psi(-t), \quad (2)$$

de unde înlocuind t cu $-t$ obținem și

$$\psi(-2t) = 3\psi(-t) + \psi(t) \quad (3)$$

Prin scădere din (2) și (3), cu notația $\delta(t) = \psi(t) - \psi(-t)$ obținem

$$\delta(2t) = 2\delta(t). \quad (2)$$

Prin ipoteză φ este de două ori derivabilă și $\varphi'(0) = 0$, $\varphi''(0) = -1$. Rezultă că ψ , δ sînt de două ori derivabile și că: $\psi(0) = \psi'(0) = 0$, $\psi''(0) = -1$, $\delta(0) = \delta'(0) = 0$.

Din (4) rezultă că $\frac{\delta(t)}{t} = \frac{\delta\left(\frac{t}{2^n}\right)}{\frac{t}{2^n}}$ de unde făcînd $n \rightarrow \infty$ obținem

că $\frac{\delta(t)}{t} = \delta'(0) = 0$, deci $\delta \equiv 0$ și prin urmare $\psi(t) = \psi(-t)$. În acest caz (2) devine $\psi(2t) = 4\psi(t)$ de unde din formula lui Taylor obținem

$$\frac{\psi(t)}{t^2} = \frac{\psi\left(\frac{t}{2^n}\right)}{\left(\frac{t}{2^n}\right)^2} = \frac{\psi(0) + \frac{t}{2^n} \psi'(0) + \frac{1}{2} \left(\frac{t}{2^n}\right)^2 \psi''(0) + o\left(\frac{t}{2^n}\right)^2}{\left(\frac{t}{2^n}\right)^2} = -\frac{1}{2} + o\left(\frac{t}{2^n}\right)$$

(unde $o(t)$ este o funcție așa încît $\lim_{t \rightarrow 0} o(t) = 0$)

Reținând primul și ultimul membru din egalitățile de mai sus și făcând $n \rightarrow \infty$ obținem că $\frac{\psi(t)}{t^2} = -\frac{1}{2}$, deci $\psi(t) = -\frac{t^2}{2}$ și prin urmare

$$\varphi(t) = c e^{-\frac{t^2}{2}}.$$

Observație. Problema precedentă admite următoarea generalizare. Dacă variabilele aleatoare f, g sînt independente, identic repartizate și $f + g, f - g$ sînt independente atunci f și g au repartiție normală (deci nu este necesar să le presupunem de dispersie finită).

Soluție. Fie φ funcția caracteristică a lui f și g . Trebuie să arătăm că există $m \in \mathbb{R}, \sigma \geq 0$ astfel încît $\varphi(t) = \exp(imt - \sigma^2 t^2/2)$ pentru orice t (prin definiție $N(m, 0) = \varepsilon_m$).

$$\text{Fie } \psi_1(t) = \varphi(t) \varphi(-t) = |\varphi(t)|^2.$$

Pasul 1. *Arătăm că $\psi_1(nx) = \psi_1(x)^{n^2}$ pentru orice $n \in \mathbb{N}, x \in \mathbb{R}$.* Din independență rezultă că pentru $t_1, t_2 \in \mathbb{R}$ are loc egalitatea

$$\begin{aligned} \varphi(t_1 + t_2) \varphi(t_1 - t_2) &= \varphi_{f+g, f-g}(t_1, t_2) = \varphi_{f+g}(t_1) \varphi_{f-g}(t_2) = \\ &= \varphi(t_1)^2 \varphi(t_2) \varphi(-t_2) \end{aligned} \quad (1)$$

Dacă în (1) luăm $t_1 = t_2 = x$ obținem

$$\varphi(2x) = \varphi^3(x) \varphi(-x) \quad (2)$$

de unde $\psi_1(2x) = \psi_1(x)^4$.

Să presupunem acum că $\psi_1((n-1)x) = \psi_1(x)^{(n-1)^2}$, $\psi_1(nx) = \psi_1(x)^{n^2}$. Punînd $t_1 = x, t_2 = nx$ în (1) obținem

$$\varphi((n+1)x) \varphi((n-1)x) = \varphi(nx)^2 \varphi(x) \varphi(-x)$$

de unde

$$\psi_1((n+1)x) \psi_1((n-1)x) = \psi_1((n+1)x) \psi_1(x)^{(n-1)^2} = \psi_1(nx)^2 \psi_1(x)^2$$

și deci

$$\psi_1((n+1)x) = \psi_1(x)^{2n^2 - (n-1)^2 + 2} = \psi_1(x)^{(n+1)^2}$$

Pasul 2. *Arătăm că φ, ψ_1 nu se anulează în nici un punct.* Prin absurd fie un real t așa încît $\psi_1(t) = 0$. Atunci din pasul (1) deducem că $\psi_1(t/n)^{n^2} = \psi_1(t)$ pentru orice natural n , deci $\psi_1(t/n) = 0$ și cum ψ_1 este continuă rezultă că $\psi_1(0) = 0$ ceea ce nu este posibil ($\psi_1(0) = 1$). Deci ψ_1 nu se anulează nicăieri și ca urmare și φ are aceeași proprietate.

Pasul 3. *Arătăm că există $b \in \mathbb{C}$ așa încît $\psi_1(x) = \exp(bx^2)$ dacă $x \in \mathbb{R}$.* Deoarece $\psi_1(1) \neq 0$ rezultă că există $b \in \mathbb{C}$ așa încît $\psi_1(t) =$

= exp (2b). Din pasul (1) avem $\psi_1(n) = \exp(2bn^2)$ pentru $n \in N$ și $\psi_1(n) = \psi_1(n/m)^m$ așa că $\psi_1(n/m) = \exp(2bn^2/m^2)$, adică $\psi_1(r) = \exp(2br^2)$ dacă $r \in Q_+$. Apoi continuitatea lui ψ_1 încheie probarea afirmației.

Pasul 4. Fie $\psi_2(t) = \varphi(t)/\varphi(-t)$. Similar ca în pasul (1) se arată că

$$\psi_2(nx) = \psi_2(x)^n \text{ pentru } n \in N; x \in R;$$

de unde raționînd ca în pasul (3) deducem egalitatea

$$\psi_2(x) = \exp(2ax) \text{ pentru } x \in R \text{ și un anumit } a \in C. \quad (3)$$

Pasul 5. Avem că $\varphi(t)^2 = \psi_1(t)\psi_2(t) = \exp(2at + 2bt^2)$, de unde $\varphi(t) = \pm \exp(at + bt^2)$. Din continuitatea lui φ și din faptul că nu se anulează rezultă că are peste tot același semn și anume + (deoarece $\varphi(0) = 1$).

Pasul 6. Arătăm că $b \leq 0$. Deoarece $|\varphi(t)|^2 = \exp(2bt^2)$ pentru orice $t \in R$ rezultă că $b \in R$. Dacă $b > 0$, atunci pentru t suficient de mare

$$|\varphi(t)| = |\exp(at + bt^2)| = \exp(t \operatorname{Re} a + bt^2) > 1 \text{ fapt imposibil}$$

$$(|\varphi(t)| \leq 1).$$

Rezultă că $b \leq 0$.

Pasul 7. Arătăm că $a/i \in R$. Din egalitatea $|\varphi(-t)| = |\varphi(t)|$ rezultă

$$|\exp(-at + bt^2)| = |\exp(at + bt^2)| \text{ pentru orice } t \in R, \text{ deci}$$

$$\exp(-t \operatorname{Re} a + bt^2) = \exp(t \operatorname{Re} a + bt^2) \text{ și prin urmare } \operatorname{Re} a = 0.$$

În fine punînd $m = a/i$, $b = -\sigma^2/2$ obținem că $\varphi(t) = \exp(imt - \sigma^2 t^2/2)$.

2.70. Fie f, g variabile aleatoare independente cu repartiția $N(0,1)$.

Să se arate că $\frac{f}{g}$ are repartiția $C_{1,0}$.

Soluție. Folosim următorul rezultat. Fie $u: R \rightarrow R$ așa încît

$$\int_{-\infty}^{+\infty} |u(x)| dx < \infty. \text{ Atunci este valabilă egalitatea}$$

$$\int_{-\infty}^{+\infty} u\left(x - \frac{t}{x}\right) dx = \int_{-\infty}^{+\infty} u(x) dx \text{ pentru orice } t > 0. \quad (1)$$

Într-adevăr aplicația $x \rightarrow x - \frac{t}{x}$ duce bijectiv $(-\infty, 0)$ pe R și $(0, \infty)$ pe R .

Pentru prima bijecție inversa este $x = \frac{y - \sqrt{y^2 + 4t}}{2}$ deci $dx = \frac{dy}{2} - \frac{ydy}{2\sqrt{y^2 + 4t}}$, iar pentru a doua bijecție inversa este $x = \frac{y + \sqrt{y^2 + 4t}}{2}$, deci $dx = \frac{dy}{2} + \frac{ydy}{2\sqrt{y^2 + 4t}}$.

Rezultă

$$\begin{aligned} \int_{-\infty}^{+\infty} u\left(x - \frac{t}{x}\right) dx &= \int_{-\infty}^0 u\left(x - \frac{t}{x}\right) dx + \int_0^{+\infty} u\left(x - \frac{t}{x}\right) dx = \\ &= \int_{-\infty}^{+\infty} u(y) \frac{dy}{2} - \int_{-\infty}^{+\infty} \frac{yu(y)}{\sqrt{y^2 + 4t}} \frac{dy}{2} + \int_{-\infty}^{+\infty} u(y) \frac{dy}{2} + \\ &\quad + \int_{-\infty}^{+\infty} \frac{yu(y)}{\sqrt{y^2 + 4t}} \frac{dy}{2} = \int_{-\infty}^{+\infty} u(y) dy \end{aligned}$$

(deoarece $|y| \leq \sqrt{y^2 + 4t}$ rezultă că integrala cu radical este finită). Mai departe avem:

$$\begin{aligned} \varphi_{\frac{f}{t}}(t) &= \int e^{it \frac{f}{t}} dP = \int_{R^2} \frac{1}{2\pi} e^{i \frac{y}{x}} e^{-\frac{x^2}{2}} e^{-\frac{y^2}{2}} dx dy = \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} e^{i \frac{y}{x} - \frac{y^2}{2}} dy \right) e^{-\frac{x^2}{2}} dx = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} e^{-\frac{(y - \frac{it}{x})^2}{2}} e^{-\frac{t^2}{x^2}} dy \right) e^{-\frac{x^2}{2}} dx = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{x^2} - \frac{x^2}{2}} dx. \end{aligned}$$

Fie acum $t > 0$ și $u(x) = e^{-\frac{x^2}{2} - t}$. Atunci $u\left(x - \frac{t}{x}\right) = e^{-\frac{t^2}{x^2} - \frac{x^2}{2}}$ și

prin aplicarea egalității (1) obținem

$$\varphi_{\frac{f}{g}}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2} - tx} dx = \frac{e^{-t^2}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = e^{-t^2} = e^{-|t|}$$

Dacă $t \leq 0$ atunci $\varphi_{\frac{f}{g}}(t) = \overline{\varphi_{\frac{f}{g}}}(-t) = e^{-|t|}$.

Prin urmare pentru orice t , $\varphi_{\frac{f}{g}}(t) = e^{-|t|}$ și cum $e^{-|t|}$ este funcția caracteristică a lui $C_{1,0}$ (vezi problema 2.55) rezultă că $\frac{f}{g}$ are repartiția $C_{1,0}$.

2.71. Fie vectorul aleator (f, g) a cărui repartiție este dată de coeficienții dezvoltării în serie a funcției $u(x, y) = \exp \{ \lambda (p_1 x + p_2 y + p_3 xy - 1) \}$ unde $\lambda > 0$, $p_1 + p_2 + p_3 = 1$, $p_i > 0$.

Să se arate că f și g verifică formula de multiplicare fără să fie independente.

Soluție. Cum $u(1,1) = 1$ și coeficienții dezvoltării lui u sînt pozitivi rezultă că aceștia definesc o repartiție pe $\{0, 1, \dots\} \times \{0, 1, \dots\}$. Mai departe din egalitatea

$$u(x, y) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} P(f = m, g = n) x^m y^n$$

se obțin relațiile

$$\begin{aligned} M(f) &= \sum_{m=0}^{\infty} m P(f = m) = \sum_{m=0}^{\infty} m \sum_{n=0}^{\infty} P(f = m, g = n) = \\ &= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} m P(f = m, g = n) = \frac{\partial u}{\partial x}(1,1) = \lambda(1 - p_2) \end{aligned}$$

$$M(g) = \lambda(1 - p_1) \quad (\text{se obține asemănător ca } M(f))$$

$$\begin{aligned} M(fg) &= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} mn P(f = m, g = n) = \\ &= \frac{\partial^2 u}{\partial x \partial y}(1, 1) = \lambda^2 (1 - p_1)(1 - p_2) \end{aligned}$$

Prin urmare $M(fg) = M(f)M(g)$.

Pe de altă parte putem scrie

$$P(f = 0) = \sum_{n=0}^{\infty} P(f = 0, g = n) = u(0, 1) = \exp \{ \lambda(p_2 - 1) \}$$

$$P(g = 0) = \sum_{m=0}^{\infty} P(f = m, g = 0) = u(1, 0) = \exp \{ \lambda(p_1 - 1) \}$$

$$P(f = 0, g = 0) = u(0, 0) = \exp(-\lambda)$$

deci $P(f = 0, g = 0) \neq P(f = 0) P(g = 0)$, fapt ce arată că f și g nu sînt independente.

2.72. Să se arate că dacă funcția caracteristică a sumei a două variabile aleatoare este egală cu produsul funcțiilor caracteristice ale variabilelor aleatoare atunci nu rezultă că variabilele aleatoare sînt independente.

Soluție. Fie f o variabilă aleatoare cu repartiția $C_{1,0}$. Atunci $\varphi_{f+f}(t) = \varphi_{2f}(t) = \varphi_f(2t) = e^{-2|t|} = e^{-|t|} \cdot e^{-|t|} = \varphi_f(t) \varphi_f(t)$ dar f nu este independentă de f întrucît $P(f < 0, f \geq 0) = 0$ și $P(f < 0) \neq 0$, $P(f \geq 0) \neq 0$.

2.73. Variabila aleatoare f are densitatea $p(x) = Cu(x)$ unde

$$u(x) = \begin{cases} 1 & 0 \leq x \leq 2 \\ \frac{1}{x^2 \ln x} & x > 2 \\ u(-x) & x < 0 \end{cases}$$

$$C = \left[\int_{\mathbb{R}} u(x) dx \right]^{-1}$$

Să se arate că $M(|f|) = \infty$ și $|\varphi'(0)| < \infty$.

Soluție. Deoarece pentru $x > 2$, $\frac{1}{x \ln x} \geq \frac{1}{x \ln 2}$ și $\int_2^{\infty} \frac{1}{x} dx = +\infty$

rezultă că $\int_2^{\infty} \frac{1}{x \ln x} dx = +\infty$.

Avem atunci

$$\begin{aligned} M(|f|) &= 2 \int_0^{\infty} x p(x) dx = 2 \int_0^2 x p(x) dx + 2 \int_2^{\infty} x p(x) dx = \\ &= 4 + 2C \int_2^{\infty} \frac{1}{x \ln x} dx = +\infty \end{aligned}$$

Mai departe avem

$$|1 - \varphi_f(t)| = 2 \int_0^{\infty} (1 - \cos tx) p(x) dx \leq$$

(dacă $T \geq 2$ și avînd în vedere inegalitatea $1 - \cos y \leq \frac{y^2}{2}$)

$$\begin{aligned} &\leq Ct^2 T \int_0^T p(x) dx + \int_T^{\infty} \frac{1 - \cos tx}{x^2 \ln x} dx = Ct^2 T \int_0^T p(x) dx + t \int_{tT}^{\infty} \frac{1 - \cos u}{u^2 \ln \frac{u}{t}} du = \\ &= Ct^2 T \int_0^T p(x) dx + \frac{t}{\ln T} \int_{tT}^{\infty} \frac{1 - \cos u}{u^2} du \end{aligned}$$

deci

$$\frac{|1 - \varphi_f(t)|}{t} \leq CtT \int_0^T p(x) dx + \frac{1}{\ln T} \int_{tT}^{\infty} \frac{1 - \cos u}{u^2} du$$

de unde $\overline{\lim}_{t \rightarrow 0} \frac{|1 - \varphi_f(t)|}{t} \leq \frac{1}{\ln T} \int_0^{\infty} \frac{1 - \cos u}{u^2} du$ și făcînd apoi $T \rightarrow \infty$ obținem că

$$\lim_{t \rightarrow 0} \frac{|1 - \varphi_f(t)|}{t} = 0$$

egalitatea care ne arată că $\varphi_f'(0) = 0$.

2.74. Fie f o variabilă aleatoare de medie finită. Să se probeze egalitatea

$$\lim_{t \rightarrow 0} \frac{2 - \varphi_f(t) - \varphi_f(-t)}{t^2} = M_2(f)$$

indiferent dacă membrul din dreapta este ori nu finit.

Să se deducă apoi echivalența următoarelor afirmații

- a) $M_2(f) < \infty$
- b) φ_f este de două ori derivabilă
- c) φ_f este de două ori derivabilă în zero.

Soluție. Avem

$$(1) \quad \frac{2 - \varphi_f(t) - \varphi_f(-t)}{t^2} = \int \frac{2 - e^{ift} - e^{-ift}}{t^2} dP = \int \frac{2(1 - \cos tf)}{t^2} dP.$$

Deoarece $1 - \cos u \leq \frac{u^2}{2}$ rezultă că $0 \leq \frac{(1 - \cos tf)}{t^2} \leq f^2$ pentru orice $t \neq 0$, x . Dacă $M_2(f) < \infty$ atunci prin aplicarea teoremei de convergență dominată putem trece la limită în (1):

$$\lim_{t \rightarrow 0} \frac{2 - \varphi_f(t) - \varphi_f(-t)}{t^2} \geq \int \lim_{t \rightarrow 0} \frac{2(1 - \cos tf)}{t^2} dP = \int f^2 dP = M_2(f)$$

Să presupunem acum că $M_2(f) = \infty$. Prin aplicarea lemei lui Fatou deducem că

$$\lim_{t \rightarrow 0} \frac{2 - \varphi_f(t) - \varphi_f(-t)}{t^2} \geq \int \lim_{t \rightarrow 0} \frac{2(1 - \cos tf)}{t^2} dP = M_2(f) = \infty,$$

deci $\lim_{t \rightarrow 0} \frac{2 - \varphi_f(t) - \varphi_f(-t)}{t^2} = \infty$

- a) \Rightarrow b) Este cunoscută
- b) \Rightarrow c) Este imediată
- c) \Rightarrow a) Prin aplicarea regulei lui l'Hospital obținem

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{2 - \varphi_f(t) - \varphi_f(-t)}{t^2} &= \lim_{t \rightarrow 0} \frac{\varphi_f'(-t) - \varphi_f'(t)}{2t} = \\ &= -\frac{1}{2} \lim_{t \rightarrow 0} \frac{\varphi_f'(-t) - \varphi_f'(0)}{-t} - \frac{1}{2} \lim_{t \rightarrow 0} \frac{\varphi_f'(t) - \varphi_f'(0)}{t} = -\varphi_f''(0). \end{aligned}$$

2.75. Să se deducă că dacă variabila aleatoare f are funcția caracteristică de forma $\varphi_f(t) = \exp\left\{ \int_{|u| \leq 1} (e^{itu} - itu - 1) dn(u) \right\}$, unde n este o măsură pozitivă pe R așa încît $\int_{|u| \leq 1} u^2 dn(u) < \infty$ atunci $M_2(f) < \infty$.

Soluție. Conform problemei precedente avem de arătat că $\varphi'(0)$ există. Se observă că aceasta are loc dacă $f(t) = \int_{|u| < 1} (e^{itu} - itu - 1) dn(u)$ are derivată de ordinul doi în zero.

Avem

$$\begin{aligned} \frac{f(t+h) - f(t)}{h} &= \int_{|u| < 1} \frac{e^{i(t+h)u} - 1 - i(t+h)u}{h} dn(u) = \\ &= \int_{|u| < 1} \frac{e^{itu}(e^{ihu} - 1) - ihu}{h} dn(u) + \int_{|u| < 1} iu(e^{itu} - 1) dn(u) \end{aligned}$$

(integrala a doua există deoarece $|iu(e^{itu} - 1)| \leq |t|u^2$)

Deoarece $\left| \frac{e^{ihu}(e^{ihu} - 1) - ihu}{h} \right| \leq \frac{h^2 u^2}{h} = hu^2$ și $\int_{|u| < 1} u^2 dn(u) < \infty$ putem aplica teorema de convergență dominată și obținem că

$$\begin{aligned} f'(t) &= \lim_{t \rightarrow 0} \frac{f(t+h) - f(t)}{h} = \int_{|u| < 1} \lim_{t \rightarrow 0} \frac{e^{itu}(e^{ihu} - 1) - ihu}{h} dn(u) + \\ &+ \int_{|u| < 1} iu(e^{itu} - 1) dn(u) = \int_{|u| < 1} iu(e^{itu} - 1) dn(u). \end{aligned}$$

Mai departe $\frac{f'(h) - f'(0)}{h} = \int_{|u| < 1} \frac{iu(e^{ihu} - 1)}{h} dn(u)$ și cum

$$\left| \frac{iu(e^{ihu} - 1)}{h} \right| \leq u^2, \quad \frac{iu(e^{ihu} - 1)}{h} \rightarrow -u^2 \text{ cînd } h \rightarrow 0$$

prin aplicarea teoremei de convergență dominată obținem

$$f''(0) = \lim_{h \rightarrow 0} \frac{f'(h) - f'(0)}{h} = - \int_{|u| < 1} u^2 dn(u) < \infty.$$

CAPITOLUL III

CONVERGENȚA VARIABILELOR ALEATOARE

Fie f, f_n ($n = 1, 2, \dots$) variabile aleatoare d -dimensionale definite pe un câmp de probabilitate (E, \mathcal{H}, P) .

Vom spune că șirul f_n converge către f

– *aproape sigur* (simbolic $f_n \xrightarrow{a.s.} f$) dacă $P(f_n \rightarrow f) = 1$, deci dacă mulțimea pe care f_n converge către f are probabilitatea 1.

– *în probabilitate* (simbolic $f_n \xrightarrow{P} f$) dacă $\lim_{n \rightarrow \infty} P(|f_n - f| > \varepsilon) = 0$ pentru orice $\varepsilon > 0$

– *în L^p* ($p \geq 1$) (simbolic $f_n \xrightarrow{L^p} f$) dacă $f, f_n \in L^p$ și $\lim_{n \rightarrow \infty} \|f_n - f\|_p = 0$ sau echivalent $\lim_{n \rightarrow \infty} \overline{M}_p(f_n - f) = 0$.

– *în repartiție* (simbolic $f_n \xrightarrow{r} f$) dacă $\lim_{n \rightarrow \infty} \int u(f_n) dP = \int u(f) dP$ pentru orice $u: R^d \rightarrow R$ continuă și mărginită.

Convergența în repartiție a variabilelor aleatoare provine din convergența slabă a repartițiilor care se definește astfel: șirul de repartiții F_n converge slab către repartiția F (simbolic $F_n \Rightarrow F$) dacă $\lim_{n \rightarrow \infty} \int u dF_n = \int u dF$ pentru orice $u: R^d \rightarrow R$ continuă și mărginită. Prin urmare $f_n \xrightarrow{r} f$ dacă $P \circ f_n^{-1} \Rightarrow P \circ f^{-1}$.

Convenție. Dacă F este o repartiție, vom spune că f_n converge slab către F (simbolic $f_n \Rightarrow F$) dacă $P \circ f_n^{-1} \Rightarrow F$.

Au loc implicațiile următoare:

Convergența aproape singură

Convergența în L^p

↓

↓

Convergența în probabilitate \Rightarrow Convergența în repartiție

Șirul f_n converge către f aproape sigur, în probabilitate, în L^p , în repartiție dacă și numai dacă f_n este șir Cauchy în același mod (pentru primele trei convergențe definiția șirului Cauchy este imediată iar pentru convergența în repartiție aceasta înseamnă

că $\left\{ \int u(f_n) dP \right\}_n$ este șir Cauchy pentru orice $u: R^d \rightarrow R$ continuă și mărginită).

Au loc următoarele descrieri ale convergențelor introduse mai sus.

a) $f_n \xrightarrow{a.s} f$ dacă și numai dacă $\sup_{m \geq n} |f_m - f| \xrightarrow{p} 0$ cu alte cuvinte dacă și numai dacă $\lim_{n \rightarrow \infty} P(\sup_{m \geq n} |f_m - f| > \varepsilon) = 0$ pentru orice $\varepsilon > 0$.

b) $f_n \xrightarrow{p} f$ dacă și numai dacă din orice subșir al lui (f_n) se poate extrage alt subșir aproape sigur convergent către f .

c) $f_n \xrightarrow{L^1} f$ dacă și numai dacă $\lim_{n \rightarrow \infty} \sup_A \left| \int_A f_n dP - \int_A f dP \right| = 0$

d) Următoarele afirmații sînt echivalente:

d₁) $f_n \xrightarrow{r} f$

d₂) $\lim_{n \rightarrow \infty} \int u(f_n) dP = \int u(f) dP$ pentru orice $u: R^d \rightarrow R$ uniform continuă și mărginită

d₃) $\lim_{n \rightarrow \infty} F_{f_n}(x) = F_f(x)$ pentru orice $x = (x_1, \dots, x_d) \in R^d$ așa încît x_i este punct de continuitate pentru F_{f_i} .

d₄) $\lim_{n \rightarrow \infty} P(f_n \in A) = P(f \in A)$ dacă $P(f \in \partial A) = 0$

e) Din inegalitatea lui Cebîșev rezultă că $f_n \xrightarrow{p} 0$ dacă

$$\lim_{n \rightarrow \infty} M_2(f_n) = 0, \quad (1)$$

și în particular $f_n - M(f_n) \xrightarrow{p} 0$ dacă

$$\lim_{n \rightarrow \infty} D^2(f_n) = 0. \quad (2)$$

Conceptul de uniform integrabilitatea ce-l vom introduce în continuare permite stabilirea unor implicații inverse celor listate mai sus.

O familie \mathcal{U} de variabile aleatoare de medie finită este uniform integrabilă dacă

$$\lim_{a \rightarrow \infty} \sup_{f \in \mathcal{U}} \int_{|f| > a} |f| dP = 0.$$

Dacă familia \mathcal{U} are proprietatea că are elementele majorate în norma din R^d de aceeași variabilă aleatoare g de medie finită sau mai general dacă există $p > 1$ așa încît $\sup_{f \in \mathcal{U}} \overline{M}_p(f) < \infty$ atunci \mathcal{U} este uniform integrabilă.

Următoarea afirmație este o generalizare a teoremei de convergență dominată.

f) Presupunem că $f_n \xrightarrow{p} f$ și $f_n \in L^p$. Atunci $f_n \xrightarrow{L^p} f$ dacă și numai dacă $(f_n)_n$ este uniform integrabilă.

Următorul rezultat cunoscut sub numele de *teorema de convergență a funcțiilor caracteristice* este adesea util.

g) Dacă $F_n \Rightarrow F$ atunci $\varphi_{F_n}(t) \rightarrow \varphi_F(t)$ uniform pe orice compact și reciproc dacă $\varphi_{F_n}(t)$ converge punctual către o funcție φ continuă în origine, atunci φ este funcția caracteristică a unei repartiții și F_n converge slab către acea repartiție.

Această afirmație se poate folosi și pentru a verifica că o funcție este funcție caracteristică.

Șirul (f_n) de variabile aleatoare de medie finită satisface *legea slabă a numerelor mari* dacă $S_n = \frac{1}{n} \sum_{i=1}^n [f_i - M(f_i)] \xrightarrow{P} 0$. Se observă că dacă $M(f_n) = a$ pentru

orice n atunci șirul (f_n) satisface legea slabă dacă $S'_n = \frac{1}{n} \sum_{i=1}^n f_i \xrightarrow{P} a$.

Criteriu. Șirul (f_n) satisface legea slabă a numerelor mari dacă și numai dacă

$$\lim_{n \rightarrow \infty} M\left(\frac{S_n^2}{1 + S_n^2}\right) = 0 \quad (3)$$

O condiție suficientă ca (f_n) să satisfacă legea slabă a numerelor mari este următoarea

$$\lim_{n \rightarrow \infty} D^2\left(\frac{1}{n} \sum_{i=1}^n f_i\right) = 0 \quad (4)$$

În cazul cînd variabilele aleatoare f_1, \dots, f_n, \dots sînt independente două cite două condiția (4) devine

$$\lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i=1}^n D^2(f_n) = 0 \quad (5)$$

Șirul (f_n) satisface *legea tare a numerelor mari* dacă $S_n \xrightarrow{a.s.} 0$ (dacă $M(f_n) = a$ pentru orice n atunci (f_n) satisface legea tare a numerelor mari dacă $S'_n \xrightarrow{a.s.} a$).

Dacă variabilele aleatoare f_1, \dots, f_n, \dots sînt independente, identic repartizate și $M(f_1) = a < \infty$ atunci șirul (f_n) satisface legea tare a numerelor mari.

Șirul de variabile aleatoare (f_n) se spune că satisface

— *teorema limită centrală* dacă $0 < D^2(f_n) < \infty$ pentru orice n și

$$\frac{\sum_{k=1}^n [f_k - M(f_k)]}{D(f_1 + \dots + f_n)} \Rightarrow N(0, 1) \quad (6)$$

— *condiția Lindeberg* dacă $0 < D^2(f_n) < \infty$ pentru orice n și

$$\lim_{n \rightarrow \infty} \frac{1}{D^2(f_1 + \dots + f_n)} \sum_{k=1}^n \int \frac{|f_k - M(f_k)|}{D(f_1 + \dots + f_n)} [f_k - M(f_k)]^2 dP = 0 \quad (7)$$

pentru orice $\varepsilon > 0$.

Următoarele afirmații sînt echivalente (*teorema limită centrală*):

i) Șirul (f_n) satisface teorema limită centrală și

$$\lim_{n \rightarrow \infty} \max_{1 \leq k \leq n} \frac{D^2(f_k)}{D^2(f_1 + \dots + f_n)} = 0. \quad (8)$$

ii) Șirul (f_n) satisface condiția Lindeberg.

Din teorema limită centrală se obțin următoarele două rezultate:

iii) Dacă variabilele aleatoare f_1, \dots, f_n, \dots sînt independente, identic repartizate și $0 < \sigma = D(f_1) < \infty$ atunci $\frac{f_1 + \dots + f_n - nM(f_1)}{\sigma\sqrt{n}} \Rightarrow N(0, 1)$.

iv) (Teorema Leapunov). Dacă variabilele aleatoare f_1, \dots, f_n, \dots sînt independente, $0 < D^2(f_n) < \infty$ pentru orice n și

$$\lim_{n \rightarrow \infty} \frac{1}{[D(f_1 + \dots + f_n)]^{2+\delta}} \sum_{k=1}^n M(|f_k - M(f_k)|^{2+\delta}) = 0 \quad (9)$$

pentru un anumit $\delta > 0$, atunci șirul (f_n) satisface teorema limită centrală.

Teorema limită centrală multidimensională pentru variabile aleatoare identic repartizate. Fie f_1, \dots, f_n, \dots un șir de vectori aleatori d -dimensionali, independenți, identic repartizați și cu $M(|f_1|)^2 < \infty$, C_{f_1} pozitiv definită. Atunci

$$\frac{1}{\sqrt{n}} \sum_{k=1}^n [f_k - M(f_k)] \Rightarrow N_d(0, C_{f_1}).$$

§ 1. Tipuri de convergență a variabilelor aleatoare

3.1. Să se demonstreze următoarele:

- convergența aproape sigură nu implică convergența în L^p ;
- convergența în L^p nu implică convergența aproape sigură;
- convergența în probabilitate nu implică convergența aproape sigură;
- convergența în probabilitate nu implică convergența în L^p ;
- convergența în repartiție nu implică convergența în probabilitate.

Soluție. a) Pe cîmpul de probabilitate $([0, 1], \mathfrak{B}_{[0, 1]}, m)$, m măsura Lebesgue, definim pentru orice n variabila aleatoare f_n prin

$$f_n(\omega) = \begin{cases} n^p & \text{dacă } \omega \in \left[0, \frac{1}{n^p}\right] \\ 0 & \text{în rest} \end{cases}$$

Pentru orice $\omega \in (0, 1]$ există n_ω așa încît $\omega \notin \left[0, \frac{1}{n^p}\right]$ pentru $n \geq n_\omega$. Prin urmare pentru $n \geq n_\omega$ avem că $f_n(\omega) = 0$, deci $f_n \rightarrow 0$ peste tot în $(0, 1]$ deci în particular a.s. în $[0, 1]$.

Dar $f_n \xrightarrow{L^p} 0$ deoarece $\int |f_n|^p dm = \int_{\left[0, \frac{1}{n^p}\right]} n^p dm = 1 \rightarrow 0$.

b). Orice $n \in \mathbb{N}$ se scrie unic sub forma $n = 2^k + m$, $0 \leq m \leq 2^k$, $k = 0, 1, \dots$. Pe câmpul de probabilitate precedent definim variabila aleatoare

$$f_n(\omega) = \begin{cases} 1 & \text{dacă } \omega \in \left[\frac{m}{2^k}, \frac{m+1}{2^k}\right] \\ 0 & \text{în rest} \end{cases}$$

Avem că $\int |f_n|^p dm = \int_{\left[\frac{m}{2^k}, \frac{m+1}{2^k}\right]} dm = \frac{1}{2^k} \rightarrow 0$ când $n \rightarrow \infty$, deci $f_n \xrightarrow{L^p} 0$.

Însă pentru orice $\omega \in [0, 1]$ avem $f_n(\omega) = 1$ pentru o infinitate de valori ale lui n , deci în particular $f_n(\omega) \not\rightarrow 0$, așa că $f_n \not\xrightarrow{\text{a.s.}} 0$.

c) Pentru orice n construim variabilele aleatoare $f_1^{(n)}, \dots, f_n^{(n)}$ prin

$$f_i^{(n)}(\omega) = \begin{cases} 1 & \text{dacă } \omega \in \left[\frac{i-1}{n}, \frac{i}{n}\right] \\ 0 & \text{în rest} \end{cases}$$

Considerăm șirul $f_1^{(1)}, f_1^{(2)}, f_2^{(2)}, \dots, f_i^{(n)}, \dots, f_n^{(n)}, \dots$.

Avem că $P(|f_i^{(n)}| \geq \varepsilon) \leq \frac{1}{n} \rightarrow 0$ deci $f_i^{(n)} \xrightarrow{p} 0$.

Dar pentru orice $\omega \in [0, 1]$ avem o infinitate de 1 în șirul $f_1^{(1)}(\omega), \dots, f_1^{(n)}(\omega), \dots$, fapt ce arată că șirul nu converge a.s. către 0.

d) Fie $n = 2^k + m$ scrierea unică a lui n ca la punctul b). Definim

$$f_n(\omega) = \begin{cases} 2^k & \omega \in \left[\frac{m}{2^{kp}}, \frac{m+1}{2^{kp}}\right] \\ 0 & \text{în rest} \end{cases}$$

Pentru $\varepsilon > 0$ avem $P(|f_n| \geq \varepsilon) \leq \frac{1}{2^{kp}} \rightarrow 0$ când $n \rightarrow \infty$, deci $f_n \xrightarrow{p} 0$ și

$$\int |f_n|^p dm = \int_{\left[\frac{m}{2^{kp}}, \frac{m+1}{2^{kp}}\right]} (2^k)^p dm = 1 \rightarrow 0 \text{ deci } f_n \not\xrightarrow{L^p} 0.$$

e) Pe câmpul de probabilitate din punctul a) definim variabilele aleatoare

$$f_n(\omega) = \begin{cases} 1 & \omega \in \left[0, \frac{1}{2}\right) \\ 0 & \omega \in \left[\frac{1}{2}, 1\right] \end{cases}$$

$$f(\omega) = \begin{cases} 0 & \omega \in \left[0, \frac{1}{2}\right) \\ 1 & \omega \in \left[\frac{1}{2}, 1\right] \end{cases}$$

Se observă că f și f_n au aceeași repartiție așa că $f_n \xrightarrow{r} f$. În schimb pentru $1 > \varepsilon > 0$ avem $P(|f_n - f| > \varepsilon) = m([0, 1]) = 1 \not\rightarrow 0$, cu alte cuvinte $f_n \not\xrightarrow{p} f$.

3.2. Fiind date șirurile de variabile aleatoare d -dimensionale (f_n) , (g_n) și variabilele aleatoare f, g cu proprietățile

a) $f_n \xrightarrow{a.s} f$ (respectiv $f_n \xrightarrow{p} f, f_n \xrightarrow{L^p} f$).

b) $g_n \xrightarrow{a.s} g$ (respectiv $g_n \xrightarrow{p} g, g_n \xrightarrow{L^p} g$).

c) Pentru orice $n, f_n = g_n$ a.s.

să se arate că $f = g$ a.s.

Soluție. Presupunem întâi că $f \xrightarrow{a.s} f, g_n \xrightarrow{a.s} g$ și să notăm $A_n = \{f_n = g_n\}$, $B = \{f_n \rightarrow f\}$, $C = \{g_n \rightarrow g\}$.

Prin ipoteză avem că $P(A_n) = P(B) = P(C) = 1$, deci avem și că $P(\left(\bigcap_n A_n\right) \cap B \cap C) = 1$ (intersecție numărabilă de mulțimi de probabilitate 1 are probabilitatea 1).

Cum este adevărată incluziunea $(\bigcap_n A_n) \cap B \cap C \subset \{f = g\}$ rezultă că $P(f = g) = 1$.

Presupunem acum că $f_n \xrightarrow{p} f, g_n \xrightarrow{p} g$. Alegem un subșir (n_i) așa încît $f_{n_i} \xrightarrow{a.s} f, g_{n_i} \xrightarrow{a.s} g$. Acuma totul se reduce la cazul precedent,

Altfel: Deoarece $P(f - g = 0) = \lim_{n \rightarrow \infty} P\left(|f - g| > \frac{1}{n}\right)$ rezultă că totul se reduce la a arăta că $P(|f - g| > \varepsilon) = 0$ pentru orice $\varepsilon > 0$,

Avem că

$$P(|f - g| > \varepsilon) \leq P\left(|f_n - f| > \frac{\varepsilon}{3}\right) + P\left(|f_n - g_n| > \frac{\varepsilon}{3}\right) + P\left(|g_n - g| > \frac{\varepsilon}{3}\right) = P\left(|f_n - g| > \frac{\varepsilon}{3}\right) + P\left(|g_n - g| > \frac{\varepsilon}{3}\right)$$

de unde făcînd $n \rightarrow \infty$ obținem că $P(|f - g| > \varepsilon) = 0$.

Presupunem că $f_n \xrightarrow{L^p} f$, $g_n \xrightarrow{L^p} g$. Deoarece în particular $f_n \xrightarrow{p} f$, $g_n \xrightarrow{p} g$ afirmația rezultă din cea corespunzătoare convergenței în probabilitate. Altfel: Din inegalitatea lui Minkowski rezultă că

$$\begin{aligned} \|f - g\|_p &\leq \|f - f_n\|_p + \|g_n - f_n\|_p + \|g_n - g\|_p = \\ &= \|f - f_n\|_p + \|g - g_n\|_p \end{aligned}$$

Făcînd $n \rightarrow \infty$ obținem că $\|f - g\|_p = 0$, deci $f = g$ a.s.

3.3. Să se arate că dacă $f_n \xrightarrow{r} f$, $g_n \xrightarrow{r} g$ și pentru orice n , f_n are aceeași repartiție cu g_n atunci f și g au aceeași repartiție.

Soluție. Fie $u: R^d \rightarrow R$ continuă și mărginită. Prin ipoteză avem că

$$\begin{aligned} \lim_{n \rightarrow \infty} \int u(f_n) dP &= \int u(f) dP; \quad \lim_{n \rightarrow \infty} \int u(g_n) dP = \int u(g) dP; \\ P \circ f_n^{-1} &= P \circ g_n^{-1} \end{aligned}$$

Din formula de transport rezultă că

$$\int u(f_n) dP = \int u dP \circ f_n^{-1}; \quad \int u(g_n) dP = \int u dP \circ g_n^{-1}$$

deci avem egalitatea $\int u(f_n) dP = \int u(g_n) dP$, de unde prin trecere la limită $\int u(f) dP = \int u(g) dP$ sau echivalent $\int u dP \circ f^{-1} = \int u dP \circ g^{-1}$, așa că $P \circ f^{-1} = P \circ g^{-1}$ (vezi problema 1.34 din Cap. I).

3.4. Să se arate că $f^n = (f_1^n, \dots, f_d^n)$ converge către $f = (f_1, \dots, f_d)$ a.s., în probabilitate sau în L^p dacă și numai dacă pentru orice $1 \leq i \leq d$, f_i^n converge către f_i în același mod.

Soluție. Pentru convergența a.s. afirmația este o consecință a faptului că mulțimea de convergență a vectorului este intersecția mulțimilor de convergență ale componentelor.

Pentru convergența în probabilitate afirmația se reduce la convergența a.s. prin aplicarea rezultatului b) din rezumatul teoretic sau se poate obține direct din inegalitățile :

$$P(|f^n - f| > \varepsilon) \leq \sum_{i=1}^d P\left(|f_i^n - f_i| > \frac{\varepsilon}{\sqrt{d}}\right);$$

$$P(|f_i^n - f_i| > \varepsilon) \leq P(|f^n - f| > \varepsilon).$$

Pentru convergența în L^p afirmația este o consecință a inegalităților :

$$\|f^n - f\|_p \leq \sum_{i=1}^d \|f_i^n - f_i\|_p \quad (\text{inegalitatea lui Minkowski})$$

$$\|f_i^n - f_i\|_p \leq \|f^n - f\|_p.$$

3.5. Să se arate că dacă $f^n = (f_1^n, \dots, f_d^n)$ converge în repartiție către $f = (f_1, \dots, f_d)$ atunci pentru orice $1 \leq i \leq d$, f_i^n converge în repartiție către f_i .

Dacă în plus f_1^n, \dots, f_d^n sînt independente pentru orice n atunci f_1, \dots, f_d sînt independente.

Soluție. Fie $u: R \rightarrow R$ o funcție continuă și mărginită și fie $\bar{u}: R^d \rightarrow R$ funcție continuă și mărginită definită prin $\bar{u}(x) = u(x_i)$.

Atunci prin ipoteză $\int \bar{u}(f_1^n, \dots, f_d^n) dP \rightarrow \int \bar{u}(f_1, \dots, f_d) dP$ sau echivalent

(avînd în vedere definiția lui \bar{u}) $\int u(f_i^n) dP \rightarrow \int u(f_i) dP$, cu

alte cuvinte $f_i^n \xrightarrow{r} f_i$.

Să presupunem acumă că f_1^n, \dots, f_d^n sînt independente pentru orice n și să artăm că f_1, \dots, f_d sînt independente. Fie x_i punct de continuitate pentru funcția de repartiție a lui f_i . Avem că $P(f_1^n < x_1, \dots, f_d^n < x_d) \rightarrow P(f_1 < x_1, \dots, f_d < x_d)$ și din independență $P(f_1^n < x_1, \dots, f_d^n < x_d) = P(f_1^n < x_1) \dots P(f_d^n < x_d)$. Pe de altă parte din $f_i^n \xrightarrow{r} f_i$ avem și $P(f_i^n < x_i) \rightarrow P(f_i < x_i)$.

Avem deci relațiile :

$$\begin{aligned} \lim_{n \rightarrow \infty} P(f_1^n < x_1, \dots, f_d^n < x_d) &= P(f_1 < x_1, \dots, f_d < x_d) = \\ &= P(f_1 < x_1) \dots P(f_d < x_d) \end{aligned}$$

Reținem deci egalitatea

$$P(f_1 < x_1, \dots, f_d < x_d) = P(f_1 < x_1) \dots P(f_d < x_d)$$

valabilă pentru x_i ; punct de continuitate pentru funcția de repartiție a lui f_i .

Or cum $\mathfrak{M}_i = \{(-\infty, x)\}$; x_i punct de continuitate pentru funcția de repartiție a lui f_i formează un sistem de generatori închis la intersecția finită rezultă că f_1, \dots, f_d sînt independente.

Putem proceda și astfel. Este suficient să arătăm că

$$\varphi_f(t_1, \dots, t_d) = \varphi_{f_1}(t_1) \dots \varphi_{f_d}(t_d)$$

Or din teorema de convergență a funcțiilor caracteristice avem:

$$\varphi_f(t_1, \dots, t_d) = \lim_{n \rightarrow \infty} \varphi_{f_n}(t_1, \dots, t_d)$$

$$\varphi_{f_i}(t_i) = \lim_{n \rightarrow \infty} \varphi_{f_i^n}(t_i)$$

așa că este suficient să arătăm că

$$\varphi_{f_n}(t_1, \dots, t_d) = \varphi_{f_1^n}(t_1) \dots \varphi_{f_d^n}(t_d).$$

Or aceasta este o consecință a independenței variabilelor aleatoare f_1^n, \dots, f_d^n .

3.6. Fie (f_n) un șir de variabile aleatoare d -dimensionale și f o variabilă aleatoare d -dimensională.

Să se arate că dacă $u: R^d \rightarrow R$ este continuă și $f_n \xrightarrow{a.s.} f$ (resp. $f_n \xrightarrow{p} f$, $f_n \xrightarrow{r} f$) atunci $u(f_n) \xrightarrow{a.s.} u(f)$ (resp. $u(f_n) \xrightarrow{p} u(f)$, $u(f_n) \xrightarrow{r} u(f)$).

Dacă $f_n \xrightarrow{L^p} f$ și familia $\{|u(f_n)|^p\}_n$ este uniform integrabilă să se arate că $u(f_n) \xrightarrow{L^p} u(f)$ (de exemplu este cazul cînd u este și mărginită sau are una din formele $u(x) = \langle \alpha, x \rangle$, $u(x) = \max(x_1, \dots, x_d)$, $u(x) = \min(x_1, \dots, x_d)$).

Soluție. Pentru convergența a.s. afirmația rezultă imediat din incluziunea $\{f_n \rightarrow f\} \subset \{u(f_n) \rightarrow u(f)\}$

Presupunem acum că $f_n \xrightarrow{p} f$ și fie (n') un subsir. Alegem (n'') un subsir al lui (n') așa încît $f_{n''} \xrightarrow{a.s.} f$. Atunci $u(f_{n''}) \xrightarrow{a.s.} u(f)$ și prin urmare $u(f_n) \xrightarrow{p} u(f)$ (vezi afirmația b) din rezumatul teoretic). Putem proceda și astfel. Fie $\varepsilon > 0$, $T < \infty$ și alegem $\delta(\varepsilon, T) > 0$ așa încît $|x| \leq T$, $|y| \leq T$, $|x - y| \leq \delta$ implică $|u(x) - u(y)| \leq \varepsilon$ (u este uniform continuă pe compacți).

Avem că :

$$P(|u(f_n) - u(f)| > \varepsilon) \leq P(|f_n| \leq T, |f| \leq T, |u(f_n) - u(f)| > \varepsilon) \\ + P(|f_n| > T) + P(|f| > T) \leq P(|f_n - f| > \delta) + P\left(|f_n - f| > \frac{T}{2}\right) + \\ + P\left(|f| > \frac{T}{2}\right) + P(|f| > T).$$

de unde făcînd $n \rightarrow \infty$ obținem $\overline{\lim}_{n \rightarrow \infty} P(|u(f_n) - u(f)| > \varepsilon) \leq P\left(|f| > \frac{T}{2}\right) + P(|f| > T)$ și în fine făcînd $T \rightarrow \infty$ deducem că $\overline{\lim}_{n \rightarrow \infty} P(|u(f_n) - u(f)| > \varepsilon) = 0$ (întrucît $\lim_{n \rightarrow \infty} P(|f| > a) = P(|f| = \infty) = 0$;

$|f|$ ia valori în R_+). Presupunem că $f_n \xrightarrow{r} f$ și fie $\varphi : R \rightarrow R$ o funcție continuă și mărginită. Deoarece $\varphi \circ u$ este continuă și mărginită rezultă că $\int (\varphi \circ u)(f_n) dP \rightarrow \int (\varphi \circ u)(f) dP$, fapt ce arată că $u(f_n) \xrightarrow{r} u(f)$,

În fine fie $f_n \xrightarrow{L^p} f$. Atunci $f_n \xrightarrow{L^p} f$ și deci $u(f_n) \xrightarrow{p} u(f)$. Afirmația f) din rezumatul teoretic arată că $u(f_n) \xrightarrow{L^p} u(f)$.

3.7. Fie $f, f_1, \dots, f_n \dots$ variabile aleatoare și $\varepsilon_1, \dots, \varepsilon_n, \dots$ un șir de numere reale.

Să se arate că

a) dacă $\varepsilon_n \searrow 0$ și $\sum_{n=1}^{\infty} P(|f_n - f| \geq \varepsilon_n) < \infty$ atunci $f_n \xrightarrow{a.s.} f$;

b) $\sum_{n=1}^{\infty} \varepsilon_n < \infty$ și $\sum_{n=1}^{\infty} P(|f_{n+1} - f_n| \geq \varepsilon_n) < \infty$ atunci f_n converge

aproape sigur către o variabilă aleatoare f^* .

Soluție. a) Dacă punem $A_n = \{|f_n - f| \geq \varepsilon_n\}$ atunci din lema Borel-Cantelli obținem că $P(\overline{\lim}_{n \rightarrow \infty} A_n) = 0$, sau echivalent

$$P(\lim_{n \rightarrow \infty} \complement A_n) = P(\complement \overline{\lim}_{n \rightarrow \infty} A_n) = 1.$$

Dacă $\omega \in A = \lim_{n \rightarrow \infty} \complement A_n$ atunci $|f_n(\omega) - f(\omega)| \leq \varepsilon_n$ pentru n suficient de mare, deci în particular $f_n(\omega) \rightarrow f(\omega)$.

Prin urmare $A \subset \{f_n \rightarrow f\}$ și cum $P(A) = 1$ rezultă că $P(f_n \rightarrow f) = 1$, cu alte cuvinte $f_n \xrightarrow{a.s.} f$.

b) Dacă punem $A_n = \{|f_{n+1} - f_n| \geq \varepsilon_n\}$ și $A = \lim_{n \rightarrow \infty} A_n$ atunci ca la punctul precedent rezultă că $P(A) = 1$ și pentru $\omega \in A$ avem $|f_{n+1}(\omega) - f_n(\omega)| \leq \varepsilon_n$ pentru n suficient de mare, deci în particular seria $\sum_n [f_{n+1}(\omega) - f_n(\omega)]$ este convergentă.

Atunci obținem că și șirul $f_n(\omega) = f_1(\omega) + \sum_{l=1}^{n-1} [f_{l+1}(\omega) - f_l(\omega)]$ este convergent.

Dacă punem

$$f^*(\omega) = \begin{cases} \lim_{n \rightarrow \infty} f_n(\omega) & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

atunci este ușor de văzut că $f_n \xrightarrow{a.s.} f^*$

3.8. Fie f_1, \dots, f_n, \dots un șir de variabile aleatoare așa încît $f_n \xrightarrow{p} 0$ și pentru orice $n f_n \geq 0$, $f_n \geq f_{n+1}$ a.s.

Să se arate că $f_n \xrightarrow{a.s.} 0$.

Soluție. Avem de arătat că $P(\sup_{m \geq n} |f_m| \geq \varepsilon) \rightarrow 0$ pentru orice $\varepsilon > 0$.

Fie $A_n = \{f_n \geq f_{n+1}\}$, $B_n = \{f_n \geq 0\}$. Prin ipoteză $P(A_n) = 1$, $P(B_n) = 1$ deci $P(\bigcap_n (A_n \cap B_n)) = 1$.

Rezultă că

$$\begin{aligned} P(\sup_{m \geq n} |f_m| \geq \varepsilon) &= P(\{\sup_{m \geq n} |f_m| \geq \varepsilon\} \cap (\bigcap_m (A_m \cap B_m))) = \\ &= P(\{|f_n| \geq \varepsilon\} \cap \{\bigcap_m (A_m \cap B_m)\}) = P(|f_n| \geq \varepsilon) \rightarrow 0 \end{aligned}$$

3.9. Fie (f_n) , (g_n) , (h_n) trei șiruri de variabile aleatoare cu proprietățile

a) $f_n \xrightarrow{a.s.} f$ (resp. $f_n \xrightarrow{p} f$, $f_n \xrightarrow{L^p} f$, $f_n \xrightarrow{r} f$);

b) $h_n \xrightarrow{a.s.} f$ (resp. $h_n \xrightarrow{p} f$, $h_n \xrightarrow{L^p} f$, $h_n \xrightarrow{r} f$);

c) Pentru orice $n, f_n \leq g_n \leq h_n$ a.s.

Să se arate că $g_n \xrightarrow{a.s.} f$ (resp. $g_n \xrightarrow{p} f, g_n \xrightarrow{L^p} f, g_n \xrightarrow{r} f$)

Soluție. Presupunem întâi că $f_n \xrightarrow{a.s.} f, h_n \xrightarrow{a.s.} f$ și să notăm

$$A_n = \{f_n \leq g_n \leq h_n\}, B = \{f_n \rightarrow f\}, C = \{h_n \rightarrow f\}$$

Prin ipoteză avem că $P(A_n) = P(B) = P(C) = 1$, deci avem și $P(B \cap C \cap (\bigcap_{n_j} A_n)) = 1$

Din incluziunea $B \cap C \cap (\bigcap_n A_n) \subset \{g_n \rightarrow f\}$ rezultă că $P(g_n \rightarrow f) = 1$.

Pentru cazul convergenței în probabilitate prin alegerea unui subșir (n_i) așa încît $f_{n_i} \xrightarrow{a.s.} f, h_{n_i} \xrightarrow{a.s.} f$ se observă că afirmația este o consecință a cazului convergenței a.s.

Altfel:

$$\begin{aligned} P(|g_n - f| > \varepsilon) &\leq P\left(|f_n - f| > \frac{\varepsilon}{2}\right) + P\left(|f_n - g_n| > \frac{\varepsilon}{2}\right) \\ &\leq P\left(|f_n - f| > \frac{\varepsilon}{2}\right) + P\left(h_n - f_n > \frac{\varepsilon}{2}\right) \leq P\left(|f_n - f| > \frac{\varepsilon}{2}\right) + \\ &+ P\left(|h_n - f| > \frac{\varepsilon}{4}\right) + P\left(|f_n - f| > \frac{\varepsilon}{4}\right) \rightarrow 0 \text{ cînd } n \rightarrow \infty \end{aligned}$$

Presupunem că $f_n \xrightarrow{L^p} f, h_n \xrightarrow{L^p} f$. Avem că

$$\begin{aligned} \|g_n - f\|_p &\leq \|f_n - f\|_p + \|g_n - f_n\|_p \leq \|h_n - f_n\|_p + \|h_n - f_n\|_p \leq \\ &\leq \|f_n - f\|_p + \|h_n - f\|_p + \|f_n - f\|_p = 2\|f_n - f\|_p + \\ &+ \|h_n - f\|_p \rightarrow 0 \text{ cînd } n \rightarrow \infty \end{aligned}$$

Altfel: în particular avem că $f_n \xrightarrow{p} f, h_n \xrightarrow{p} f$, deci din cazul convergenței în probabilitate obținem că $g_n \xrightarrow{p} f$. Din afirmația f) din rezumatul teoretic rezultă că este suficient să arătăm că familia $(|g_n|^p)_n$ este uniform integrabilă.

Ori aceasta este o consecință a faptului că familiile $(|f_n|^p)_n, (|h_n|^p)_n$ sînt uniform integrabile, deci în particular și familia $\{\max(|f_n|^p, |h_n|^p)\}_n$ este uniform integrabilă și a inegalității $|g_n|^p \leq \max(|f_n|^p, |h_n|^p)$.

În sfârșit să presupunem că $f_n \xrightarrow{r} f$, $h_n \xrightarrow{r} f$. Fie F_n, G_n, H_n , funcțiile de repartiție ale lui f_n, g_n, h_n . Afirmatia în acest caz (adică $g_n \xrightarrow{r} f$) este o consecință imediată a inegalităților $H_n(x) \leq G_n(x) \leq F_n(x)$ și a faptului că extremele acestor inegalități au aceeași limită.

3.10. Fie $f_1, \dots, f_n, \dots, f$ variabile aleatoare și $v(1), v(2), \dots, \dots, v(n), \dots$ un șir de variabile aleatoare cu valori în $\{1, 2, \dots\}$.

a) Dacă $f_n \xrightarrow{a.s.} f$ și $v(n) \xrightarrow{a.s.} \infty$ să se arate că $f_{v(n)} \xrightarrow{a.s.} f$

b) Dacă $f_1, \dots, f_n, \dots, v(1), \dots, v(n), \dots$ sînt independente și $f_n \xrightarrow{p} f$, $v(n) \xrightarrow{p} \infty$ să se arate că $f_{v(n)} \xrightarrow{p} f$.

Soluție. a) Pentru $\varepsilon > 0$ avem din ipoteză că

$$\lim_{m \rightarrow \infty} P(\sup_{n \geq m} |f_n - f| \geq \varepsilon) = 0, \quad \lim_{m \rightarrow \infty} P(\inf_{n \geq m} v(n) \geq \varepsilon) = 1.$$

Avem că

$$\begin{aligned} P(\sup_{n \geq N} |f_{v(n)} - f| \geq \varepsilon) &= P(\sup_{n \geq N} |f_{v(n)} - f| \geq \varepsilon, \inf_{n \geq N} v(n) \geq N_1) + \\ &+ P(\sup_{n \geq N} |f_{v(n)} - f| \geq \varepsilon, \inf_{n \geq N} v(n) < N_1) \leq P(\sup_{n \geq N_1} |f_n - f| \geq \varepsilon) + \\ &+ P(\inf_{n \geq N} v(n) < N_1) \end{aligned}$$

de unde făcînd pe rînd $N \rightarrow \infty$, $N_1 \rightarrow \infty$ obținem afirmația.

b) Presupunem că $f_n \xrightarrow{p} f$ și fie $\varepsilon > 0$. Avem

$$\begin{aligned} P(|f_{v(n)} - f| \geq \varepsilon) &= P(|f_{v(n)} - f| \geq \varepsilon, v(n) < N) + \\ &+ \sum_{k=N}^{\infty} P(|f_{v(n)} - f| \geq \varepsilon, v(n) = k) \leq P(v(n) < N) + \\ &+ \sum_{k=N}^{\infty} P(|f_k - f| \geq \varepsilon) P(v(n) = k) \leq P(v(n) < N) + \\ &+ P(v(n) \geq N) \cdot \sup_{k \geq N} P(|f_k - f| \geq \varepsilon) \leq P(v(n) < N) + \\ &+ \sup_{k \geq N} P(|f_k - f| \geq \varepsilon), \end{aligned}$$

de unde, făcînd pe rînd $n \rightarrow \infty$, $N \rightarrow \infty$, obținem că $f_{v(n)} \xrightarrow{p} f$.

Fie acum $f_n \xrightarrow{r} f$ și x punct de continuitate al repartiției lui f .

Avem că

$$P(f_{\nu(n)} < x) = P(f_{\nu(n)} < x, \nu(n) < N) + \sum_{k=N}^{\infty} P(f_{\nu(n)} < x, \nu(n) = k) \leq \\ \leq P(\nu(n) < N) + \sup_{k \geq N} P(f_k < x)$$

de unde făcînd pe rînd $n \rightarrow \infty$, $N \rightarrow \infty$ obținem că $f_{\nu(n)} \xrightarrow{p} f$.

3.11. Fie (f_n) un șir de variabile aleatoare independente, identic repartizate și cu $P \circ f_1^{-1} = C_{1,0}$. Să se arate că

$$\frac{1}{n^{1+\lambda}} \sum_{k=1}^n f_k \xrightarrow{a.s.} 0 \text{ pentru orice } \lambda > 0.$$

Soluție. Fie $g_n = \frac{1}{n^{1+\lambda}} \sum_{k=1}^n f_k$. Observăm întîi că $g_n \xrightarrow{p} 0$ deoarece $P(|g_n| \geq \varepsilon) \leq \frac{1}{\pi \varepsilon n^\lambda}$.

Conform lemei Borel-Cantelli $P(\lim_{n \rightarrow \infty} g_{2^n} = 0) = 1$. Pentru a arăta că $g_n \xrightarrow{a.s.} 0$ trebuie să arătăm că $P(\overline{\lim}_{n \rightarrow \infty} A_n) = 0$ sau echivalent $\lim_{n \rightarrow \infty} P\left(\bigcup_{i=n}^{\infty} A_i\right) = 0$, unde $A_n = \{|g_n| \geq \varepsilon\}$.

Pentru $n \geq 1$ alegem un întreg $\nu(n)$ așa încît $2^{\nu(n)} \leq n \leq 2^{\nu(n)+1}$. Deoarece $g_n \xrightarrow{p} 0$ pentru $\varepsilon > 0$ există N_ε astfel încît pentru $n \geq N_\varepsilon$ să avem

$$P\left(|h_{2^{\nu(n)+1}} + h_n| < \frac{\varepsilon}{2} n^{1+\lambda}\right) \geq \frac{1}{2} \text{ unde } h_n = \sum_{k=1}^n f_k.$$

Într-adevăr

$$\frac{h_n}{n^{1+\lambda}} \xrightarrow{p} 0 \text{ și } \frac{1}{n^{1+\lambda}} h_{2^{\nu(n)+1}} = \frac{2^{(\nu(n)+1)(1+\lambda)}}{n^{1+\lambda}} \frac{1}{2^{(\nu(n)+1)(1+\lambda)}} h_{2^{\nu(n)+1}} \xrightarrow{p} 0$$

deoarece $2^{(\nu(n)+1)(1+\lambda)} n^{-1-\lambda} \leq 2^{1+\lambda}$.

Fie

$$B_n = \left\{ |h_{2^{\nu(n)+1}} - h_n| < \frac{\varepsilon}{2} n^{1+\lambda} \right\}; C_n = \left\{ |h_{2^{\nu(n)+1}| \geq \frac{\varepsilon}{2} n^{1+\lambda} \right\}.$$

Observăm că $A_n \cap B_n \subset C_n$ și că A_i, B_i sînt independente pentru $i \leq n$. Pentru $n \geq N$ avem:

$$\begin{aligned}
 P\left(\bigcup_{i=n}^{\infty} A_i\right) &= P(A_n) + P(A_{n+1} \setminus A_n) + \dots + P\left(A_{n+k} \setminus \bigcup_{i=0}^{k-1} A_{n+i}\right) + \dots \leq \\
 &\leq 2 \left[P(A_n)P(B_n) + P(A_{n+1} \setminus A_n) P(B_{n+1}) + \dots + \right. \\
 &\quad \left. + P\left(A_{n+k} \setminus \bigcup_{i=0}^{k-1} A_{n+i}\right)P(B_{n+k}) + \dots \right] = 2[P(A_n \cap B_n) + \\
 &\quad + P(A_{n+1} \cap B_{n+1} \cap \bar{C}A_n) + \dots + P(A_{n+k} \cap B_{n+k} \cap \bar{C}A_n \cap \dots \\
 &\quad \dots \cap \bar{C}A_{n+k-1}) + \dots] \leq 2[P(A_n \cap B_n) + P(A_{n+1} \cap \bar{C}(A_n \cap B_n)) + \dots \\
 &\quad \dots P(A_{n+k} \cap B_{n+k} \cap \bar{C}(A_n \cap B_n) \cap \dots \cap \bar{C}(A_{n+k-1} \cap B_{n+k-1})) + \dots] = \\
 &= 2P\left(\bigcup_{i=n}^{\infty} (A_i \cup B_i)\right) \leq 2\bar{P}\left(\bigcup_{i=n}^{\infty} \bar{C}_i\right).
 \end{aligned}$$

Dar deoarece $g_{2n} \xrightarrow{a.s.} 0$ rezultă că $\lim_{n \rightarrow \infty} P\left(\bigcup_{i=n}^{\infty} \bar{C}_i\right) = 0$.

3.12. Pentru orice n fie f_1^n, \dots, f_n^n variabile aleatoare independente cu repartiția $N(0, 1)$.

Pentru $\varepsilon, r, > 0$ definim $B_{r, \varepsilon} = \{x = (x_1, \dots, x_n); (1 - \varepsilon)r \leq |x_1| + \dots + |x_n| \leq (1 + \varepsilon)r\}$, $C_{r, \varepsilon} = \{x = (x_1, \dots, x_n); (1 - \varepsilon)r \leq ||x|| = (x_1^2 + \dots + x_n^2)^{\frac{1}{2}} \leq (1 + \varepsilon)r\}$.

Să se arate că

$$\lim_{n \rightarrow \infty} P\left((f_1^n, \dots, f_n^n) \in B_{\sqrt{\frac{2}{\pi}}, \varepsilon}\right) = 1, \quad \lim_{n \rightarrow \infty} P\left((f_1^n, \dots, f_n^n) \in C_{\sqrt{\frac{2}{\pi}}, \varepsilon}\right) = 1.$$

Soluție. Avem că

$$P\left((f_1^n, \dots, f_n^n) \in B_{\sqrt{\frac{2}{\pi}}, \varepsilon}\right) = P\left(\left|\frac{|f_1^n| + \dots + |f_n^n|}{n \sqrt{\frac{2}{\pi}}} - 1\right| \leq \varepsilon\right)$$

și cum $|f_1^n|, \dots, |f_n^n|$ sînt variabile aleatoare independente, identic

$$\text{repartizate și } M(|f_1^n|) = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} x e^{-\frac{x^2}{2}} dx = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{-u} du = \sqrt{\frac{2}{\pi}}$$

$$D^2(|f_1^n|) = M[(f_1^n)^2] - \left(\sqrt{\frac{2}{\pi}}\right)^2 = 1 - \frac{2}{\pi},$$

prima relație din enunț rezultă prin aplicarea inegalității lui Cebîșev. Mai departe

$$(*) P\left((f_1^n, \dots, f_n^n) \in C_{\sqrt{n}, \varepsilon}\right) = P\left(\left|\frac{[(f_1^n)^2 + \dots + (f_n^n)^2]^{\frac{1}{2}}}{n^{\frac{1}{2}}} - 1\right| \leq \varepsilon\right)$$

$(f_1^n)^2, \dots, (f_n^n)^2$ sînt variabilele aleatoare independente, identic repartizate și $M[(f_1^n)^2] = 1$, $D^2[(f_1^n)^2] = M[(f_1^n)^4] - 1 = 3 - 1 = 2$ așa că din inegalitatea lui Cebîșev rezultă că

$$\frac{(f_1^n)^2 + \dots + (f_n^n)^2}{n} \xrightarrow{p} 1 \text{ de unde și } \frac{[(f_1^n)^2 + \dots + (f_n^n)^2]^{\frac{1}{2}}}{n^{\frac{1}{2}}} \xrightarrow{p} 1$$

sau echivalent că șirul din (*) converge către 1.

3.13. Fie f_1, \dots, f_n, \dots un șir de variabile aleatoare independente, identic repartizate și pentru orice n fie $u_n: R^n \rightarrow R$ o aplicație boreliană și simetrică (simetrică în sensul că $u_n(x_{i_1}, \dots, x_{i_n}) = u_n(x_1, \dots, x_n)$ pentru orice permutare i_1, \dots, i_n a numerelor $1, \dots, n$).

Să se arate că dacă $u_n(f_1, \dots, f_n) \xrightarrow{p} f$ atunci există o constantă a așa încît $f = a$ a.s.

Soluție. Fie $\varphi: R \rightarrow R$ o funcție continuă, injectivă și mărginită (de exemplu $\varphi(x) = \arctg x$).

Atunci din problema 3.6. rezultă că $\varphi(u_n(f_1, \dots, f_n)) \xrightarrow{p} \varphi(f)$. Problema este rezolvată dacă vom arăta că există o constantă c așa încît $\varphi(f) = c$ a.s.

Deoarece șirul $\varphi(u_n(f_1, \dots, f_n)) - \varphi(u_{2n}(f_1, \dots, f_{2n})) \xrightarrow{p} 0$ și este uniform mărginit rezultă din teorema de convergență dominată că

$$\lim_{n \rightarrow \infty} M[|\varphi(u_n(f_1, \dots, f_n)) - \varphi(u_{2n}(f_1, \dots, f_{2n}))|^2] = 0 \quad (1)$$

Din independență și identic repartizare obținem că

$$P \circ (f_{i_1}, \dots, f_{i_m})^{-1} = P \circ (f_1, \dots, f_m)^{-1}$$

pentru orice permutare i_1, \dots, i_m a numerelor $1, \dots, m$.

Avînd în vedere acest lucru (1) se mai scrie sub forma

$$\lim_{n \rightarrow \infty} M[|\varphi(u_n(f_{n+1}, \dots, f_{2n})) - \varphi(u_{2n}(f_{n+1}, \dots, f_{2n}, f_1, \dots, f_n))|^2] = 0 \quad (2)$$

sau ținând cont de simetria lui u_n

$$\lim_{n \rightarrow \infty} M [|\varphi(u_n(f_{n+1}, \dots, f_{2n})) - \varphi(u_{2n}(f_1, \dots, f_{2n}))|^2] = 0. \quad (3)$$

Din (1) și (2) obținem

$$\lim_{n \rightarrow \infty} M [|\varphi(u_n(f_1, \dots, f_n)) - \varphi(u_n(f_{n+1}, \dots, f_{2n}))|^2] = 0. \quad (4)$$

sau întrucît $M[\varphi(u_n(f_{i_1}, \dots, f_{i_n}))] = M[\varphi(u_n(f_1, \dots, f_n))]$ pentru orice permutare i_1, \dots, i_n a numerelor $1, \dots, n$ și variabilele aleatoare $\varphi(u_n(f_1, \dots, f_n)) - M[\varphi(u_n(f_1, \dots, f_n))]$, $\varphi(u_n(f_{n+1}, \dots, f_{2n})) - M[\varphi(u_n(f_{n+1}, \dots, f_{2n}))]$ sînt independente (din asociativitatea independenței) rezultă că (4) se mai scrie sub forma (folosindu-se egalitatea (10) din rezumatul teoretic):

$$\begin{aligned} & \lim_{n \rightarrow \infty} M [|\varphi(u_n(f_1, \dots, f_n)) - \varphi(u_n(f_{n+1}, \dots, f_{2n}))|^2] = \\ & = \lim_{n \rightarrow \infty} \{D^2[\varphi(u_n(f_1, \dots, f_n))] + D^2[\varphi(u_n(f_{n+1}, \dots, f_{2n}))]\} = \end{aligned} \quad (5)$$

$$= 2 \lim_{n \rightarrow \infty} D^2[\varphi(u_n(f_1, \dots, f_n))] = 0, \text{ deci reținem că}$$

$$\lim_{n \rightarrow \infty} D^2[\varphi(u_n(f_1, \dots, f_n))] = 0. \quad (6)$$

Pe de altă parte din teorema de convergență dominată avem și relația

$$\lim_{n \rightarrow \infty} D^2[\varphi(u_n(f_1, \dots, f_n))] = D^2[\varphi(f)]. \quad (7)$$

Din (6) și (7) deducem că $D^2[\varphi(f)] = 0$ cu alte cuvinte

$\varphi(f) = M[\varphi(f)]$ a.s. ($M[\varphi(f)]$ este deci constanta c căutată).

3.14. Fie f_1, \dots, f_n, \dots un șir de variabile aleatoare independente unde f_n are funcția de repartiție

$$F_n(x) = \begin{cases} 1 - \frac{1}{x+n} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

Să se arate că $f_n \xrightarrow{p} 0$ dar $\frac{1}{n} \sum_{l=1}^n f_l \xrightarrow{p} 0$ (aceasta arată că nu există

o normă pe spațiul variabilelor aleatoare care să dea convergența în probabilitate).

Soluție. Pentru $\varepsilon > 0$ avem că $P(f_n < \varepsilon) = F_n(\varepsilon) = 1 - \frac{1}{\varepsilon+n} \rightarrow 1$, cînd $n \rightarrow \infty$, deci $f_n \xrightarrow{p} 0$ (deoarece $P(f_n < 0) = 0$ rezultă că $f_n \geq 0$

ă.s.) Fie $g_n = \max(f_1, \dots, f_n)$. Ca să arătăm că $\frac{1}{n} \sum_{i=1}^n f_i \xrightarrow{p} 0$ este suficient să verificăm că $\frac{g_n}{n} \xrightarrow{p} 0$ (deoarece are loc inegalitatea $\frac{g_n}{n} \leq \frac{1}{n} \sum_{i=1}^n f_i$).

Or avem:

$$P\left(\frac{g_n}{n} < \varepsilon\right) = P(f_1 < n\varepsilon, \dots, f_n < n\varepsilon) = \prod_{i=1}^n P(f_i < n\varepsilon) = \\ = \prod_{i=1}^n \left(1 - \frac{1}{\varepsilon + n}\right) = \left(1 - \frac{1}{\varepsilon + n}\right)^n < \left[1 - \frac{1}{n(1+\varepsilon)}\right]^n \rightarrow e^{-(1+\varepsilon)} < 1,$$

deci $\overline{\lim}_{n \rightarrow \infty} P\left(\frac{g_n}{n} < \varepsilon\right) \leq e^{-(1+\varepsilon)} < 1$, fapt ce arată că $\frac{g_n}{n} \xrightarrow{p} 0$.

3.15. Fie f_1, \dots, f_n, \dots un șir de variabile aleatoare nenegative, de medie finită și fie f o variabilă aleatoare așa încît $f_n \xrightarrow{p} f$.

Să se arate că f are medie finită și $f_n \xrightarrow{L} f$ dacă și numai dacă f are medie finită și $\lim_{n \rightarrow \infty} M(f_n) = M(f)$.

Soluție. Este suficient să arătăm că dacă $M(f_n) \rightarrow M(f) < \infty$, atunci $f_n \xrightarrow{L} f$ (cealaltă implicație este imediată).

Putem presupune că $f_n \xrightarrow{a.s.} f$ (în caz contrar se trece la subșiruri). Deoarece $\min(f_n, f) \xrightarrow{a.s.} f$ rezultă prin folosirea teoremei de convergență dominată că $M[\min(f_n, f)] \rightarrow M(f)$.

Apoi din egalitatea $f + f_n = \min(f, f_n) + \max(f, f_n)$ rezultă și faptul că $M[\max(f, f_n)] \rightarrow M(f)$.

În fine din egalitatea $|f - f_n| = \max(f, f_n) - \min(f, f_n)$ rezultă că $M(|f_n - f|) \rightarrow 0$, cu alte cuvinte $f_n \xrightarrow{L} f$.

3.16. Fie f_1, \dots, f_n, \dots un șir de variabile aleatoare și $a \in \mathbb{R}^d$ o constantă. Să se arate că $f_n \xrightarrow{p} a$ dacă și numai dacă $f_n \xrightarrow{r} a$.

Soluție. Faptul că întotdeauna convergența în probabilitate implică convergența în repartiție este cunoscut.

Să presupunem acum că $f_n \xrightarrow{r} a$ și fie $\varepsilon > 0$ și $A = \mathcal{C} S_a(\varepsilon)$; $S_a(\varepsilon)$ este sfera închisă de centru a și raza ε . Deoarece $P(a \in \partial A) = 0$

rezultă din punctul d_4) al rezumatului teoretic că $P(f_n \in A) \rightarrow P(a \in A) = 0$, dar $P(f_n \in A) = P(|f_n - a| > \varepsilon)$, deci $P(|f_n - a| > \varepsilon) \rightarrow 0$, cu alte cuvinte $f_n \xrightarrow{p} a$.

3.17. Să se arate că $f^n = (f_1^n, \dots, f_d^n) \xrightarrow{r} f = (f_1, \dots, f_d)$ dacă și numai dacă pentru orice $\alpha_1, \dots, \alpha_d$ din R avem

$$\alpha_1 f_1^n + \dots + \alpha_d f_d^n \xrightarrow{r} \alpha_1 f_1 + \dots + \alpha_d f_d.$$

Soluție. Fie $u: R \rightarrow R$ continuă și mărginită și $L(x) = \alpha_1 x_1 + \dots + \alpha_d x_d$. Atunci deoarece $u \circ L: R^d \rightarrow R$ este continuă și mărginită avem că:

$$\begin{aligned} \int (\alpha_1 f_1^n + \dots + \alpha_d f_d^n) dP &= \int (u \circ L)(f_1^n, \dots, f_d^n) dP \rightarrow \\ &\rightarrow \int (u \circ L)(f_1, \dots, f_d) dP = \int u(\alpha_1 f_1 + \dots + \alpha_d f_d) dP \end{aligned}$$

deci $\alpha_1 f_1^n + \dots + \alpha_d f_d^n \xrightarrow{r} \alpha_1 f_1 + \dots + \alpha_d f_d$.

Implicația inversă: fie $t = (t_1, \dots, t_d)$. Prin ipoteză avem că $\langle t, f^n \rangle \xrightarrow{r} \langle t, f \rangle$, de unde folosind teorema de convergență a funcțiilor caracteristice obținem că

$$\varphi_{\langle t, f^n \rangle}(s) \rightarrow \varphi_{\langle t, f \rangle}(s) \text{ pentru orice } s \in R, \text{ adică } \varphi_{f_n}(st) \rightarrow \varphi_f(st).$$

În particular luînd $s = 1$ obținem că $\varphi_{f_n}(t) \rightarrow \varphi_f(t)$, deci $f_n \xrightarrow{r} f$ dacă apelăm din nou la teorema de convergență a funcțiilor caracteristice.

3.18. Fie $f, f_1, \dots, f_n, \dots$ variabilele aleatoare cu valori în $\{0, 1, 2, \dots\}$. Să se arate că $f_n \xrightarrow{r} f$ dacă și numai dacă $\lim_{n \rightarrow \infty} P(f_n = k) = P(f = k)$ pentru orice $k = 0, 1, \dots$

În particular dacă $P(f_n = 1) = \frac{1}{n}$, $P(f_n = 0) = 1 - \frac{1}{n}$ pentru orice n și f_1, \dots, f_n, \dots sînt independente să se arate că $f_n \xrightarrow{p} 0$ dar $f_n \xrightarrow{a.s.} 0$.

Soluție. Să presupunem că $f_n \xrightarrow{r} f$ și pentru $k \in \{0, 1, \dots\}$ alegem r_1, r_2 din R așa încît $k - 1 < r_1 < k < r_2 < k + 1$.

Atunci $P(f_n = k) = F_{f_n}(k+1) - F_{f_n}(k) = F_{f_n}(r_2) - F_{f_n}(r_1) \rightarrow F_f(r_2) - F_f(r_1) = P(f_f = k)$ (am folosit mai sus faptul că r_1, r_2 sînt puncte de continuitate pentru F_f).

Reciproc să presupunem [că $\lim_{n \rightarrow \infty} P(f_n = k) = P(f = k)$ pentru orice $k = 0, 1, \dots$. Deoarece $R \setminus \{0, 1, \dots\}$ este mulțimea punctelor de continuitate ale lui F_f rezultă că avem de arătat că $\lim_{n \rightarrow \infty} F_{f_n}(r) = F_f(r)$ pentru orice $r \in R \setminus \{0, 1, \dots\}$.

Fie k cel mai mare număr natural așa încît $k < r$. Atunci

$$F_{f_n}(r) = \sum_{l=1}^k P(f_n = l) \rightarrow \sum_{l=1}^k P(f = l) = F_f(r).$$

Fie acum

$$P(f_n = 1) = \frac{1}{n}, \quad P(f_n = 0) = 1 - \frac{1}{n}.$$

Deoarece

$$P(f_n = 1) \rightarrow P(0 = 1) = 0, \quad P(f_n = 0) \rightarrow P(0 = 0) = 1,$$

rezultă că $f_n \xrightarrow{r} 0$ deci și $f_n \xrightarrow{p} 0$ (vezi probelema 3.16)

Fie $A_n = \{f_n = 1\}$; A_n sînt evenimente independente și $\sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty$, deci din lema Borel-Cantelli rezultă că $P(\overline{\lim}_{n \rightarrow \infty} A_n) = P(f_n = 1 \text{ pentru o infinitate de } n) = 1$, fapt ce arată că f_n nu poate să convergă a.s. către 0.

3.19. Fie (f_n, g_n) două șiruri de variabile aleatoare și g o variabilă aleatoare.

Să se arate că dacă $f_n \xrightarrow{p} 0$, $g_n \xrightarrow{r} g$ atunci $f_n + g_n \xrightarrow{r} g$ și $g_n(1 + f_n) \xrightarrow{r} g$.

Soluție. Fie $\varepsilon > 0$ și $x \in R$ așa încît $x + \varepsilon$, $x - \varepsilon$ să fie puncte de continuitate pentru funcția de repartiție a lui g .

Din egalitatea

$$P(f_n + g_n < x) = P(|f_n| \leq \varepsilon, f_n + g_n < x) + P(|f_n| > \varepsilon, f_n + g_n < x)$$

și incluziunile

$$\{g_n < x - \varepsilon, |f_n| \leq \varepsilon\} \subset \{|f_n| \leq \varepsilon, f_n + g_n < x\} \subset \{g_n < x + \varepsilon, |f_n| \leq \varepsilon\}$$

rezultă inegalitățile

$$P(g_n < x - \varepsilon, |f_n| \leq \varepsilon) + P(|f_n| > \varepsilon, f_n + g_n < x) \leq P(f_n + g_n < x) \leq \\ \leq P(g_n < x + \varepsilon, |f_n| \leq \varepsilon) + P(|f_n| > \varepsilon, f_n + g_n < x)$$

de unde avînd în vedere că $P(A \cap B) \geq P(A) - P(\bar{B})$ obținem

$$P(g_n < x - \varepsilon) - P(|f_n| > \varepsilon) + P(|f_n| > \varepsilon, f_n + g_n < x) \leq \\ \leq P(f_n + g_n < x) \leq P(g_n < x + \varepsilon) + P(|f_n| > \varepsilon).$$

Utilizînd faptul că $P(|f_n| > \varepsilon, f_n + g_n < x) \leq P(|f_n| > \varepsilon) \rightarrow 0$, obținem prin trecere la limită cînd $n \rightarrow \infty$ în inegalitățile precedente

$$P(g < x - \varepsilon) \leq \overline{\lim}_{n \rightarrow \infty} P(f_n + g_n < x) \leq P(g < x + \varepsilon)$$

din care făcînd $\varepsilon \rightarrow 0$ deducem că

$$P(g \leq x) \leq \overline{\lim}_{n \rightarrow \infty} P(f_n + g_n < x) \leq P(g < x)$$

cu alte cuvinte

$$\lim_{n \rightarrow \infty} P(f_n + g_n < x) = P(g < x)$$

Putem proceda și altfel folosind teorema de convergență a funcțiilor caracteristice. Trebuie să arătăm că $\varphi_{f_n + g_n}(t) \rightarrow \varphi_g(t)$ adică

$$\int e^{i(t f_n + g_n)} dP \rightarrow \int e^{i t g} dP.$$

Avem

$$\left| \int e^{i(t f_n + g_n)} dP - \int e^{i t g} dP \right| \leq \left| \int e^{i(t f_n + g_n)} dP - \int e^{i t g_n} dP \right| + \\ + \left| \int e^{i t g_n} dP - \int e^{i t g} dP \right| \leq \int |e^{i t f_n} - 1| dP + |\varphi_{g_n}(t) - \varphi_g(t)|.$$

Deoarece $|e^{i t x} - 1|$ este continuă, rezultă din problema 3.6 că $|e^{i t f_n} - 1| \xrightarrow{p} 0$ și cu $|e^{i t x} - 1| \leq 2$ rezultă din teorema de convergență dominată că

$$\int |e^{i t f_n} - 1| dP \rightarrow 0.$$

Pe de altă parte $\varphi_{g_n}(t) \rightarrow \varphi_g(t)$ conform teoremei de convergență a funcțiilor caracteristice. Acum afirmația rezultă de îndată.

În continuare alegem $\varepsilon > 0$ și $x \in R$ așa încît $x, \frac{x}{1-\varepsilon}, \frac{x}{1+\varepsilon}$ să fie puncte de continuitate ale funcției de repartiție a lui g .

Avem

$$\begin{aligned} P(g_n(1+f_n) < x) &= P(g_n(1+f_n) < x, |f_n| \leq \varepsilon) + \\ &+ P(g_n(1+f_n) < x, |f_n| > \varepsilon) \\ \left\{g_n < \frac{x}{1-\varepsilon}, |f_n| \leq \varepsilon\right\} &\subset \{g_n(1+f_n) < x, |f_n| \leq \varepsilon\} \subset \\ &\subset \left\{g_n < \frac{x}{1+\varepsilon}, |f_n| \leq \varepsilon\right\}. \end{aligned}$$

Se continuă apoi raționamentul de la cazul precedent.

3.20. Fie f_1, \dots, f_n, \dots un șir de variabile aleatoare avînd repartiții geometrice de parametri p_1, \dots, p_n, \dots unde $p_n \rightarrow 1$. Să se arate că $g_n = (1-p_n)f_n$ converge slab către repartiția exponențială de parametru 1.

Soluție. Vom aplica teorema de convergență a funcțiilor caracteristice (vezi g) din rezumatul teoretic).

Funcția caracteristică a lui f_n este $\varphi_n(t) = \frac{1-p_n}{1-p_n e^{it}}$, deci funcția caracteristică a lui g_n este $\psi_n(t) = \frac{1-p_n}{1-p_n e^{it(1-p_n)}}$. Avem prin aplicarea regulii lui l'Hospital că

$$\lim_{n \rightarrow \infty} \psi_n(t) = \lim_{x \rightarrow 1} \frac{1-x}{1-xe^{it(1-x)}} = \lim_{x \rightarrow 1} \frac{-1}{-e^{it(1-x)} + itxe^{it(1-x)}} = \frac{1}{1-it} = \psi(t).$$

Însă $\psi(t)$ este funcția caracteristică a repartiției e_1 , deci conform teoremei de convergență a funcțiilor caracteristice rezultă că $g_n \Rightarrow e_1$.

Putem proceda direct astfel: va trebui să arătăm că $\lim_{n \rightarrow \infty} P(g_n < x) = 1 - e^{-x}$, $x \geq 0$ ($1 - e^{-x}$ este funcția de repartiție a lui e_1 pentru $x \geq 0$; pentru $x < 0$ relația de demonstrat este imediată).

Avem

$$\begin{aligned}
 P(g_n < x) &= P\left(f_n < \frac{x}{1-p_n}\right) = \\
 &= \begin{cases} \sum_{i=0}^{\left[\frac{x}{1-p_n}\right]} (1-p_n)p_n^i & \text{dacă } \frac{x}{1-p_n} \text{ nu este întreg} \\ \sum_{i=0}^{\left[\frac{x}{1-p_n}\right]-1} (1-p_n)p_n^i & \text{în caz contrar} \end{cases} = \\
 &= \begin{cases} (1-p_n)p_n \frac{1-p_n^{\left[\frac{x}{1-p_n}\right]+1}}{1-p_n} & \\ (1-p_n)p_n \frac{1-p_n^{\left[\frac{x}{1-p_n}\right]}}{1-p_n} & \end{cases} = \\
 &= \begin{cases} p_n(1-p_n^{\left[\frac{x}{1-p_n}\right]+1}) & \text{dacă } \frac{x}{1-p_n} \text{ nu este întreg} \\ p_n(1-p_n^{\left[\frac{x}{1-p_n}\right]}) & \text{în caz contrar} \end{cases}
 \end{aligned}$$

Prin urmare totul se reduce la a arăta că $\lim_{n \rightarrow \infty} p_n^{\left[\frac{x}{1-p_n}\right]} = e^{-x}$.

Avem că

$$\begin{aligned}
 p_n^{\left[\frac{x}{1-p_n}\right]} &= (1+p_n-1)^{\left[\frac{x}{1-p_n}\right]} = \left[(1+p_n-1)^{\frac{1}{p_n-1}}\right]^{-(1-p_n)\left[\frac{x}{1-p_n}\right]} \rightarrow \\
 &\rightarrow e^{-\lim_{n \rightarrow \infty} (1-p_n)\left[\frac{x}{1-p_n}\right]}.
 \end{aligned}$$

Din inegalitățile

$$\begin{aligned}
 x-1+p_n &= (1-p_n)\left(\frac{x}{1-p_n}-1\right) \leq (1-p_n)\left[\frac{x}{1-p_n}\right] \leq \\
 &\leq (1-p_n)\left(\frac{x}{1-p_n}+1\right) = x+1-p_n
 \end{aligned}$$

obținem că $\lim_{n \rightarrow \infty} (1-p_n)\left[\frac{x}{1-p_n}\right] = x$.

3.21. Dacă f este o variabilă aleatoare cu valori în $[0, 1]$ și cu densitatea $p(x)$ continuă să se arate că șirul $g_n = \{nf\}$ ($\{x\}$ este partea zecimală a lui x) converge în repartiție către repartiția uniformă pe $[0, 1]$.

Soluție. Avem de arătat că pentru orice $0 \leq x \leq 1$, $\lim_{n \rightarrow \infty} P(g_n < x) = x$.

Putem scrie

$$\begin{aligned} P(g_n < x) &= \sum_{k=0}^{n-1} P(k \leq nf < k+1, g_n < x) = \\ &= \sum_{k=0}^{n-1} P\left(\frac{k}{n} \leq f < \frac{k+x}{n}\right) = \sum_{k=0}^{n-1} \int_{\frac{k}{n}}^{\frac{k+x}{n}} p(y) dy = \\ &= \sum_{k=0}^{n-1} \frac{x}{n} p(\xi_{n,k}) = x \sum_{k=0}^{n-1} \frac{1}{n} p(\xi_{n,k}) \rightarrow x \int_0^1 p(y) dy = x \end{aligned}$$

(teorema de medie afirmă că există $\xi_{n,k} \in \left[\frac{k}{n}, \frac{k+x}{n}\right]$).

3.22. Fie A o submulțime a lui R și pentru orice $x \in A$ fie $f_1(x), \dots, f_n(x), \dots$ un șir de variabile aleatoare cu proprietățile

$$M[f_n(x)] = x, \sup_{x \in A} D^2[f_n(x)] \rightarrow 0.$$

Să se arate că pentru orice funcție φ uniform continuă și mărginită are loc egalitatea $\lim_{n \rightarrow \infty} M[\varphi(f_n(x))] = \varphi(x)$ uniform pe A , cu alte cuvinte $f_n(x) \xrightarrow{r} x$ uniform pentru $x \in A$.

Soluție. Pentru $\varepsilon > 0$ fie $\delta > 0$ așa încît $|\varphi(x) - \varphi(y)| \leq \varepsilon$ dacă $|x - y| \leq \delta$.

Atunci avem

$$\begin{aligned} |M[\varphi(f_n(x))] - \varphi(x)| &\leq M[|\varphi(f_n(x)) - \varphi(x)|] = \\ &= \int_{|f_n(x) - x| \leq \delta} |\varphi(f_n(x)) - \varphi(x)| dP + \int_{|f_n(x) - x| > \delta} |\varphi(f_n(x)) - \varphi(x)| dP \leq \\ &\leq \varepsilon + 2 \sup_y |\varphi(y)| P(|f_n(x) - x| > \delta) \leq \\ &\leq \varepsilon + 2 \sup_y |\varphi(y)| \cdot \frac{1}{\delta^2} D^2[f_n(x)] \end{aligned}$$

de unde $\overline{\lim}_{n \rightarrow \infty} \sup_{x \in A} |M[\varphi(f_n(x))] - \varphi(x)| \leq \varepsilon$ și cum ε este arbitrar obținem că $\lim_{n \rightarrow \infty} \sup_{x \in A} |M[\varphi(f_n(x))] - \varphi(x)| = 0$.

Observație. Se poate obține ca o consecință teorema lui Bernstein. Dacă $f: [0, 1] \rightarrow R$ este o funcție continuă atunci șirul de polinoame $B_n(f, x) = \sum_{k=0}^n C_n^k f\left(\frac{k}{n}\right) x^k (1-x)^{n-k}$ converge uniform către f .

Într-adevăr se ia $\bar{f}_1(x), \dots, \bar{f}_n(x), \dots$ un șir de variabile aleatoare independente, identic repartizate așa încît $P(f_n(x) = 1) = x$, $P(f_n(x) = 0) = 1 - x$. Atunci se aplică rezultatul problemei șirului

$$\bar{f}_n(x) = \frac{\bar{f}_1(x) + \dots + \bar{f}_n(x)}{n} \text{ și funcției } f.$$

3.23. Să se arate că în teorema de convergență a funcțiilor caracteristice cerința ca φ să fie continuă în 0 nu se poate suprima.

Soluție. Fie F_n repartiția pe R cu funcția de repartiție

$$\bar{F}_n(x) = \begin{cases} 0, & x \leq -n \\ \frac{1}{2} + \frac{x}{2n}, & x \in [-n, n] \\ 1, & x \geq n \end{cases}$$

Avem

$$\varphi_{F_n}(t) = \frac{\sin nt}{nt} \text{ și } \lim_{n \rightarrow \infty} \varphi_{F_n}(t) = \varphi(t) = \begin{cases} 1 & t = 0 \\ 0 & t \neq 0 \end{cases}$$

Evident φ nu este continuă în origine. Pe de altă parte pentru orice x avem că $\lim_{n \rightarrow \infty} \bar{F}_n(x) = \bar{F}(x) = \frac{1}{2}$, deci \bar{F} nu este funcție de repartiție și în particular φ nu este funcția caracteristică a unei repartiții.

3.24. Să se arate că în raport cu convergența slabă a repartițiilor :

- Repartiția Dirac ε_a depinde continuu în a și mulțimea $\{\varepsilon_a; a \in R^d\}$ este închisă;
- repartiția Poisson π_λ depinde continuu în λ și mulțimea $\{\pi_\lambda; \lambda \geq 0\}$ este închisă.
- Repartiția normală $N(a, \sigma^2)$ depinde continuu în (a, σ) și mulțimea $\{N(a, \sigma^2); a \in R, \sigma \geq 0\}$ este închisă.
- Repartiția Cauchy $C_{\alpha, \beta}$ depinde continuu în (α, β) și mulțimea $\{C_{\alpha, \beta}; \alpha \geq 0, \beta \in R\}$ este închisă.

(Prin convenție $\pi_0 = \varepsilon_0$, $N(a, 0) = \varepsilon_a$, $C_{0, \beta} = \varepsilon_\beta$)

Soluție. Pentru continuitate este suficient să observăm (conform teoremei de convergență a funcțiilor caracteristice) că funcțiile caracteristice depind continuu de parametrii listați mai sus.

Să arătăm închiderea mulțimilor, respective.

Fie $\varepsilon_{a_n} \Rightarrow F$. Atunci $\varphi_{\varepsilon_{a_n}}(t) = \exp(i\langle t, a_n \rangle) \rightarrow \varphi_F(t)$.

Arătăm că șirul (a_n) este mărginit. Prin absurd dacă (a_n) nu este mărginit atunci el conține un subșir $a'_n \rightarrow \infty$. Din teorema de convergență dominată rezultă că pentru orice s avem

$$\int_0^s \exp(i\langle t, a'_n \rangle) dt \rightarrow \int_0^s \varphi_F(t) dt.$$

Dar

$$\left| \int_0^s \exp(i\langle t, a'_n \rangle) dt \right| = \left| \prod_{j=1}^d \left[\frac{\exp\{ia'_n(j) s_j\} - 1}{ia'_n(j)} \right] \right| \leq \frac{1}{\prod_{j=1}^d a'_n(j)} \rightarrow 0$$

deci $\int_0^s \varphi_F(t) dt = 0$ și cum φ_F este continuă rezultă că $\varphi_F = 0$, fapt imposibil ($\varphi_F(0) = 1$).

Arătăm acum că (a_n) este convergent. Prin absurd dacă nu este așa putem găsi $a' \neq a''$ și două subșiruri (a'_n) , (a''_n) așa încît $a'_n \rightarrow a'$, $a''_n \rightarrow a''$. În particular obținem că $\varphi_F(t) = \exp(i\langle t, a' \rangle) = \exp(i\langle t, a'' \rangle)$, deci $\exp(i\langle t, a' - a'' \rangle) = 1$ ceea ce nu este posibil (se ia $t = \frac{\pi}{a' - a''}$).

Prin urmare $a_n \rightarrow a$ și deci în particular $\varphi_F(t) = \exp(i\langle t, a \rangle)$ sau echivalent $F = \varepsilon_a$.

Fie acum $\pi_{\lambda_n} \Rightarrow F$. Atunci $\varphi_{\pi_{\lambda_n}}(t) = \exp\{\lambda_n(e^{it} - 1)\} \rightarrow \varphi_F(t)$ și în particular $|\varphi_{\pi_{\lambda_n}}(t)| = \exp\{-\lambda_n(1 - \cos t)\} \rightarrow |\varphi_F(t)|$, deci $\lambda_n \rightarrow \lambda$ cu $0 \leq \lambda < \infty$.

Fie $N(a_n, \sigma_{n_k}^2) \Rightarrow F$. Arătăm că (σ_n) este șir mărginit. În caz contrar am extrage $\sigma_{n_k} \rightarrow \infty$ și am avea

$$\begin{aligned} |\varphi_F(t)| &= \lim_{n \rightarrow \infty} |\varphi_{N(a_{n_k}, \sigma_{n_k}^2)}(t)| = \lim_{n \rightarrow \infty} \left| \exp\left(ita_{n_k} - \frac{1}{2}t^2\sigma_{n_k}^2\right) \right| = \\ &= \lim_{n \rightarrow \infty} \exp\left(-\frac{1}{2}t^2\sigma_{n_k}^2\right) = \begin{cases} 1 & t = 0 \\ 0 & t \neq 0 \end{cases} \end{aligned}$$

ceea ce contrazice continuitatea lui φ_F .

Arătăm că (a_n) este mărginit. Prin absurd fie un subșir $a'_n \rightarrow \infty$ și trecînd eventual la alt subșir putem presupune că $\sigma'_n \rightarrow \sigma$.

Atunci

$$\varphi_{\sigma'_n}(t) = \lim_{n \rightarrow \infty} \varphi_{N(a'_n, \sigma'^n_2)}(t) \exp\left(\frac{1}{2} t^2 \sigma'^n_2\right) \rightarrow \varphi_F(t) \exp\left(\frac{1}{2} \sigma^2 t^2\right)$$

de unde ar rezulta că $\varphi_F \equiv 0$ (vezi considerațiile precedente legate de repartiția Dirac).

Arătăm acum că (a_n) , (σ_n) sînt convergente. Într-adevăr, dacă de exemplu σ_n nu ar converge, am putea extrage două subșiruri

$$\sigma'_n \rightarrow \sigma', \quad \sigma''_n \rightarrow \sigma'', \quad \sigma' \neq \sigma'' \quad \text{și} \quad a'_n \rightarrow a', \quad a''_n \rightarrow a''.$$

Ar urma că $\varphi_{N(a', \sigma')} = \varphi_{N(a'', \sigma'')}$ fost incompatibil cu $\sigma' \neq \sigma''$.

În fine fie $C_{\alpha_n, \beta_n} \Rightarrow F$. Atunci $|\varphi_{C_{\alpha_n, \beta_n}}(t)| = \exp(it\beta_n - |t|\alpha_n) = \exp(-|t|\alpha_n) \rightarrow |\varphi_F(t)|$, deci $\alpha_n \rightarrow \alpha \in [0, \infty]$.

Dacă $\alpha = +\infty$, atunci $|\varphi_F(t)| = 0$ pentru orice $t \neq 0$, fapt imposibil (contrazice continuitatea lui φ_F).

Ca la repartiția Dirac și normală se arată că $\beta_n \rightarrow \beta$.

3.25. Să se arate că dacă $N_d(a^n, A^n) \Rightarrow F$ atunci $a^n \rightarrow a$, $A^n \rightarrow A$ (deci în particular $F = N_d(a, A)$, dacă A este pozitiv definită).

Soluție. Fie φ_n , φ funcțiile caracteristice corespunzătoare lui $N_d(a^n, A^n)$ și F . Se știe că $\varphi_n(t) = \exp(i\langle t, a^n \rangle) - \frac{1}{2} \langle A^n t, t \rangle$ și că $\varphi_n \rightarrow \varphi$ uniform pe orice compact, deci în particular $\lim_{n \rightarrow \infty} |\varphi_n(t)| = \exp\left(-\frac{1}{2} \langle A^n t, t \rangle\right) = |\varphi(t)|$ uniform pe orice compact.

Întrucît φ este continuă și $\varphi(0) = 1$ rezultă că există $\delta > 0$ așa încît $\varphi(t) \neq 0$ în cubul cu laturile egale cu 2δ și centrul în origine. Dacă $t = (\delta y)_i$, atunci $\langle A^n t, t \rangle = A^n_{jj} \rightarrow \ln |\varphi(t)|$.

Dacă $t = (0, \dots, 0, \delta, 0, \dots, 0, \delta, 0, \dots)$ atunci $\langle A^n t, t \rangle = \delta^2(A^n_{jj} + A_{ii} + 2A^n_{ji})$, de unde rezultă că $\lim_{n \rightarrow \infty} A^n_{jj}$ există, cu alte cuvinte $A^n \rightarrow A$.

$$\text{Fie } \psi_n(t) = \varphi_n(t) \exp\left(\frac{1}{2} \langle A^n t, t \rangle\right) = \exp(i\langle a^n, t \rangle)$$

$$\psi(t) = \varphi(t) \exp\left(\frac{1}{2} \langle A t, t \rangle\right)$$

Din convergența lui φ_n și A^n rezultă convergența lui $\exp(i\langle t, a^n \rangle)$, în particular luînd $t = (\delta_{jj})_j$ obținem că $\lim_{n \rightarrow \infty} \exp(ia_j^n)$ există. Fie $S = \{0\} \times \dots \times \{0\} \times [0, 1] \times \{0\} \times \dots \times \{0\}$. Atunci $\varphi_n \rightarrow \varphi$ uniform

pe S și $\psi_n(t) = \exp(ia_j^n t_j)$, $\psi(t) = \varphi(t) \exp\left(\frac{1}{2} A_{jj} t_j^2\right)$ dacă $t \in S$.

Deoarece

$$\begin{aligned} & \left| \exp\left(\frac{1}{2} A_{jj}^n t_j^2\right) - \exp\left(\frac{1}{2} A_{jj} t_j^2\right) \right| = \\ & = \begin{cases} \exp\left(\frac{1}{2} A_{jj} t_j^2\right) \left[\exp\left\{\frac{1}{2} (A_{jj}^n - A_{jj}) t_j^2\right\} - 1 \right] & \text{dacă } A_{jj}^n \geq A_{jj} \\ \exp\left(\frac{1}{2} A_{jj}^n t_j^2\right) \left[\exp\left\{\frac{1}{2} (A_{jj} - A_{jj}^n) t_j^2\right\} - 1 \right] & \text{dacă } A_{jj}^n \leq A_{jj} \end{cases} \end{aligned}$$

rezultă că maximul pe S este luat în $t_j = 1$.

Dar pentru $t_j = 1$ din $A_{jj}^n \rightarrow A_{jj}$ rezultă că $\lim_{n \rightarrow \infty} \left| \exp\left(\frac{1}{2} A_{jj}^n\right) - \exp\left(\frac{1}{2} A_{jj}\right) \right| = 0$, așa că $\lim_{n \rightarrow \infty} \exp\left(\frac{1}{2} A_{jj}^n t_j^2\right) = \exp\left(\frac{1}{2} A_{jj} t_j^2\right)$ uniform pentru $t_j \in [0, 1]$ și deci ψ_n converge uniform pe S .

Atunci pentru $\varepsilon > 0$ există n_0 așa încît $|\psi_{n'}(t) - \psi_{n''}(t)| \leq \varepsilon$ pentru $t \in S$, $n', n'' \geq n_0$. Fie Γ_n drumul definit de $z(t) = \psi_n(t) = \exp(it_j - a_j^n)$, $t \in S$. Γ_n pleacă din $z = 1$ și merge pe cercul unitate în sensul determinat de sign a_j^n pînă în punctul $z = \exp(ia_j^n)$.

Din ultima inegalitate obținem că

$$\int_{\Gamma_{n'}} \frac{1}{z} dz - \int_{\Gamma_{n''}} \frac{1}{z} dz \leq \varepsilon \text{ dacă } n', n'' \geq n_0.$$

Dar $\int_{\Gamma_n} \frac{1}{z} dz = \int_0^1 ia_j^n dt_j = ia_j^n$, deci $|a_j^{n'} - a_j^{n''}| \leq \varepsilon$ dacă $n', n'' \geq n_0$.

Prin urmare șirul $(a_j^n)_n$ este Cauchy, deci convergent.

3.26. Fie $f_n \in N(a_n, \sigma_n^2)$ și f o variabilă aleatoare. Să se arate că $f_n \xrightarrow{p} f$ dacă și numai dacă $f_n \xrightarrow{L^p} f$ pentru orice $1 \leq p < \infty$.

Soluție. Să presupunem că $f_n \xrightarrow{p} f$. Atunci din problema 3.24 rezultă că $a_n \rightarrow a$, $\sigma_n \rightarrow \sigma$.

Prin urmare, șirurile $M[\exp(\pm \alpha f_n)] = \exp\left(\pm a_n - \frac{\alpha^2 \sigma_n^2}{2}\right)$ sînt mărginite (fiind convergente) pentru orice $\alpha \in R$, deci în particular obținem că $A = \sup M[\exp(f_n) + \exp(-f_n)] < \infty$.

Cum avem $\sup_n M(|f_n|^{p+2}) \leq A$ rezultă că familia $(|f_n|^p)_n$ este uniform integrabilă (vezi rezumatul teoretic) și cum $f_n \xrightarrow{p} f$ rezultă din punctul f) al rezumatului teoretic că $f_n \xrightarrow{L^p} f$. Cealaltă implicație este cunoscută.

§2. Legea numerelor mari. Teorema limită centrală

3.27. Fie (f_n) , (g_n) două șiruri de variabile aleatoare cu proprietățile:

a) $M(f_n) = M(g_n)$ pentru orice n .

b) $\sum_{n=1}^{\infty} P(f_n \neq g_n) < \infty$.

c) (f_n) satisface legea tare (legea slabă) a numerelor mari.

Să se arate că și (g_n) satisface legea tare (legea slabă) a numerelor mari.

Soluție. Presupunem întâi că (f_n) satisface legea tare a numerelor mari.

Fie $A = \lim_{n \rightarrow \infty} \{f_n = g_n\}$. Din proprietatea b) și lema Borel-Cantelli rezultă că $P(A) = 1$.

Din incluziunea

$$\left\{ \frac{1}{n} \sum_{i=1}^n [f_i - M(f_i)] \rightarrow 0 \right\} \cap A \subset \left\{ \frac{1}{n} \sum_{i=1}^n [g_i - M(g_i)] \rightarrow 0 \right\}$$

și faptul că $P\left(\frac{1}{n} \sum_{i=1}^n [f_i - M(f_i)] \rightarrow 0\right) = 1$ deci și

$$P\left(\left\{ \frac{1}{n} \sum_{i=1}^n [f_i - M(f_i)] \rightarrow 0 \right\} \cap A\right) = 1 \text{ rezultă că}$$

$P\left(\frac{1}{n} \sum_{i=1}^n [g_i - M(g_i)] \rightarrow 0\right) = 1$, cu alte cuvinte (g_n) satisface legea tare a numerelor mari.

Presupunem acum că (f_n) satisface legea slabă a numerelor mari. Pentru a arăta că (g_n) satisface legea slabă este suficient de arătat că din orice subșir al lui $\frac{1}{n} \sum_{i=1}^n [g_i - M(g_i)]$ se poate extrage alt subșir a.s. convergent către zero. Fie deci (n') un subșir al lui (n) . Deoarece $\frac{1}{n} \sum_{i=1}^n [f_i - M(f_i)] \xrightarrow{p} 0$ rezultă că (n') conține un subșir (n'') așa încît $\frac{1}{n''} \sum_{i=1}^{n''} [f_i - M(f_i)] \xrightarrow[n'' \rightarrow \infty]{a.s.} 0$, deci conform celor de mai sus și $\frac{1}{n''} \sum_{i=1}^{n''} [g_i - M(g_i)] \xrightarrow[n'' \rightarrow \infty]{a.s.} 0$.

3.28. Fie (f_n^1, f_n^2) un șir de variabile aleatoare cu valori în R^2 care sînt independente, identic repartizate.

Dacă $0 < a_1 = M(f_1^1) < \infty$, $0 < a_2 = M(f_1^2) < \infty$ să se arate că $g_n = \frac{f_1^1 + \dots + f_n^1}{f_1^1 + \dots + f_n^2}$ converge a.s. către $\frac{a_1}{a_2}$.

Soluție. Putem scrie $g_n = \frac{f_1^1 + \dots + f_n^1}{n} \cdot \frac{n}{f_1^1 + \dots + f_n^2}$. Din legea numerelor mari rezultă că $\frac{f_1^i + \dots + f_n^i}{n} \xrightarrow{a.s.} a_i$, $i = 1, 2$ deci $g_n \xrightarrow{a.s.} \frac{a_1}{a_2}$.

3.29. Fie (f_n) un șir de variabile aleatoare independente, identic repartizate ce iau valori în $(0, 1]$ aproape sigur și fie $u, v: (0, 1] \rightarrow (0, \infty)$ funcții continue așa încît $u \leq cv$, c constantă.

Să se arate că

$$\lim_{n \rightarrow \infty} M \left(\frac{u(f_1) + \dots + u(f_n)}{v(f_1) + \dots + v(f_n)} \right) = \frac{M[u(f_1)]}{M[v(f_1)]}.$$

Soluție. Deoarece $g_n = \frac{u(f_1) + \dots + u(f_n)}{v(f_1) + \dots + v(f_n)} \leq c$ rezultă că g_n are medie finită.

Din legea tare a numerelor mari rezultă că

$$\frac{u(f_1) + \dots + u(f_n)}{n} \xrightarrow{a.s.} M[u(f_1)]$$

$$\frac{v(f_1) + \dots + v(f_n)}{n} \xrightarrow{a.s.} M[v(f_1)]$$

deci

$$\frac{u(f_1) + \dots + u(f_n)}{v(f_1) + \dots + v(f_n)} \xrightarrow{a.s} \frac{M[u(f_1)]}{M[v(f_1)]},$$

de unde prin aplicarea teoremei de convergență dominată obținem egalitatea din enunț.

Observație. Luind (f_n) cu repartiția uniformă pe $[0, 1]$ obținem egalitatea:

$$\lim_{n \rightarrow \infty} \int_{[0, 1]^n} \frac{u(x_1) + \dots + u(x_n)}{v(x_1) + \dots + v(x_n)} dx_1 \dots dx_n = \frac{\int_0^1 u(x) dx}{\int_0^1 v(x) dx}.$$

3.30. Fie $u: \mathbb{R} \rightarrow \mathbb{R}$ continuă și mărginită. Să se arate egalitățile următoare

a) $\lim_{n \rightarrow \infty} \sum_{k=0}^{\infty} u\left(x + \frac{k}{n}\right) e^{-nh} \frac{(nh)^k}{k!} = u(x+h)$ pentru orice $x \in \mathbb{R}$ și $h > 0$.

b) $\lim_{n \rightarrow \infty} \int_{[0, 1]^n} u\left(\frac{1}{n} \sum_{k=1}^n x_k\right) dx_1 \dots dx_n = u\left(\frac{1}{2}\right).$

c) $\lim_{n \rightarrow \infty} \int_{[0, 1]^n} u(\sqrt[n]{x_1 \dots x_n}) dx_1 \dots dx_n = u\left(\frac{1}{e}\right).$

d) $\lim_{n \rightarrow \infty} \int_{[0, 1]^n} \cos^{2m} \frac{\pi}{2n} (x_1 + \dots + x_n) dx_1 \dots dx_n = 2^{-m}.$

Soluție. a) Fie $f_1, \dots, f_n \dots$ un șir de variabile aleatoare independente, identic repartizate, cu f_1 având repartiția π_h . Din legea tare a numerelor mari avem că $\frac{f_1 + \dots + f_n}{n} \xrightarrow{a.s} h$, deci

$u\left(x + \frac{f_1 + \dots + f_n}{n}\right) \xrightarrow{a.s} u(x+h)$ și prin urmare $M\left[u\left(x + \frac{f_1 + \dots + f_n}{n}\right)\right] \rightarrow u(x+h)$ dacă avem în vedere teorema de convergență dominată.

$$\text{Însă } M\left[u\left(x + \frac{f_1 + \dots + f_n}{n}\right)\right] = \int_{\bigcup_{k=0}^{\infty} \{f_1 + \dots + f_n = k\}} u\left(x + \frac{f_1 + \dots + f_n}{n}\right) dP = \\ = \sum_{k=0}^{\infty} u\left(x + \frac{k}{n}\right) P(f_1 + \dots + f_n = k) = \sum_{k=0}^{\infty} u\left(x + \frac{k}{n}\right) e^{-nh} \frac{(nh)^k}{k!}.$$

b) Fie f_1, \dots, f_n, \dots un șir de variabile aleatoare independente, identic repartizate cu f_1 avînd repartiția uniformă pe $[0, 1]$.

Din legea tare a numerelor mari avem $\frac{f_1 + \dots + f_n}{n} \xrightarrow{a.s.} \frac{1}{2}$ deci și $u\left(\frac{f_1 + \dots + f_n}{n}\right) \xrightarrow{a.s.} u\left(\frac{1}{2}\right)$ și prin urmare $M\left[u\left(\frac{f_1 + \dots + f_n}{n}\right)\right] \rightarrow u\left(\frac{1}{2}\right)$ conform teoremei de convergență dominată. Însă conform formulei de transport $M\left[u\left(\frac{f_1 + \dots + f_n}{n}\right)\right] = \int_{[0, 1]^n} u\left(\frac{x_1 + \dots + x_n}{n}\right) dx_1 \dots dx_n.$

c) Se consideră șirul de la punctul b. Din legea tare a numerelor mari rezultă că $\ln \sqrt[n]{f_1 \dots f_n} = \frac{\ln f_1 + \dots + \ln f_n}{n} \xrightarrow{a.s.} M(\ln f_1) = -1$, deci $\sqrt[n]{f_1 \dots f_n} \xrightarrow{a.s.} \frac{1}{e}$ și în particular $u(\sqrt[n]{f_1 \dots f_n}) \xrightarrow{a.s.} u\left(\frac{1}{e}\right).$

Mai departe se continuă ca la punctul b).

d) Se aplică punctul b) luînd $u(x) = \cos^{2m} \frac{\pi}{2} x.$

3.31. Să se calculeze limitele

a) $\lim_{n \rightarrow \infty} \int_{\{x_1^2 + x_2^2 + \dots + x_n^2 \leq \sqrt{n}\} \cap [0, 1]^n} dx_1 dx_2 \dots dx_n$

b) $\lim_{n \rightarrow \infty} \int_{\{x_1^2 + x_2^2 + \dots + x_n^2 \leq \frac{n}{4}\} \cap [0, 1]^n} dx_1 dx_2 \dots dx_n$

c) $\lim_{n \rightarrow \infty} \int_{\{x_1^2 + \dots + x_n^2 \leq \frac{n}{2}\}} dx_1 dx_2 \dots dx_n.$

Soluție. a) Considerăm un câmp de probabilitate $\{E, \mathfrak{X}, P\}$ și un șir (f_n) de variabile aleatoare independente uniform repartizate

pe $[0, 1]$. Utilizând formula de transport avem:

$$\int_{\{x_1^2 + \dots + x_n^2 \leq \sqrt{n}\} \cap [0, 1]^n} dx_1 \dots dx_n = P\{f_1^2 + f_2^2 + \dots + f_n^2 \leq \sqrt{n}\} =$$

$$= P\left\{\frac{f_1^2 + \dots + f_n^2}{n} \leq \frac{1}{\sqrt{n}}\right\}.$$

Dar conform legii slabe a numerelor mari

$$P\left\{\left|\frac{f_1^2 + \dots + f_n^2}{n} - \frac{1}{3}\right| > \varepsilon\right\} \rightarrow 0 \text{ pentru orice } \varepsilon < 0.$$

Observăm apoi că pentru $\varepsilon = \frac{1}{4}$ (de exemplu) există un rang N astfel încît pentru orice $n > N$

$$\left\{\frac{f_1^2 + \dots + f_n^2}{n} \leq \frac{1}{\sqrt{n}}\right\} \subset \left\{\left|\frac{f_1^2 + \dots + f_n^2}{n} - \frac{1}{3}\right| > \varepsilon\right\}.$$

Rezultă deci că $\lim_{n \rightarrow \infty} \int_{\{x_1^2 + \dots + x_n^2 \leq \sqrt{n}\} \cap [0, 1]^n} dx_1 \dots dx_n = 0$.

b) Se raționează analog luînd pentru ε o valoare mai mică decît $\frac{1}{12}$, obținîndu-se pentru limită valoarea 0.

c) De această dată luăm $\varepsilon = \frac{1}{6}$ (de exemplu)

$$\left\{\left|\frac{f_1^2 + \dots + f_n^2}{n} - \frac{1}{3}\right| \leq \varepsilon\right\} \subset \left\{\frac{f_1^2 + \dots + f_n^2}{n} \leq \frac{1}{2}\right\}$$

obținîndu-se pentru limită valoarea 0.

3.32. Fie $f \in C^2([0, 1])$. Să se arate că este adevărată egalitatea:

$$\lim_{n \rightarrow \infty} n \int_{[0, 1]^n} \left[f\left(\frac{x_1 + \dots + x_n}{n}\right) - f\left(\frac{1}{2}\right) \right] dx = \frac{f''\left(\frac{1}{2}\right)}{24}.$$

Soluție: Fie u_1, \dots, u_n, \dots un șir de variabile aleatoare independente cu repartiția uniformă pe $[0, 1]$.

$$\text{Atunci } \int_{[0, 1]^n} \left[f\left(\frac{x_1 + \dots + x_n}{n}\right) - f\left(\frac{1}{2}\right) \right] dx = M \left[f\left(\frac{u_1 + \dots + u_n}{n}\right) - f\left(\frac{1}{2}\right) \right],$$

aşa că avem de dovedit egalitatea

$$\lim_{n \rightarrow \infty} nM \left[f\left(\frac{u_1 + \dots + u_n}{n}\right) - f\left(\frac{1}{2}\right) \right] = \frac{f''\left(\frac{1}{2}\right)}{24}.$$

Din formula lui Taylor obținem :

$$f(x) = f\left(\frac{1}{2}\right) + f'\left(\frac{1}{2}\right)\left(x - \frac{1}{2}\right) + \frac{1}{2}f''\left(\frac{1}{2} + \theta\left(x - \frac{1}{2}\right)\right)\left(x - \frac{1}{2}\right)^2,$$

$$0 < \theta < 1.$$

și cum $M\left(\frac{u_1 + \dots + u_n}{n} - \frac{1}{2}\right) = 0$ rezultă egalitatea

$$\begin{aligned} & M \left[f\left(\frac{u_1 + \dots + u_n}{n}\right) - f\left(\frac{1}{2}\right) \right] = \\ & = \frac{1}{2} M \left[f''\left(\frac{1}{2} + \theta\left(\frac{u_1 + \dots + u_n}{n} - \frac{1}{2}\right)\right) \left(\frac{u_1 + \dots + u_n}{n} - \frac{1}{2}\right)^2 \right]. \end{aligned}$$

Din legea tare a numerelor avem că $\frac{u_1 + \dots + u_n}{n} \xrightarrow{a.s.} \frac{1}{2}$ și cum

$$nM \left[\left(\frac{u_1 + u_2 + \dots + u_n}{n} - \frac{1}{2}\right)^2 \right] = \frac{1}{12}$$

rezultă că

$$\begin{aligned} & nM \left[f\left(\frac{u_1 + \dots + u_n}{n}\right) - f\left(\frac{1}{2}\right) \right] - \frac{f''\left(\frac{1}{2}\right)}{24} = \\ & = \frac{n}{2} M \left[f''\left(\frac{1}{2} + \theta\left(\frac{u_1 + \dots + u_n}{n} - \frac{1}{2}\right)\right) \cdot \left(\frac{u_1 + \dots + u_n}{n} - \frac{1}{2}\right)^2 \right] - \\ & \quad - \frac{nf''\left(\frac{1}{2}\right)}{2} M \left[\left(\frac{u_1 + \dots + u_n}{n} - \frac{1}{2}\right)^2 \right]. \end{aligned}$$

3.33. Fie șirul $(f_n)_{n \geq 1}$ de variabile aleatoare independente care iau cu probabilități egale doar două valori $\pm nx_n$, $\alpha_n > \alpha > 0$, $n = 1, 2, \dots$

Să se arate că șirul $(f_n)_{n \geq 1}$ nu satisface legea slabă a numerelor mari.

Soluție. Evident $M(f_n) = 0$. Avem apoi

$$\begin{aligned} P\left\{\frac{1}{n}\left|\sum_{k=1}^n f_k\right| < \varepsilon\right\} &= P\left\{\frac{1}{n}\left|\sum_{k=1}^{n-1} f_k + f_n\right| < \varepsilon\right\} = \\ &= \frac{1}{2} P\left\{\frac{1}{n}\left|\sum_{k=1}^{n-1} f_k + n\alpha_n\right| < \varepsilon\right\} + \frac{1}{2} P\left\{\frac{1}{n}\left|\sum_{k=1}^{n-1} f_k - n\alpha_n\right| < \varepsilon\right\}. \end{aligned}$$

Fie $\varepsilon > 0$ astfel încît de la un rang încolo

$$\varepsilon < \alpha_n \text{ și } \frac{1}{n}\left|\sum_{k=1}^{n-1} f_k + n\alpha_n\right| < \varepsilon.$$

Atunci

$$\begin{aligned} \frac{1}{n}\left|\sum_{k=1}^{n-1} f_k - n\alpha_n\right| &= \left|2\alpha_n - \frac{1}{n}\left(\sum_{k=1}^{n-1} f_k + n\alpha_n\right)\right| \geq \\ &\geq 2\alpha_n - \frac{1}{n}\left|\sum_{k=1}^{n-1} f_k + n\alpha_n\right| > \varepsilon. \end{aligned}$$

Rezultă deci că pentru un astfel de ε

$$P\left\{\frac{1}{n}\left|\sum_{k=1}^n f_k\right| < \varepsilon\right\} \leq \frac{1}{2}.$$

3.34. Fie $(f_n)_{n \geq 1}$ un șir de variabile aleatoare independente care iau cu probabilități egale două valori $\pm \alpha_n$ care satisfac relația $\alpha_n \geq \sum_{k=1}^{n-1} \alpha_k + \beta$, $\beta > 0$. Să se arate că șirul (f_n) nu satisface legea slabă a numerelor mari.

Soluție. Evident $M(f_n) = 0$. Apoi

$$\left|\sum_{k=1}^{n-1} f_k\right| \leq \sum_{k=1}^{n-1} |f_k| = \sum_{k=1}^{n-1} \alpha_k \leq \alpha_n - \beta \leq \alpha_n = |f_n|.$$

Atunci

$$\begin{aligned} P\left\{\frac{1}{n}\left|\sum_{k=1}^n f_k\right| < \varepsilon\right\} &= P\left\{\left|\frac{1}{n}f_n + \frac{1}{n}\sum_{k=1}^{n-1} f_k\right| < \varepsilon\right\} \leq \\ &\leq P\left\{\frac{1}{n}|f_n| - \frac{1}{n}\left|\sum_{k=1}^{n-1} f_k\right| < \varepsilon\right\} = P\left\{\frac{\alpha_n}{n} - \frac{1}{n}\left|\sum_{k=1}^{n-1} f_k\right| < \varepsilon\right\}. \end{aligned}$$

Dar $\frac{1}{n} \left| \sum_{k=1}^{n-1} f_k \right| \leq \frac{1}{n} \sum_{k=1}^{n-1} |f_k| = \frac{1}{n} \sum_{k=1}^{n-1} \alpha_k$, deci

$$\frac{\alpha_n}{n} - \frac{1}{n} \left| \sum_{k=1}^{n-1} f_k \right| \geq \beta.$$

Rezultă că pentru $\varepsilon < \beta$ obținem

$$P \left\{ \frac{1}{n} \left| \sum_{k=1}^n f_k \right| < \varepsilon \right\} = 0.$$

3.35. Fie (f_n) un șir de variabile aleatoare independente cu densitățile $p_n(x)$ triunghiulare, adică

$$p_n(x) = \begin{cases} 0 & \text{dacă } x \in (-\infty, -n^\delta] \cup (n^\delta, \infty) \\ \frac{n^\delta + x}{n^{2\delta}} & \text{dacă } x \in (-n^\delta, 0] \\ \frac{n^\delta - x}{n^{2\delta}} & \text{dacă } x \in (0, n^\delta] \end{cases} \quad \delta < \frac{1}{2}$$

Să se arate că (f_n) satisface legea slabă a numerelor mari.

Soluție. Avem succesiv

$$M(f_n) = \int_{-n^\delta}^0 x \frac{n^\delta + x}{n^{2\delta}} dx + \int_0^{n^\delta} x \frac{n^\delta - x}{n^{2\delta}} dx = 0.$$

$$D^2(f_n) = \int_{-n^\delta}^0 x^2 \frac{n^\delta + x}{n^{2\delta}} dx + \int_0^{n^\delta} x^2 \frac{n^\delta - x}{n^{2\delta}} dx = \frac{1}{6} n^{2\delta}.$$

În continuare:

$$\frac{1}{n^2} D^2 \left(\sum_{k=1}^n f_k \right) = \frac{1}{n^2} \sum_{k=1}^n \frac{k^{2\delta}}{6} < \frac{1}{6n^2} \int_1^{n+1} x^{2\delta} dx < \frac{(n+1)^{2\delta+1}}{6(2\delta+1)n^2} \rightarrow 0$$

deoarece cum $\delta < \frac{1}{2}$ avem și $2\delta + 1 < 2$.

Prin urmare (f_n) satisface legea slabă a numerelor mari, vezi condiția (4) din rezumatul teoretic.

3.36. Fie f_1, \dots, f_n, \dots un șir de variabile aleatoare independente, identic repartizate și g_1, \dots, g_n, \dots un alt șir de variabile aleatoare independente, identic repartizate așa încît

$$M(f_n) = M(g_n) = a, \quad D^2(f_n) = D^2(g_n) = \sigma^2 < \infty.$$

Să se arate că șirul $h_{2n} = f_n, h_{2n-1} = g_n$ satisface legea slabă a numerelor mari.

Soluție. Afirmația rezultă dacă vom arăta că are loc incluziunea

$$\left\{ \left| \frac{h_1 + \dots + h_n}{n} - a \right| > \varepsilon \right\} \subset \left\{ \left| \frac{f_1 + \dots + f_{\lfloor \frac{n}{2} \rfloor}}{\lfloor \frac{n}{2} \rfloor} - a \right| > \varepsilon \right\} \cup \left\{ \left| \frac{g_1 + \dots + g_{\lfloor \frac{n+1}{2} \rfloor}}{\lfloor \frac{n+1}{2} \rfloor} - a \right| > \frac{\varepsilon}{2} \right\}$$

(șirurile (f_n) și (g_n) satisfac legea slabă (chiar tare) a numerelor mari). Presupunem pentru început că $n = 2k$. Avem

$$\left| \frac{h_1 + \dots + h_n}{n} - a \right| > \varepsilon \Leftrightarrow \left| \frac{f_1 + \dots + f_k + g_1 + \dots + g_k}{2k} - a \right| > \varepsilon \Leftrightarrow \left| \frac{f_1 + \dots + f_k + g_1 + \dots + g_k}{k} - 2a \right| > 2\varepsilon \Rightarrow \left| \frac{f_1 + \dots + f_k}{k} - a \right| > \varepsilon$$

$$\text{sau } \left| \frac{g_1 + \dots + g_k}{k} - a \right| > \varepsilon \Leftrightarrow \left| \frac{f_1 + \dots + f_{\lfloor \frac{n}{2} \rfloor}}{\lfloor \frac{n}{2} \rfloor} - a \right| > \varepsilon \text{ sau}$$

$$\left| \frac{g_1 + \dots + g_{\lfloor \frac{n+1}{2} \rfloor}}{\lfloor \frac{n+1}{2} \rfloor} - a \right| > \varepsilon.$$

Presupunem acum că $n = 2k + 1$. Avem

$$\left| \frac{h_1 + \dots + h_n}{n} - a \right| > \varepsilon \Leftrightarrow \left| \frac{f_1 + \dots + f_k - ka + g_1 + \dots + g_{k+1} - (k+1)a}{2k+1} \right| > \varepsilon \Rightarrow$$

$$\begin{aligned}
&\Rightarrow \left| \frac{f_1 + \dots + f_k - ka}{2k+1} \right| > \frac{\varepsilon}{2} \text{ sau } \left| \frac{g_1 + \dots + g_{k+1} - (k+1)a}{2k+1} \right| > \frac{\varepsilon}{2} \Rightarrow \\
&\Rightarrow \left| \frac{f_1 + \dots + f_k - ka}{2k} \right| > \frac{\varepsilon}{2} \text{ sau } \frac{k+1}{2k+1} \left| \frac{g_1 + \dots + g_{k+1} - (k-1)a}{k+1} \right| > \frac{\varepsilon}{2} \Rightarrow \\
&\Rightarrow \left| \frac{f_1 + \dots + f_k}{k} - a \right| > \varepsilon \text{ sau } \left| \frac{g_1 + \dots + g_{k+1}}{k+1} - a \right| > \frac{\varepsilon}{2} \Leftrightarrow \\
&\Leftrightarrow \left| \frac{f_1 + \dots + f_{\left[\frac{n}{2} \right]}}{\left[\frac{n}{2} \right]} - a \right| > \varepsilon \text{ sau } \left| \frac{g_1 + \dots + g_{\left[\frac{n+1}{2} \right]}}{\left[\frac{n+1}{2} \right]} - a \right| > \frac{\varepsilon}{2}.
\end{aligned}$$

3.37. Variabilele aleatoare f_1, \dots, f_n, \dots sînt independente și $M(f_n) = a$, $D^2(f_n) = \sigma^2 < \infty$ pentru orice n . Dacă $h_n = \sum_{1 \leq i < j \leq n} f_i f_j$ să se arate că $\frac{h_n}{C_n^2} \xrightarrow{p} a^2$. Dacă în plus variabilele aleatoare f_1, \dots, f_n, \dots sînt identic repartizate să se arate că $\frac{h_n}{C_n^2} \xrightarrow{a.s.} a^2$.

Soluție. Avem $M(h_n) = \sum_{1 \leq i < j \leq n} M(f_i f_j) = \sum_{1 \leq i < j \leq n} M(f_i) M(f_j) = C_n^2 a^2$, deci pentru a arăta că $\frac{h_n}{C_n^2} \xrightarrow{p} a^2$ este suficient să demonstrăm că $D^2\left(\frac{h_n}{C_n^2}\right) \rightarrow 0$ (vezi afirmația din punctul e) al rezumatului teoretic).

Dar $D^2\left(\frac{h_n}{C_n^2}\right) = \frac{1}{(C_n^2)^2} D^2(h_n)$, așa că este suficient să arătăm că $D^2(h_n)$ este un polinom în n de grad mai mic sau egal cu 3. Cum $M^2(h_n) = \frac{a^4}{4} n^4 + P_3(n)$ ($P_3(n)$ polinom de gradul 3 în n) rezultă că totul se reduce la a arăta că $M(h_n^2) = \frac{a^4}{4} n^4 + \bar{P}_3(n)$ cu $\bar{P}_3(n)$ polinom de grad 3 în n (atunci $D^2(h_n) = M(h_n^2) - M^2(h_n) = P_3(n) - \bar{P}_3(n)$). Pentru $j = 1, \dots, n-1$ fie $g_j = \sum_{l=j+1}^n f_l$. Atunci $h_n = \sum_{j=1}^{n-1} f_j g_j$.

Avem

$$\begin{aligned}
M(g_j^2) &= \sum_{l=j+1}^n M(f_l^2) + 2 \sum_{j+1 \leq l < k \leq n} M(f_l) M(f_k) = \\
&= (n-j)(\sigma^2 + a^2) + 2a^2 C_{n-j}^2
\end{aligned}$$

$$h_n^2 = \sum_{j=1}^{n-1} f_j^2 g_j^2 + 2 \sum_{j=1}^{n-2} f_j f_{j+1} g_j g_{j+1} + 2 \sum_{j=1}^{n-3} \sum_{j=1}^{n-j-1} f_j f_{j+1} g_j g_{j+1} = \\ = I_1 + 2I_2 + 2I_3.$$

$$M(I_1) = \sum_{j=1}^{n-1} M(f_j^2) M(g_j^2) = (\sigma^2 + a^2) \sum_{j=1}^{n-1} [(n-j)(\sigma^2 + a^2) + 2a^2 C_{n-j}^2] = \\ = (\sigma^2 + a^2) \sum_{j=1}^{n-1} [(n-j) + a^2(\sigma^2 + a^2)(n-j)(n-j-1)] = (\sigma^2 + a^2) \cdot \\ \cdot \frac{n(n-1)}{2} + a^2(\sigma^2 + a^2) \left[\sum_{j=1}^{n-1} (n-j)^2 - \sum_{j=1}^{n-1} (n-j) \right] = (\sigma^2 + a^2)^2 \frac{n(n-1)}{2} + \\ + a^2(\sigma^2 + a^2) \left[\frac{(n-1)n(2n-1)}{6} - \frac{n(n-1)}{2} \right] = Q_3(n) \text{ cu } Q_3(n) \text{ polinom de} \\ \text{grad 3 în } n \text{ (reșinem din acest calcul că } \sum_{j=1}^{n-1} M(g_j^2) \text{ este polinom de} \\ \text{grad 3 în } n). \text{ Mai departe}$$

$$f_j f_{j+1} g_j g_{j+1} = f_j f_{j+1}^2 (f_{j+2} + \dots + f_n) + f_j f_{j+1} (f_{j+2} + \dots + f_n)^2 \\ M(f_j f_{j+1} g_j g_{j+1}) = M(f_j) M(f_{j+1}^2) [M(f_{j+2}) + \dots + M(f_n)] + \\ + M(f_j) M(f_{j+1}) M[(f_{j+2} + \dots + f_n)^2] = a^2(\sigma^2 + a^2)(n-j-1) + \\ + a^2 [(n-j-1)(\sigma^2 + a^2) + 2a^2 C_{n-j-1}^2]$$

$$M(I_2) = \sum_{j=1}^{n-2} M(f_j f_{j+1} g_j g_{j+1}) = 2a^2(\sigma^2 + a^2) \sum_{j=1}^{n-2} (n-j-1) + \\ + a^4 \sum_{j=1}^{n-2} (n-j-1)(n-j-2) = \bar{Q}_3(n) \text{ cu } \bar{Q}_3(n) \text{ polinom de grad 3.}$$

Prin urmare, rămâne de arătat că $M(2I_3) = \frac{a^4}{4} n^4 + \tilde{Q}_3(n)$ cu $\tilde{Q}_3(n)$ polinom de gradul 3.

Pentru $l > 1$ avem

$$f_j f_{j+l} g_j g_{j+l} = f_j f_{j+l} (f_{j+1} + \dots + f_{j+l-1})(f_{j+l+1} + \dots + f_n) + \\ + f_j f_{j+l}^2 (f_{j+l+1} + \dots + f_n) + f_j f_{j+l} (f_{j+l+1} + \dots + f_n)^2.$$

$$\begin{aligned}
M(f_j f_{j+1} g_j g_{j+1}) &= a^4 (l-1)(n-j-l) + a^2(\sigma^2 + a^2)(n-j-l) + \\
&\quad + a^2 [(\sigma^2 + a^2)(n-j-l) + 2a^2 C_{n-j-1}^2] \\
M(I_3) &= \sum_{j=1}^{n-3} \sum_{l=2}^{n-j-1} M(f_j f_{j+1} g_j g_{j+1}) = \sum_{j=1}^{n-3} \sum_{l=2}^{n-j-1} \{a^4(l-1)(n-j-l) + \\
&\quad + a^2(\sigma^2 + a^2)(n-j-l) + a^2 [(\sigma^2 + a^2)(n-j-l) + \\
&\quad + a^2(n-j-l)(n-j-l-1)]\} = \sum_{j=1}^{n-3} \left\{ a^4(n-j) \sum_{l=2}^{n-j-1} (l-1) - \right. \\
&\quad - a^4 \sum_{l=2}^{n-j-1} l^2 + a^4 \sum_{l=2}^{n-j-1} l + 2a^2(\sigma^2 + a^2) \sum_{l=2}^{n-j-1} (n-j-l) + \\
&\quad \left. + a^4 \sum_{l=2}^{n-j-1} (n-j-l)^2 - a^4 \sum_{l=2}^{n-j-1} (n-j-l) \right\} = \\
&= \sum_{j=1}^{n-3} \left\{ a^4(n-j) \frac{(n-j-2)(n-j-1)}{2} - a^4 \left[\frac{(n-j-1)(n-j)(2n-2j-1)}{6} - 1 \right] + \right. \\
&\quad + a^4 \left[\frac{(n-j-1)(n-j)}{2} - 1 \right] + 2a^2(\sigma^2 + a^2) \frac{(n-j-2)(n-j-1)}{2} + \\
&\quad \left. + a^4 \frac{(n-j-2)(n-j-1)(2n-2j-3)}{6} - a^4 \frac{(n-j-2)(n-j-1)}{2} \right\} = \\
&= \sum_{j=1}^{n-3} \left\{ \frac{a^4}{2} [(n-j) - 3(n-j)^2 + 2(n-j)] - \frac{a^4}{6} [2(n-j)^3 - \right. \\
&\quad - 3(n-j)^2 + (n-j)] + \frac{a^4}{6} [2(n-j)^3 - 9(n-j)^2 + 13(n-j) - 6] \Big\} + \\
&\quad + \frac{a^4}{2} (n-j-1)(n-j) + a^2(\sigma^2 + a^2)(n-j-2)(n-j-1) - \\
&\quad - \frac{a^4}{2} (n-j-2)(n-j-1) = \frac{a^4}{2} \sum_{j=1}^{n-3} (n-j)^3 + H_3(n) \text{ cu } H_3(n) \text{ poli-} \\
&\text{nom de grad 3. Dar}
\end{aligned}$$

$$\sum_{j=1}^{n-3} (n-j)^3 = \frac{(n-1)^2 n^2}{4} - 1^3 - 2^3$$

deci

$$M(2I_3) = 2M(I_3) = 2 \left[\frac{a^4}{4} \sum_{j=1}^{n-3} (n-j)^3 + H_3(n) \right] = a^4 \sum_{j=1}^{n-3} (n-j)^3 + 2H_3(n) = \frac{a^4}{4} n^4 + \tilde{Q}_3(n).$$

Presupunem acum că f_1, \dots, f_n, \dots sînt identic repartizate. Avem că $(f_1 + \dots + f_n)^2 = f_1^2 + \dots + f_n^2 + 2h_n$, deci

$$\frac{h_n}{C_n^2} = \frac{(f_1 + \dots + f_n)^2}{2C_n^2} - \frac{f_1^2 + \dots + f_n^2}{2C_n^2}$$

Din legea tare a numerelor mari rezultă că $\frac{f_1 + \dots + f_n}{n} \xrightarrow{a.s.} a$ și $\frac{f_1^2 + \dots + f_n^2}{n} \xrightarrow{a.s.} \sigma^2 + a^2$, deci

$$\begin{aligned} \frac{(f_1 + \dots + f_n)^2}{2C_n^2} &= \frac{(f_1 + \dots + f_n)}{n^2} \frac{n}{n-1} \xrightarrow{a.s.} a^2 \text{ și } \frac{f_1^2 + \dots + f_n^2}{2C_n^2} = \\ &= \frac{1}{n+1} \frac{f_1^2 + \dots + f_n^2}{n} \xrightarrow{a.s.} 0, \text{ fapt ce implică convergența } \frac{h_n}{C_n^2} \xrightarrow{a.s.} a^2. \end{aligned}$$

3.38. (O teoremă a lui Bernstein). Într-un câmp de probabilitate (E, \mathcal{X}, P) fie (f_n) un șir de variabile aleatoare astfel încît $D^2 f_n \leq C$ și $r_{ij} \rightarrow \infty$ odată cu $|i-j| \rightarrow \infty$, (unde r_{ij} este coeficientul de corelație dintre f_i și f_j).

Să se arate că (f_n) satisface legea slabă a numerelor mari.

$$\text{Soluție. Știind că } r_{ij} = \frac{M[(f_i - M(f_i))(f_j - M(f_j))]}{D(f_i)D(f_j)}$$

avem

$$D^2 \left(\sum_{k=1}^n f_k \right) = \sum_{k=1}^n D^2(f_k) + 2 \sum_{1 \leq i < j \leq n} D(f_i) D(f_j) r_{ij}$$

Fie $\varepsilon > 0$ arbitrar. Există N natural încît dacă $|i-j| > N$, $|r_{ij}| < \varepsilon$. Atunci

$$D^2 \left(\sum_{k=1}^n f_k \right) = \sum_{k=1}^n D^2(f_k) + 2 \sum_{\substack{1 \leq i < j \leq n \\ |i-j| \leq N}} D(f_i) D(f_j) \frac{r_{ij}}{|r_{ij}| \leq 1} +$$

$$2 + \sum_{\substack{1 \leq i < j \leq n \\ |i-j| > N}} D(f_i) D(f_j) \frac{r_{ij}}{|r_{ij}| < \varepsilon} \leq$$

$$\leq nC + 4CnN + \varepsilon Cn(n+1).$$

Rezultă

$$\frac{1}{n^2} D^2 \left(\sum_{k=1}^n f_k \right) \leq \frac{1}{n} C_1 + \varepsilon \cdot C_2$$

și de aici

$$\frac{1}{n^2} D_2 \left(\sum_{k=1}^n f_k \right) \rightarrow 0.$$

Deci (f_n) satisface legea slabă a numerelor mari (vezi condiția (4) din rezumatul teoretic).

3.38. a) Fie (f_n) un șir de variabile aleatoare independente așa încît $\lim_{n \rightarrow \infty} \frac{1}{n} D^2(f_n) = 0$. Să se arate că (f_n) satisface legea slabă a numerelor mari.

b) Fie (f_n) un șir de variabile aleatoare așa încît pentru orice n, f_n este independentă de $\mathfrak{F}(f_l; l \neq n-1, n+1)$.

Să se arate că dacă $D^2(f_n)$ este șir crescător și $\lim_{n \rightarrow \infty} \frac{D^2(f_n)}{n} \rightarrow 0$ atunci (f_n) satisface legea slabă a numerelor m mari.

Soluție. a) Este suficient să arătăm că $\lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{k=1}^n D^2(f_k) = 0$ (vezi condiția 5) din rezumatul teoretic).

Ori conform lemei lui Stoltz avem

$$\lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{k=1}^n D^2(f_k) = \lim_{n \rightarrow \infty} \frac{D^2(f_n)}{n^2 - (n-1)^2} = \lim_{n \rightarrow \infty} \frac{D^2(f_n)}{2n-1} = 0$$

b) Avem:

$$\begin{aligned} D^2 \left(\sum_{k=1}^n f_n \right) &= \sum_{k=1}^n D^2(f_k) + 2 \sum_{k=1}^n \text{cov}(f_k, f_{k+1}) + 2 \sum_{\substack{h>k \\ h-k \geq 2}} \text{cov}(f_h, f_k) = \\ &= \sum_{k=1}^n D^2(f_k) + 2 \sum_{k=1}^n \text{cov}(f_k, f_{k+1}) \text{ deoarece din ipoteza de independență} \\ &\text{obținem } \text{cov}(f_k, f_h) = 0 \text{ dacă } h - k \geq 2. \end{aligned}$$

Dar din inegalitatea lui Schwartz cov $(f_k, f_{k+1}) \leq \sqrt{D^2(f_k)D^2(f_{k+1})} \leq D^2(f_{k+1})$ așa că

$$\frac{1}{n^2} D^2\left(\sum_{k=1}^n f_k\right) \leq \frac{3}{n^2} \sum_{k=1}^n D^2(f_k) \leq 3 \frac{D^2(f_n)}{n} \rightarrow 0,$$

cu alte cuvinte (f_n) satisface legea slabă a numerelor mari (vezi condiția 4) din rezumatul teoretic).

3.40. Fie f_1, \dots, f_n, \dots un șir de variabile aleatoare independente cu media zero și așa încît $\lim_{n \rightarrow \infty} \frac{1}{n} M(f_n^2) = 0$ și fie τ o variabilă aleatoare cu repartiția Poisson de parametru λ și independentă de șirul (f_n) .

Să se deducă convergența $\frac{1}{\tau+1} \sum_{k=1}^{\tau+1} f_k \xrightarrow[\lambda \rightarrow \infty]{o} 0$.

Soluție. Fie $\varepsilon > 0$ și alegem δ, N așa încît $P\left(\frac{1}{n+1} |f_1 + \dots + f_{n+1}| \geq \varepsilon\right) \leq \delta$ dacă $n > N$, (f_n satisface legea slabă a numerelor mari, vezi problema precedentă).

Atunci avem:

$$\begin{aligned} P\left(\frac{1}{\tau+1} \left|\sum_{k=1}^{\tau+1} f_k\right| \geq \varepsilon\right) &= \sum_{n=0}^{\infty} P(\tau = n) P\left(\frac{1}{n+1} \left|\sum_{k=1}^{n+1} f_k\right| \geq \varepsilon \mid \tau = n\right) = \\ &= \sum_{n=0}^{\infty} P(\tau = n) P\left(\frac{1}{n+1} \left|\sum_{k=1}^{n+1} f_k\right| \geq \varepsilon\right) = \sum_{n=0}^N \frac{\lambda^n}{n!} e^{-\lambda} P\left(\frac{1}{n+1} \left|\sum_{k=1}^{n+1} f_k\right| \geq \varepsilon\right) + \\ &+ \sum_{n=N+1}^{\infty} P(\tau = n) P\left(\frac{1}{n+1} \left|\sum_{k=1}^{n+1} f_k\right| \geq \varepsilon\right) \leq \sum_{n=0}^N \frac{\lambda^n}{n!} e^{-\lambda} + \delta \end{aligned}$$

de unde făcînd pe rînd $\lambda \rightarrow \infty, \delta \rightarrow 0$ obținem

$$\lim_{n \rightarrow \infty} P\left(\frac{1}{\tau+1} \left|\sum_{k=1}^{\tau+1} f_k\right| \geq \varepsilon\right) = 0.$$

3.41. Fie (f_n) un șir de variabile aleatoare pentru care $M(f_n) = 0$ și seria $\sum_{n=1}^{\infty} \frac{1}{n} f_n$ este convergentă a.s. Să se arate că șirul (f_n) satisface legea tare a numerelor mari.

Soluție. Să notăm cu $s_n = \sum_{k=1}^n \frac{1}{k} f_k$. Din ipoteză $s_n \xrightarrow{a.s.} S =$
 $= \sum_{n=1}^{\infty} \frac{1}{n} f_n$.

Observăm apoi că

$$f_k = k(s_k - s_{k-1}), \quad k \geq 1, \text{ luând } s_0 = 0.$$

Rezultă deci

$$\begin{aligned} \frac{1}{n} \sum_{k=1}^n f_k &= \frac{1}{n} \sum_{k=1}^n k(s_k - s_{k-1}) = \frac{1}{n} [s_1 + 2(s_2 - s_1) + \dots + n(s_n - s_{n-1})] = \\ &= - \underbrace{\frac{1}{n} (s_1 + s_2 + \dots + s_n)}_{\xrightarrow{a.s.} s} + \underbrace{\frac{n+1}{n} s_n}_{\xrightarrow{a.s.} s} \xrightarrow{a.s.} 0. \end{aligned}$$

Se poate da următoarea generalizare :

Fie (f_n) un șir de variabile aleatoare independente pentru care $M(f_n) = 0$, $M(f_n^2) < \infty$, și a_n un șir de numere pozitive strict crescător astfel încît seria $\sum_{n=1}^{\infty} \frac{1}{a_n} f_n$ este convergentă a.s. Să se arate că

$$\frac{1}{a_n} \sum_{k=1}^n f_k \xrightarrow{a.s.} 0.$$

Pentru demonstrație se utilizează următorul rezultat : Dacă $x_n \rightarrow x$, atunci $\frac{1}{a_n} \sum_{k=1}^n (a_k - a_{k-1}) x_k \rightarrow x$, (a_n) fiind șirul din enunț.

3.42. Se dă șirul de variabile aleatoare independente (f_n) , unde f_n ia valorile $-n\alpha, 0, n\alpha$, $\alpha \in R$, cu probabilitățile $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$.

Folosind teorema limită centrală, să se arate că șirul (f_n) nu satisface legea slabă a numerelor mari.

Soluție. Avem $M(f_n) = 0$, $D^2(f_n) = \frac{n^2 \alpha^2}{2}$ și $\sigma_{(n)}^2 = \sum_{k=1}^n \sigma_k^2 =$
 $= \frac{\alpha^2 n(n+1)(2n+1)}{12}$, deci pentru $\varepsilon > 0$ există n_ε așa încît $\varepsilon \sigma_{(n)} > n|\alpha|$

pentru $n \geq n_\varepsilon$. Este astfel satisfăcută condiția Lindeberg (vezi (7) din rezumatul teoretic).

$$P\left(\left|\frac{1}{n} \sum_{k=1}^n f_k\right| < \varepsilon\right) \leq P\left(\frac{1}{\frac{\varepsilon}{|\alpha|} \sigma_{(n)}} \left|\sum_{k=1}^n f_k\right| < \varepsilon\right) =$$

$$= P\left(\frac{1}{\sigma_{(n)}} \left|\sum_{k=1}^n f_k\right| < \frac{\varepsilon^2}{|\alpha|}\right) \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\frac{\varepsilon^2}{|\alpha|}}^{\frac{\varepsilon^2}{|\alpha|}} e^{-\frac{t^2}{2}} dt < 1,$$

asa că șirul (f_n) nu satisface legea slabă a numerelor mari.

3.43. Se dă șirul (f_n) de variabile aleatoare independente fiecare f_n avînd densitatea de repartiție

$$p_n(x) = \begin{cases} 2^n \text{ dacă } -\frac{1}{2^{n+2}} \leq x \leq \frac{1}{2^{n+2}} \text{ și } 1 - \frac{1}{2^{n+3}} \leq |x| \leq 1 + \frac{1}{2^{n+3}}, \\ 0 \text{ în rest.} \end{cases}$$

Să se arate că șirul (f_n) satisface teorema limită centrală.

Soluție.

$$\begin{array}{ccccccc} -1 - \frac{1}{2^{n+3}} & -1 & -1 + \frac{1}{2^{n+3}} & -\frac{1}{2^{n+2}} & 0 & \frac{1}{2^{n+2}} & 1 - \frac{1}{2^{n+3}} & 1 + \frac{1}{2^{n+3}} \\ \hline & \underbrace{\hspace{1.5cm}} & & \underbrace{\hspace{1.5cm}} & & & \underbrace{\hspace{1.5cm}} & \end{array}$$

Avem

$$M(f_n) = 2^n \left(\int_{-1 - \frac{1}{2^{n+3}}}^{-1 + \frac{1}{2^{n+3}}} x dx + \int_{-\frac{1}{2^{n+2}}}^{\frac{1}{2^{n+2}}} x dx + \int_{1 - \frac{1}{2^{n+3}}}^{1 + \frac{1}{2^{n+3}}} x dx \right) = 0.$$

$$D^2(f_n) = 2^n \left(\int_{-1 - \frac{1}{2^{n+3}}}^{-1 + \frac{1}{2^{n+3}}} x^2 dx + \int_{-\frac{1}{2^{n+2}}}^{\frac{1}{2^{n+2}}} x^2 dx + \int_{1 - \frac{1}{2^{n+3}}}^{1 + \frac{1}{2^{n+3}}} x^2 dx \right) =$$

$$= 2^n \left[\frac{2}{3} \left[\left(1 + \frac{1}{2^{n+3}}\right)^3 - \left(1 - \frac{1}{2^{n+3}}\right)^3 \right] + \frac{2}{3} \left(\frac{1}{2^{n+2}}\right)^3 \right] = \frac{1}{2} + \frac{5}{3 \cdot 2^{2n+7}}$$

$$\begin{aligned}
M(|f_n|^3) &= - \int_{-\infty}^0 x^3 \rho_n(x) dx + \int_0^{\infty} x^3 \rho_n(x) dx = \\
&= 2^n \left(- \int_{-1-\frac{1}{2^{n+3}}}^{-1+\frac{1}{2^{n+3}}} x^3 dx - \int_{-\frac{1}{2^{n+2}}}^0 x^3 dx + \int_0^{\frac{1}{2^{n+2}}} x^3 dx + \int_{1-\frac{1}{2^{n+3}}}^{1+\frac{1}{2^{n+3}}} x^3 dx \right) = \\
&= 2^{n-1} \left[\left(1 + \frac{1}{2^{n+3}}\right)^4 - \left(1 - \frac{1}{2^{n+3}}\right)^4 + \left(\frac{1}{2^{n+2}}\right)^4 \right] = \frac{1}{2^3} + \frac{1}{2^{2n+8}} + \frac{1}{2^{3n+9}}.
\end{aligned}$$

Aceste calcule permit să arătăm că este satisfăcută teorema Lyapunov pentru $\delta = 1$ deci șirul (f_n) satisface teorema limită centrală.

3.44. Fie (f_n) un șir de variabile aleatoare independente, identic repartizate cu $\mu = M(f_1)$, $\tau^2 = D^2(f_1) < \infty$. Folosind teorema limită centrală să se arate că șirul $g_n = \sum_{k=1}^n f_k$ nu satisface legea slabă a numerelor mari.

Soluție. Să notăm cu $h_n = n f_n$. Observăm că $\sum_{k=1}^n g_k = \sum_{k=1}^n h_k$.

Deoarece $M(h_n) = n\mu$ și $D^2\left(\sum_{k=1}^n h_k\right) = \frac{n(n+1)(2n+1)}{6} \sigma^2$,

utilizând condiția Lindeberg, rezultă că șirul h_n satisface teorema limită centrală. Deci:

$$\begin{aligned}
&P \left\{ \frac{1}{D(h_1 + \dots + h_n)} \sum_{k=1}^n (h_k - M(h_k)) < \varepsilon \right\} = \\
&= P \left\{ \frac{1}{\sigma \sqrt{\frac{n(n+1)(2n+1)}{6}}} \sum_{k=1}^n (h_k - M(h_k)) < \varepsilon \right\} \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\varepsilon} e^{-\frac{t^2}{2}} dt.
\end{aligned}$$

Deci pentru orice $\varepsilon > 0$

$$\begin{aligned}
&P \left\{ \left| \frac{1}{n} \sum_{k=1}^n (g_k - M(g_k)) \right| \geq \varepsilon \right\} = P \left\{ \left| \frac{1}{n} \sum_{k=1}^n (h_k - M(h_k)) \right| \geq \varepsilon \right\} = \\
&= 1 - P \left\{ \frac{1}{n} \sum_{k=1}^n (h_k - M(h_k)) < \varepsilon \right\} + P \left\{ \frac{1}{n} \sum_{k=1}^n (h_k - M(h_k)) < -\varepsilon \right\} \geq
\end{aligned}$$

$$\begin{aligned} &\geq 1 - P \left\{ \frac{1}{\sqrt{\frac{\sigma(n+1)(2n+1)}{6}}} \sum_{k=1}^n (h_k - M(h_k)) < \varepsilon \right\} + \\ &+ P \left\{ \frac{1}{\sqrt{\frac{n(n+1)(2n+1)}{6}}} \sum_{k=1}^n (h_k - M(h_k)) < -\varepsilon \right\} \rightarrow \\ &1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\varepsilon} e^{-\frac{t^2}{2}} dt + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\varepsilon} e^{-\frac{t^2}{2}} dt = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\varepsilon}^{\varepsilon} e^{-\frac{t^2}{2}} dt > \frac{1}{2} \end{aligned}$$

pentru un ε suficient de mic, ceea ce arată că șirul g_n nu satisface legea slabă a numerelor mari.

3.45. Fie șirul de variabile aleatoare (f_n) așa încît :

a) $\max_{1 \leq k \leq n} |f_k| \leq \varepsilon_n$ unde ε_n este un șir convergent către zero ;

b) $D^2(f_1 + \dots + f_n) \geq C^2 > 0$.

Să se arate că (f_n) satisface teorema limită centrală.

Soluție. Vom arăta că este îndeplinită condiția Lindeberg (7) din rezumatul teoretic.

Pentru $1 \leq k \leq n$ avem că $|M_i(f_k)| \leq M(|f_k|) \leq \varepsilon_n$, deci $\max_{1 \leq k \leq n} |f_k - M(f_k)| \leq 2\varepsilon_n$ și prin urmare

$$P\left(\max_{1 \leq k \leq n} \frac{|f_k - M(f_k)|}{D(f_1 + \dots + f_n)} \leq \frac{2\varepsilon_n}{D(f_1 + \dots + f_n)}\right) = 1$$

Dar cum $\frac{2\varepsilon_n}{D(f_1 + \dots + f_n)} \leq \frac{2\varepsilon_n}{C}$, rezultă

$$P\left(\max_{1 \leq k \leq n} \frac{|f_k - M(f_k)|}{D(f_1 + \dots + f_n)} \leq \frac{2\varepsilon_n}{C}\right) = 1$$

Pentru $\varepsilon > 0$ fie n așa încît $\frac{2\varepsilon_n}{C} < \varepsilon$ dacă $n \geq n_\varepsilon$. Rezultă că avem

$P\left(\max_{1 \leq k \leq n} \frac{|f_k - M(f_k)|}{D(f_1 + \dots + f_n)} \geq \varepsilon\right) = 0$ dacă $n \geq n_\varepsilon$, fapt ce implică valabilitatea condiției Lindeberg.

3.46. Fie (f_n) un șir de variabile aleatoare independente încît $P\left(f_n = \pm \frac{1}{n^\beta}\right) = p$, $P(f_n = 0) = 1 - 2p$, unde $0 < p < 1$ și $\frac{1}{3} < \beta \leq \frac{1}{2}$.

Să se arate că șirul (f_n) satisface teorema limită centrală.

Soluție. Vom arăta că este îndeplinită condiția 9) din teorema lui Leapunov cu $\delta = 1$. Avem

$$M(f_n) = 0, D^2(f_1 + \dots + f_n) = \sum_{k=1}^n D^2(f_k) = 2p \sum_{k=1}^n \frac{1}{k^{2\beta}}$$

$$M(|f_n|) = 2p \frac{1}{n^{3\beta}}$$

deci

$$\frac{1}{[(D(f_1 + \dots + f_n))]^3} \sum_{k=1}^n M(|f_k|^3) = \frac{2p \sum_{k=1}^n \frac{1}{k^{3\beta}}}{(\sqrt{2p})^3 \left(\sum_{k=1}^n \frac{1}{k^{2\beta}} \right)^{\frac{3}{2}}} \rightarrow 0$$

deoarece $\sum_{k=1}^{\infty} \frac{1}{k^{3\beta}} < \infty$ și $\sum_{k=1}^{\infty} \frac{1}{k^{2\beta}} = \infty$ pentru $\frac{1}{3} < \beta \leq \frac{1}{2}$.

3.47. Fie f_1, \dots, f_n, \dots un șir de variabile aleatoare independente așa încît

a) $\sup |f_n| \leq C$ a.s unde C este o constantă ;

b) $\lim_{n \rightarrow \infty} D^2(f_1 + \dots + f_n) = \infty$.

Să se arate că șirul (f_n) satisface teorema limită centrală.

Soluție. Vom arăta că este îndeplinită condiția Leapunov 9) din rezumatul teoretic pentru orice $\delta > 0$. Avem

$$|M(f_k)| \leq M(|f_k|) \leq C, \text{ deci } |f_k - M(f_k)| \leq 2C \text{ a.s.}$$

Dacă $\delta > 0$, atunci

$$\frac{1}{[D(f_1 + \dots + f_n)]^{2+\delta}} \sum_{k=1}^n M(|f_k - M(f_k)|^{2+\delta}) \leq$$

$$\leq \frac{2^\delta C^\delta}{[D(f_1 + \dots + f_n)]^{2+\delta}} \sum_{k=1}^n M(|f_k - M(f_k)|^2) \leq \left[\frac{2C}{D^2(f_1 + \dots + f_n)} \right]^\delta \xrightarrow{n \rightarrow \infty} 0,$$

deci condiția Leapunov este îndeplinită.

3.48. a) Fie $x \in (0, 1)$, $\varepsilon_n(x)$ a n cifră în scrierea zecimală a lui x

$$f_n(x) = \sum_{k=1}^n \varepsilon_n(x) \text{ și } A_n(y) = \left\{ x \in (0, 1), \frac{2f_n(x) - 9n}{\sqrt{33n}} < y \right\}.$$

Dacă m este măsura Lebesgue să se arate că

$$\lim_{n \rightarrow \infty} m(A_n(y)) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{t^2}{2}} dt$$

b) Orice $x \in (0, 1)$ (exceptînd o mulțime numărabilă) se poate pune sub forma $x = \sum_{n=1}^{\infty} \frac{\varepsilon_n(x)}{q_1 q_2 \cdots q_n}$ unde $\varepsilon_n(x)$ poate lua valorile $0, 1, \dots, q_n - 1$. Punînd

$$f_n(x) = \sum_{k=1}^n \varepsilon_n(x) \text{ și } B_n(y) = \left\{ x \in (0, 1), \frac{f_n(x) - \frac{1}{2} \sum_{k=1}^n (q_k - 1)}{\sqrt{\frac{1}{12} \sum_{k=1}^n (q_k^2 - 1)}} < y \right\}$$

să se arate că dacă $\lim_{n \rightarrow \infty} \frac{\max_{1 \leq k \leq n} q_k}{\sqrt{\sum_{k=1}^n q_k^2}} = 0$, atunci

$$\lim_{n \rightarrow \infty} m(B_n(y)) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{t^2}{2}} dt.$$

Soluție. a) Se consideră cîmpul de probabilitate $((0, 1) \mathfrak{B}_{(0,1)}, m)$, m măsura Lebesgue.

Fie f variabila aleatoare repartizată uniform pe $(0, 1)$ care corespunde alegerii la întîmplare a unei număr în intervalul $(0, 1)$. Atunci $g_n = \varepsilon_n(f)$ sînt variabile aleatoare independente identic repartizate cu $M(g_n) = \frac{9}{2}$ și $D^2(g_n) = \frac{33}{4}$. Se aplică în continuare teorema limită centrală.

b) Se consideră același cîmp de probabilitate ca la punctul a) se ia din nou variabila aleatoare f și se pune în evidență șirul de variabile aleatoare $g_n = \varepsilon_n(f)$. Avem

$$M(g_n) = \frac{q_n - 1}{2} \quad D^2(g_n) = \frac{q_n^2 - 1}{12} \quad \text{și} \quad M(|g_n - M(g_n)|^3) < Kq_n^3,$$

condiția lui Leapunov fiind îndeplinită se aplică în continuare teorema limită centrală.

3.48. Fie (f_n) un șir de variabile aleatoare independente identic repartizate cu valori în $\{-1, 0, \dots\}$ și așa încît $P(f_n = l) = \frac{e^{-1}}{(l+1)!}$; $l = -1, 0, 1 \dots$ și fie $s_n = f_1 + \dots + f_n$.

Să se arate că șirurile $P(s_n < 0)$ și $P(s_n > 0)$ sînt crescătoare și au limita $1/2$. Să se deducă apoi egalitatea

$$\lim_{n \rightarrow \infty} e^{-n} \sum_{l=0}^n \frac{n^l}{l!} = \frac{1}{2}.$$

Soluție. Deoarece $1 + f_1$ este repartizată Poisson cu parametrul 1, rezultă că $M(f_1) = M(1 + f_1) - 1 = 1 - 1 = 0$ și $D^2(f_1) = D^2(1 + f_1) = 1$. Conform teoremei limită centrală (vezi iii din rezumatul teoretic) avem $\frac{s_n}{n} \Rightarrow N(0,1)$.

$$\begin{aligned} \text{În particular obținem că } P(s_n < 0) &= P\left(\frac{s_n}{n} < 0\right) \rightarrow \int_{-\infty}^0 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \\ &= \frac{1}{2} \text{ și } P(s_n > 0) = P\left(\frac{s_n}{n} > 0\right) \rightarrow \int_0^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{1}{2}. \end{aligned}$$

Deoarece $n + s_n$ este repartizată Poisson cu parametrul n , rezultă

$$P(s_n < 0) = P(n + s_n < n) = \sum_{k=0}^{n-1} \frac{n^k}{k!} e^{-n}, \text{ așa că } \sum_{k=0}^n \frac{n^k}{k!} e^{-n} \rightarrow \frac{1}{2}.$$

Din formula lui Taylor de ordinul n cu rest integral rezultă că

$$e^n = \sum_{k=0}^n \frac{n^k}{k!} + \frac{1}{n!} \int_0^n e^t (n-t)^n dt,$$

$$\text{deci } P(s_n > 0) = 1 - e^{-n} \left[e^n - \frac{1}{n!} \int_0^n e^t (n-t)^n dt \right] = \frac{1}{n!} \int_0^n e^{-t} t^n dt.$$

Atunci $P(s^{n+1} > 0) - P(s_n > 0) = \frac{1}{(n+1)!} \left[\int_0^n e^{-t} t^{n+1} dt + \int_n^{n+1} e^{-t} t^{n+1} dt - \int_0^n e^{-t} (n+1) t^n dt \right]$, de unde aplicînd ultimei integrale formula de integrare prin părți obținem :

$$P(s_{n+1} > 0) - P(s_n > 0) = \frac{e^{-n} n^{n+1}}{(n+1)!} \left[1 - \int_0^1 e^{-t} \left(1 + \frac{t}{n}\right)^n dt \right]$$

Analog cu ajutorul formulei lui Taylor de ordinul $n-1$ se obține că

$$P(s_{n+1} < 0) - P(s_n < 0) = \frac{e^{-n} n^n}{n!} \left[\int_0^1 e^{-t} \left(1 + \frac{t}{n}\right)^{n+1} dt - 1 \right]$$

În fine este vizibil că totul se reduce la inegalitățile clasice

$$\left(1 + \frac{t}{n}\right)^n \leq e^t \leq \left(1 + \frac{t}{n}\right)^{n+1}$$

pentru $0 \leq t \leq 1$.

3.50. Fie (f_n) , $n \geq 1$, un șir de variabile aleatoare independente care au valorile $\pm n^\lambda$ cu probabilitățile $\frac{1}{2n^\lambda}$ și 0 cu probabilitățile $1 - \frac{1}{n^\lambda}$, $0 < \lambda < 1$. Să se arate că pentru șirul $(f_n)_{n \geq 1}$ este îndeplinită teorema limită centrală.

Soluție. Avem $M(f_n) = 0$, $D^2(f_n) = n^\lambda$. Fie $\delta > 0$. Atunci

$$\begin{aligned} \sum_{k=1}^n \frac{M(f_k^{2+\delta})}{\left(\sum_{k=1}^n D^2(f_k)\right)^{1+\frac{\delta}{2}}} &= \frac{\sum_{k=1}^n k^{\lambda(1+\delta)}}{\left(\sum_{k=1}^n k^\lambda\right)^{1+\frac{\delta}{2}}} < K \frac{(n+1)^{\lambda+\delta\lambda+1}}{n^{(\lambda+1)\left(1+\frac{\delta}{2}\right)}} = \\ &= K \frac{\left(1 + \frac{1}{n}\right)^{\lambda+\delta\lambda+1}}{\frac{\delta}{2}(1-\lambda)} \rightarrow 0 \quad (K \text{ constantă}) \end{aligned}$$

și este deci îndeplinită condiția Leapunov.

3.51. Se dă șirul de variabile aleatoare (f_n) care iau cu probabilități egale doar două valori $\pm \sqrt{an}$. Să se arate că dacă $\lim_{n \rightarrow \infty} \frac{\sum_{k=1}^n a_k^2}{\left(\sum_{k=1}^n a_k\right)^2} = 0$,

atunci (f_n) satisface teorema limită centrală.

Soluție. Cum $M(f_n) = 0$ și $D^2(f_n) = a_n$ avem

$$\begin{aligned} & \frac{1}{D^2(f_1 + \dots + f_n)} \sum_{k=1}^n \int_{|f_k - M(f_k)| \geq \varepsilon D(f_1 + \dots + f_n)} (f_k - M(f_k))^2 dP = \\ & = \frac{1}{\sum_{k=1}^n a_k} \sum_{k=1}^n \int_{|f_k| \geq \varepsilon \sqrt{\sum_{k=1}^n a_k}} a_k dP = \frac{1}{\sum_{k=1}^n a_k} a_k P\left(|f_k| \geq \varepsilon \sqrt{\sum_{k=1}^n a_k}\right) \leq \\ & \leq \frac{1}{\sum_{k=1}^n a_k} \cdot \sum_{k=1}^n a_k \frac{a_k}{\varepsilon^2 \sum_{k=1}^n a_k} = \frac{1}{\varepsilon^2} \frac{\sum_{k=1}^n a_k^2}{\left(\sum_{k=1}^n a_k\right)^2} \rightarrow 0 \end{aligned}$$

și este deci îndeplinită condiția Lindeberg.

CAPITOLUL IV

ELEMENTE DE STATISTICĂ MATEMATICĂ

Completitudine, suficiență. Fie $X: (E, \mathfrak{E}) \rightarrow (F, \mathfrak{F})$ o variabilă aleatoare cu $F \subset \mathbb{R}^n$, $n \geq 1$, $\mathfrak{F} = \mathfrak{B}_F$ și fie $\mathfrak{P} = \{P_\theta(\cdot), \theta \in \Theta\}$ o familie de repartiții de probabilitate pe (F, \mathfrak{F}) .

O aplicație măsurabilă $t: (F, \mathfrak{F}) \rightarrow (G, \mathfrak{G})$ cu $G \subset \mathbb{R}^d$, $d \geq 1$, $\mathfrak{G} = \mathfrak{B}_G$ se numește *statistică*.

Familia de repartiții $\mathfrak{P} = \{P_\theta(\cdot), \theta \in \Theta\}$ este *completă* dacă are loc următoarea implicație:

$$\int_F f(x) dP_\theta(x) = 0 \text{ pentru orice } \theta \in \Theta \Rightarrow f(x) = 0 \text{ } P_\theta - a. s. \text{ pentru orice } \theta \in \Theta.$$

O statistică t este *completă* pentru familia de repartiții $\mathfrak{P} = \{P_\theta(\cdot), \theta \in \Theta\}$ dacă familia de probabilități indusă de t pe (G, \mathfrak{G}) este completă.

Un concept de mare însemnătate în statistica matematică este acela de suficiență. Fischer a introdus conceptul de statistică suficientă ca fiind cea statistică ce „conține toată informația relevantă”. Neyman, Halmos și Savage au dat metode de determinare a statisticilor suficiente.

a) *Cazul discret.* O statistică t este *suficientă* pentru familia de repartiții discrete $\mathfrak{P} = \{P_\theta(\cdot), \theta \in \Theta\}$ (suficientă pentru $\theta \in \Theta$) dacă repartiția condiționată de t , $P_\theta(\cdot|t)$ există aproape sigur pentru orice $\theta \in \Theta$ și este independentă de θ .

Criteriul de favorizare (Neyman): Fie $P_\theta(x) = P_\theta(X = x)$. O condiție necesară și suficientă ca statistica t să fie suficientă pentru $\theta \in \Theta$ este ca să existe factorizarea $P_\theta(x) = g_\theta(t(x)) \varphi(x)$, unde primul factor poate să depindă de θ , dar depinde de x doar prin intermediul lui $t(x)$, iar al doilea factor este nenegativ, independent de θ , astfel încât pentru orice y cu $P_\theta(y) > 0$ să avem $\sum_{x \in A} \varphi(x) > 0$, unde $A = \{x|t(x) = y\}$.

b) *Cazul absolut continuu:* $P_\theta \ll \mu$, $\theta \in \Theta$. O statistică t este *suficientă* pentru familia de repartiții $\mathfrak{P} = \{P_\theta(\cdot), \theta \in \Theta\}$ dacă pentru orice funcție f , P_θ — integrabilă pentru orice θ , valoarea medie condiționată $M_\theta(f|t)$ există aproape sigur pentru orice $\theta \in \Theta$ și este independentă de θ .

Criteriul de factorizare (Halmos și Savage). Fie $p_\theta(x)$ densitatea de repartiție în raport cu μ . O condiție necesară și suficientă ca statistica t să fie suficientă pentru $\theta \in \Theta$ este ca să existe o funcție nenegativă φ , măsurabilă, μ — integrabilă și o familie de funcții măsurabile $\{g_\theta, \theta \in \Theta\}$ așa încât să aibă loc o factorizare $p_\theta(x) = g_\theta(t(x)) \varphi(x)$ μ — aproape sigur, pentru orice $\theta \in \Theta$.

Observație privind notația. În cursul acestui capitol vom utiliza, în funcție de necesitățile scrierii următoarele notații echivalente: $P_0(\cdot)$ sau $P(\cdot; 0)$ și $p_0(x)$ sau $p(x; 0)$.

Selecție. Concordanța repartiției de selecție cu o repartiție teoretică. Considerăm o populație modelată prin cimpul de probabilitate (E, \mathfrak{E}, P) și o caracteristică modelată printr-o variabilă aleatoare $X: (E, \mathfrak{E}) \rightarrow (S, \mathfrak{S})$.

Spațiul selecțiilor bernoulliene n -dimensionale este spațiul produs $(E^n, \mathfrak{E}^n, P^n)$, iar vectorul de selecție n -dimensional este $X^n = (X_1, \dots, X_n)$, $X^n: (E^n, \mathfrak{E}^n) \rightarrow (S^n, \mathfrak{S}^n)$ cu componentele independente și identic repartizate cu X , $X_i(\omega^n) = X(\omega_i)$, unde $\omega^n = (\omega_1, \dots, \omega_n)$.

Repartiția variabilei X aflată sub cercetare, $F = P \circ X^{-1}$ este necunoscută. Ea este bine aproximată de repartiția de selecție $F_n(y, \omega^n) = \frac{1}{n} \text{card} \{X_i(\omega^n) | X_i(\omega^n) < y\}$ în sensul următoarelor teoreme:

Teorema lui Glivenko. Fie $(E^\infty, \mathfrak{F}^\infty, P^\infty)$ spațiul selecțiilor bernoulliene infinit-dimensionale și $F_n(y, \omega^\infty) = F_n(y, \omega^n)$. Avem

$$P^\infty(\{\omega^\infty | \limsup_{n \rightarrow \infty} \sup_{y \in \mathbb{R}} |F_n(y, \omega^\infty) - F(y)| = 0\}) = 1.$$

Teorema lui Kolmogorov. Dacă F este continuă, atunci

$$\lim_{n \rightarrow \infty} P^\infty \left[\sup_{y \in \mathbb{R}} \sqrt{n} |F_n(y, \omega^\infty) - F(y)| > z \right] = \begin{cases} 0 & z < 0, \\ 2 \sum_{r=1}^{\infty} (-1)^{r-1} \exp(-2r^2 z^2), & z > 0, \end{cases}$$

Pornind de la teorema lui Kolmogorov, se poate construi testul pentru verificarea concordanței între repartiția de selecție și o repartiție teoretică F complet specificată:

a) Se face o selecție bernoulliană x_1, \dots, x_n asupra lui X , de volum n suficient de mare;

b) Se calculează $K_n = \sqrt{n} \max_{1 \leq i \leq n} |F_n(x_i) - F(x_i)|$;

c) Se determină din tabelul repartiției Kolmogorov cuantila $h_{1-\alpha}$ de rang $(1 - \alpha)$;

d) Decizia de acceptare sau de respingere a ipotezei " $F(x)$ este repartiția reală a lui X " se ia conform regulii:

— dacă $K_n < h_{1-\alpha}$, se acceptă ipoteza de concordanță

— dacă $K_n > h_{1-\alpha}$, se respinge ipoteza că F este repartiția reală a lui X .

Testul descris mai sus se aplică în cazul repartiției F complet specificate. În cazul repartiției normale cu media și dispersia nespecificate a fost construită o variantă a testului lui Kolmogorov, testul lui Lilliefors pentru verificarea normalității.

Un alt binecunoscut test pentru verificarea concordanței repartiției de selecție cu o repartiție teoretică este testul χ^2 pentru datele grupate.

Presupunem că există o partiție a spațiului valorilor caracteristicii $X: S = \bigcup_{j=1}^r S_j$,

$S_i \cap S_j = \emptyset$ $i \neq j$; $S_i \in \mathfrak{S}$.

Fie $p_i = \int_{S_i} dF(x)$; $\sum_{i=1}^r p_i = 1$ și $q_i(\omega^n) = \int_{S_i} dF_n(x, \omega^n)$, $\sum_{i=1}^r q_i(\omega^n) = 1$, $\omega^n \in E^n$.

Notăm $v_i(\omega^n) = \frac{1}{n} \cdot q_i(\omega^n)$, $\omega^n \in E^n$.

Teorema lui Pearson. Variabila aleatoare $H_n^2 = \sum_{i=1}^r \frac{(v_i - np_i)^2}{np_i}$ converge în repartiție pentru n tinzând la infinit la o variabilă repartizată χ^2 cu $(r-1)$ grade de libertate (vezi problema 2.24. din cap. II pentru definiția repartiției χ^2).

Această teoremă stă la baza următorului test pentru verificarea concordanței între $F_n(x)$ și $F(x)$:

a) Se face o selecție bernoulliană x_1, \dots, x_n asupra lui X , de volum n suficient de mare;

b) Se calculează $H_n^2 = \sum_{i=1}^r \frac{(v_i - np_i)^2}{np_i}$

c) Se determină din tabelul repartiției χ^2 cu $(r-1)$ grade de libertate cuantila de rang $(1-\alpha)$, $h_{r-1}^2; 1-\alpha$;

d) Decizia de acceptare sau de respingere a ipotezei " $F(x)$ este repartiția reală a lui X " se ia conform regulii:

— dacă $H_n^2 < h_{r-1}^2; 1-\alpha$, se acceptă ipoteza de concordanță.

— dacă $H_n^2 > h_{r-1}^2; 1-\alpha$, se respinge ipoteza că F este repartiția reală a lui X .

În aplicarea practică a testului χ^2 de obicei nu se cunoaște partiția lui S . STASS 7/22-65 recomandă împărțirea lui S în $r \approx 1 + 3,322 \lg n$ intervale de lungimi egale.

De asemenea, în practică repartiția teoretică F nu este complet specificată, o parte dintre parametrii săi trebuind să fie estimați. Dacă se introduc s relații pentru estimarea parametrilor, variabila H_n^2 urmează asimptotic o repartiție χ^2 cu $(r-s-1)$ grade de libertate, iar testul χ^2 se adaptează corespunzător.

Estimarea parametrilor. Considerăm o populație $(E, \mathcal{A}, P_\theta)$ cu $0 \in \Theta \subset R^k$ și o caracteristică aleatoare X . O estimatie a parametrului θ bazată pe o selecție bernoulliană n -dimensională este o aplicație măsurabilă $t_n: (S^n, \mathfrak{S}^n) \rightarrow (\Theta, \mathfrak{B}_\Theta)$.

t_n este o estimatie nedepasată a lui θ dacă $M_\theta(t(X^n)) = \theta$, unde media se face în raport cu P_θ .

Utilitatea noțiunii de suficiență în problema estimărilor nedepasate, de dispersie minimă este dată de următoarea teoremă.

Teorema Rao-Blackwell. Fie t o statistică suficientă pentru $\theta \in \Theta$ și \tilde{g} o estimatie nedepasată a unei funcții cu valori reale $g(\theta)$. Atunci $M_\theta(\tilde{g}|t)$ este de asemenea o estimatie nedepasată a lui $g(\theta)$, iar dispersia sa este mai mică sau egală cu dispersia lui \tilde{g} pentru orice θ .

În problemele de estimare a parametrilor se pune în general problema găsirii unui estimator nedepasat, de dispersie minimă pentru parametrul θ . În anumite condiții de regularitate, se poate găsi o margine inferioară a dispersiei unui estimator.

Inegalitatea Rao-Cramer. Fie X_1, \dots, X_n variabile de selecție independente, identic repartizate, cu densitatea de repartiție $p_\theta(x)$ în raport cu o măsură σ — finită μ (adică $p_\theta(x)$ este derivata Radon-Nicodym a repartiției P_θ o X_i^{-1} în raport cu măsura μ). Fie $\delta(X_1, \dots, X_n)$ o estimatie (nu obligatoriu nedepasată a lui θ).

Dacă $M_\theta(\delta(X_1, \dots, X_n)) = \theta + b(\theta)$ și următoarele condiții de regularitate sînt satisfăcute.

- i) Θ este un interval deschis al dreptei reale;
- ii) $\frac{\partial p_\theta(x)}{\partial \theta}$ există pentru toți $\theta \in \Theta$;
- iii) $\int \prod_{i=1}^n p_\theta(x_i) \prod_i d\mu(x_i)$ poate fi derivată în raport cu θ sub integrală;
- iv) $M_\theta \left[\frac{\partial \log p_\theta(X)}{\partial \theta} \right]^2 > 0$ pentru orice $\theta \in \Theta$;
- v) $\int \delta(x_1, \dots, x_n) \prod_{i=1}^n p_\theta(x_i) \prod_{i=1}^n d\mu(x_i)$ poate fi derivată în raport cu θ sub integrală; atunci dispersia estimației δ satisface inegalitatea

$$D_\theta^2(\delta) \geq \frac{[1 + b'(\theta)]^2}{n \cdot M_\theta \left[\left(\frac{\partial \log p_\theta(x)}{\partial \theta} \right)^2 \right]}.$$

În general se verifică ușor condițiile (i), (ii) și (iv). Pentru a verifica condițiile (iii) și (v) se procedează în modul următor: Presupunem că pentru orice $\theta^0 \in \Theta$ există numerele a și b și funcțiile μ - integrabile G_1 și G_2 așa încît $\theta^0 \in (a, b)$, $\left| \frac{\partial p_\theta(x)}{\partial \theta} \right| < G_1(x)$ pentru $\theta \in (a, b)$ și $x \in S - N$, unde N este o mulțime de măsură μ egală cu zero și așa încît $\left| \delta(x) \frac{\partial p_\theta(x)}{\partial \theta} \right| < G_2(x)$ pentru $\theta \in (a, b)$ și $x \in S - N$. Atunci (iii) și (v) sînt verificate dacă

$$\int_S G_i(x) d\mu(x) < \infty; \quad i = 1, 2.$$

Valoarea $I(\theta) = M_\theta \left[\frac{\partial \log p_\theta(X_1, \dots, X_n)}{\partial \theta} \right]^2$ se numește *informația lui Fisher*.

Printr-un calcul direct se verifică că: $I(\theta) = n i(\theta)$, unde

$$i(\theta) = M_\theta \left[- \frac{\partial^2 \log p_\theta(X)}{\partial \theta^2} \right].$$

Pe de altă parte, se utilizează frecvent notația

$$U(\theta) = \frac{\partial \log p_\theta(X_1, \dots, X_n)}{\partial \theta}.$$

Avem $M_\theta(U(\theta)) = 0$, $D_\theta^2(U(\theta)) = I(\theta)$.

Eficiența unei estimații nedepasate t_n este egală cu raportul dintre dispersia sa și marginea inferioară a dispersiei dată de inegalitatea Rao-Cramer:

$$e(t_n) = \frac{D_\theta^2(t_n)}{1/I(\theta)}.$$

Pentru un parametru θ , k -dimensional, se consideră *matricea informațională Fisher* :

$$I(\theta) = \left\| M_{\theta} \left[\frac{\partial \log p_{\theta}(X_1, \dots, X_n)}{\partial \theta_i} \cdot \frac{\partial \log p_{\theta}(X_1, \dots, X_n)}{\partial \theta_j} \right] \right\|.$$

Generalizarea inegalității Rao–Cramer în cazul multidimensional pentru un estimator nedepășat \tilde{g} al lui θ stabilește faptul că matricea covariațională

$$\text{Cov}_{\theta}(\tilde{g}) = M_{\theta}[(\tilde{g} - \theta)(\tilde{g} - \theta)^t]$$

are proprietatea că $\text{Cov}_{\theta}(\tilde{g}) - [I(\theta)]^{-1}$ este o matrice pozitiv semidefinită.

Una dintre cele mai vechi și mai importante metode de estimare a parametrilor este *metoda verosimilității maxime*.

Pentru o selecție bernoulliană n -dimensională, *funcția de verosimilitate* este în cazul discret $L_n(x_1, \dots, x_n, \theta) = \prod_{i=1}^n P_{\theta}(X_i = x_i)$ iar în cazul repartițiilor absolut continue

în raport cu măsura Lebesgue $L_n(x_1, \dots, x_n, \theta) = \prod_{i=1}^n p_{\theta}(x_i)$. Metoda verosimilității

maxime revine la găsirea unei valori a parametrului care maximizează funcția de verosimilitate. Presupunind că $L_n(x_1, \dots, x_n; \theta)$ este o funcție pozitivă, diferențiabilă în raport cu θ , iar maximumul nu este atins pe frontiera mulțimii valorilor admisibile

pentru θ , problema revine la rezolvarea ecuației de verosimilitate $\frac{\partial L_n}{\partial \theta} = 0$, sau echi-

valent $\frac{\partial \log L_n}{\partial \theta} = 0$.

Teoremă. Presupunem că sînt verificate următoarele condiții :

i) $\frac{\partial \log p_{\theta}(x)}{\partial \theta}$, $\frac{\partial^2 \log p_{\theta}(x)}{\partial \theta^2}$, $\frac{\partial^3 \log p_{\theta}(x)}{\partial \theta^3}$ există pentru aproape toți x și pentru toți θ dintr-un interval Θ al dreptei reale;

ii) Pentru orice $\theta \in \Theta$ avem $\left| \frac{\partial p_{\theta}(x)}{\partial \theta} \right| < A_1(x)$, $\left| \frac{\partial^2 p_{\theta}(x)}{\partial \theta^2} \right| < A_2(x)$, $\left| \frac{\partial^3 \log p_{\theta}(x)}{\partial \theta^3} \right| < A_3(x)$, unde $A_1(x)$, $A_2(x)$ sînt funcții integrabile, iar $\int A_3(x) p_{\theta}(x) dx < M$, cu M independent de θ

iii) Pentru orice $\theta \in \Theta$, $\int_R \left(\frac{\partial \log p_{\theta}(x)}{\partial \theta} \right)^2 p_{\theta}(x) dx$ este finită și pozitivă.

Atunci ecuația de verosimilitate are o soluție care converge la θ în probabilitate (P_{θ}) pentru n tinzînd la infinit. Această soluție este asimptotic eficientă (nedepășată, iar dispersia sa converge la marginea inferioară dată de inegalitatea Rao–Cramer). Mai mult estimăția de verosimilitate maximă (E.V.M) este repartizată asimptotic normal.

E.V.M. are o serie de alte proprietăți remarcabile dintre care enumerăm :

- Dacă există o estimăție pentru care marginea inferioară a inegalității Rao–Cramer este atinsă, atunci ecuația de verosimilitate are o soluție unică.
- Dacă există o statistică suficientă pentru $\theta \in \Theta$, atunci E.V.M. este funcție doar de această statistică.

O altă metodă frecvent utilizată este estimarea prin *metoda celor mai mici pătrate*, pe care o vom prezenta doar pentru problema estimării coeficienților în *regresia polinomială*.

Considerăm modelul regresiei polinomiale $Y = X\gamma + \varepsilon$, unde $Y^t = (y_1, \dots, y_m)$ și $\varepsilon^t = (\varepsilon_1, \dots, \varepsilon_m)$ sînt vectori aleatori

$$X = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^k \\ 1 & x_2 & x_2^2 & \dots & x_2^k \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_m & x_m^2 & \dots & x_m^k \end{bmatrix}$$

este matricea constituită cu ajutorul valorilor x_i , $i = 1, \dots, m$ observate pentru o variabilă reală, iar $\gamma^t = (\gamma_0 \dots \gamma_k)$ este vectorul coeficienților de regresie.

Presupunem că $M(Y) = X\gamma$, rangul lui X este $(k + 1)$, iar matricea de covarianță a vectorului aleator Y este $\sigma^2 I$.

Fie $g^t = (g_0, \dots, g_k)$ o estimatie pentru vectorul coeficienților de regresie. Atunci eroarea de ordinul doi este

$$\varepsilon^t \varepsilon = (Y - Xg)^t (Y - Xg).$$

Estimația obținută prin metoda celor mai mici pătrate, notată $\tilde{\gamma}$ realizează eroarea pătratică minimă. Ea este soluția sistemului de ecuații normale

$$\frac{\partial}{\partial g_j} (Y - Xg)^t (Y - Xg) = 0, \quad j = 0, 1, \dots, k,$$

și anume $\gamma = (X^t X)^{-1} X^t Y$.

O altă abordare a problemei estimării parametrilor este oferită de punctul de vedere bayesian.

O *funcție de pierdere* este o aplicație $w: \Theta \times \Theta \rightarrow R_+$ măsurabilă în ambele argumente. *Funcția de risc* asociată unei estimatii g a parametrului θ este definită de relația

$$R_g(\theta) = M_\theta(w(\theta, g(X_1, \dots, X_n))).$$

Riscul maxim asociat unei estimatii g este $R(g) = \sup_{\theta \in \Theta} R_g(\theta)$. *Riscul mediu* asociat

estimației g și unei probabilități a priori pe $(\Theta, \mathfrak{B}_\Theta)$ este $\bar{R}(g) = \int_{\Theta} R_g(\theta) d\lambda(\theta)$

\tilde{g} este o *estimație minimax* dacă $R(\tilde{g}) \leq R(g)$, oricare ar fi g , estimație a lui θ .

g^* este o *estimație Bayes* dacă $\bar{R}(g^*) \leq \bar{R}(g)$, oricare ar fi g , estimație a lui θ .

În cazul unui parametru real și al unei funcții de pierdere pătratică, $w(\theta, \theta') = (\theta - \theta')^2$, estimația Bayes este media a posteriori a parametrului,

$$g^*(x_1, \dots, x_n) = \int_{\Theta} \theta p(\theta | x_1, \dots, x_n) d\theta, \text{ unde}$$

$$p(\theta | x_1, \dots, x_n) = \frac{p(\theta) \prod_{i=1}^n p_\theta(x_i)}{\int_{\Theta} \prod_{i=1}^n p_\theta(x_i) p(\theta) d\theta},$$

iar $p(\theta)$ este o densitate de repartiție a priori.

Alături de problema estimării punctuale a valorii parametrului, în statistica matematică se pune și problema construirii unor intervale de estimare, sau de încredere, pentru parametrul unei repartiții.

În cazul unui parametru real, dacă există două funcții A și B , măsurabile, așa încît

$$P_{\theta}(A(X_1, \dots, X_n) < \theta < B(X_1, \dots, X_n)) = 1 - \alpha,$$

atunci $[A(X_1, \dots, X_n); B(X_1, \dots, X_n)]$ este un *interval de estimare* pentru θ , cu coeficientul de încredere $(1 - \alpha)$.

Construcția unui interval de estimare se face utilizînd o funcție $g: S^n \times \Theta \rightarrow R$, măsurabilă în x_1, \dots, x_n , monotonă ca funcție în θ și a cărei repartiție nu depinde de θ (dacă un asemenea g există). Pentru $\alpha > 0$ fixat, fie $a(\alpha)$, $b(\alpha)$ două numere cu proprietatea

$$P_{\theta}(a(\alpha) < g(X_1, \dots, X_n); 0) < b(\alpha) = 1 - \alpha.$$

Pentru x_1, \dots, x_n fixați, din monotonia lui g rezultă că există $A(x_1, \dots, x_n; \alpha)$ și $B(x_1, \dots, x_n; \alpha)$ așa încît

$$P_{\theta}(A(X_1, \dots, X_n; \alpha) < \theta < B(X_1, \dots, X_n; \alpha)) = 1 - \alpha.$$

Verificarea ipotezelor statistice; teste parametrice. Considerăm populația $(E, \mathfrak{S}, P_{\theta})$ $\theta \in \Theta \subset R^k$, caracteristica aleatoare X și vectorul de selecție n -dimensional X^n . Notăm cu $F_n(x^n; \theta)$ funcția de repartiție a lui X^n .

Se numește *ipoteză statistică* familia

$$H = \{F_n(x^n; \theta), \theta \in \Theta_0\} \text{ cu } \Theta_0 \subset \Theta, \text{ notată } H: \{\theta \in \Theta_0\}.$$

Se numește *ipoteză alternativă* lui H familia

$$H_A = \{F_n(x^n; \theta), \theta \in \Theta \setminus \Theta_0\}, \text{ notată } H_A: \{\theta \in \Theta \setminus \Theta_0\}.$$

O ipoteză statistică se numește *simplică* dacă $\Theta_0 = \{\theta_0\}$.

O mulțime $W \subset S^n$, măsurabilă se numește *regiune critică* pentru ipoteza H dacă decidem să respingem ipoteza H atunci cînd valoarea x^n a selecției cade în mulțimea W , iar atunci cînd x^n nu aparține lui W , acceptăm ipoteza H .

Presupunem că $\Theta = \{\theta_0, \theta_1\}$ și considerăm ipoteza $H: \{\theta = \theta_0\}$ cu alternativa $H_A: \{\theta = \theta_1\}$. Pentru un test bazat pe regiunea critică W pot apărea două tipuri de erori: *eroarea de primul ordin*, care revine la respingerea unei ipoteze adevărate și a cărei probabilitate este $P_{\theta_0}^n(X^n \in W) = \alpha$ și *eroarea de ordinul doi*, care revine la acceptarea unei ipoteze false și a cărei probabilitate este $P_{\theta_1}^n(X^n \notin W) = \beta$.

Fie $\mathfrak{W}_{\alpha} = \{W \in \mathfrak{S}^n | P_{\theta_0}^n(X^n \in W) = \alpha\}$, unde $\alpha \in (0, 1)$ este o valoare fixată („prag de semnificație”). Mulțimea $W^* \in \mathfrak{W}_{\alpha}$ se numește *cea mai bună regiune critică*, pentru ipoteza $H: \{\theta = \theta_0\}$ cu alternativa $H_A: \{\theta = \theta_1\}$ dacă $P_{\theta_1}^n(X^n \in W^*) \geq P_{\theta_1}^n(X^n \in W)$ pentru orice $W \in \mathfrak{W}_{\alpha}$.

Lema Neyman-Pearson: presupunem că $\mathfrak{W}_{\alpha} \neq \emptyset$ și că există $W^* \in \mathfrak{W}_{\alpha}$ și $K \geq 0$ încît

$$P_{\theta_1}^n(X^n \in W^*) \geq K \cdot P_{\theta_0}^n(X^n \in W^*)$$

$$P_{\theta_1}^n(X^n \notin W^*) < K \cdot P_{\theta_0}^n(X^n \notin W^*).$$

Atunci W^* este cea mai bună regiune critică pentru ipoteza $H: \{\theta = \theta_0\}$ cu alternativa $H_A: \{\theta = \theta_1\}$ la pragul de semnificație α .

Lema Neyman-Pearson stabilește deci forma celei mai bune regiuni critice: $W^* = \{x^n | F_n(x^n, \theta_1) \geq K \cdot F_n(x^n, \theta_0)\}$. Constanta K se determină așa încît pragul de semnificație să fie egal cu valoarea α fixată.

Să considerăm acum cazul ipotezei compuse $H: \{\theta \in \Theta_0\}$ cu alternativa compusă $H_A: \{\theta \in \Theta - \Theta_0\}$. Pentru $0 < \alpha < 1$ fixat considerăm familia $\mathfrak{W}_\alpha = \{W \in \mathfrak{S}^n | P_\theta^n(X^n \in W) < \alpha \text{ pentru toți } \theta \in \Theta_0\}$. Mulțimea $W^* \in \mathfrak{W}_\alpha$ se numește regiunea critică uniform cea mai bună pentru ipoteza H dacă $P_\theta^n(X^n \in W^*) \geq P_\theta^n(X^n \in W)$ pentru orice $\theta \in \Theta - \Theta_0$ și orice $W \in \mathfrak{W}_\alpha$. Testul uniform cel mai puternic nu există întotdeauna.

O metodă larg utilizată pentru verificarea ipotezelor statistice este *testul raportului de verosimilitate*. Dacă notăm prin $p(x_1, \dots, x_n, \theta)$ densitatea de repartiție a vectorului de selecție X^n și presupunem că aceasta este măsurabilă, raportul de verosimilitate pentru ipoteza $H: \{\theta \in \Theta_0\}$ este definit prin

$$\Lambda(x_1, \dots, x_n) = \frac{\sup_{\theta \in \Theta_0} p(x_1, \dots, x_n, \theta)}{\sup_{\theta \in \Theta} p(x_1, \dots, x_n, \theta)}$$

Dacă se cunoaște repartiția lui $\Lambda(X_1, \dots, X_n)$, testul este dat de regiunea critică $W = \{(x_1, \dots, x_n) | \Lambda(x_1, \dots, x_n) < k_\alpha\}$, constanta k_α fiind determinată așa încît $P_\theta(W) = \alpha$ pentru orice $\theta \in \Theta_0$.

Un rezultat asimptotic important, utilizat foarte frecvent în probleme este următorul:

Teoremă. Presupunem că sînt verificate condițiile de regularitate pentru existența estimărilor de verosimilitate maximă $\hat{\theta}(x_1, \dots, x_n)$, $\hat{\theta}(x_1, \dots, x_n)$, cu

$$p(x_1, \dots, x_n, \hat{\theta}) = \sup_{\theta \in \Theta} p(x_1, \dots, x_n, \theta)$$

$$p(x_1, \dots, x_n, \hat{\theta}) = \sup_{\theta \in \Theta_0} p(x_1, \dots, x_n, \theta).$$

Atunci $-2 \log \Lambda = 2(\log p(X_1, \dots, X_n, \hat{\theta}) - \log p(X_1, \dots, X_n, \hat{\theta}))$ converge în repartiție la o variabilă repartizată χ^2 cu r grade de libertate, r fiind numărul de restricții impuse asupra lui θ pentru definirea lui Θ_0 .

Testul raportului de verosimilitate (T.R.V.) asimptotic se construiește utilizînd cuantila de rang α a repartiției χ^2 .

În cazul ipotezelor simple cu alternative simple ($\Theta = \{\theta_1, \theta_0\}$) prezentăm și *testul secvențial al raportului probabilităților* (TSRP) pentru verificarea ipotezei $H: \{\theta = \theta_0\}$ cu alternativa $H_A: \{\theta = \theta_1\}$.

Se numește *raport al probabilităților* funcția măsurabilă

$$R_n(x_1, \dots, x_n) = \frac{p(x_1, \dots, x_n, \theta_1)}{p(x_1, \dots, x_n, \theta_0)}$$

Fie constantele pozitive A și B , $B < A$. T.S.R.P. este definit la momentul n de regiunile:

$W_n^a = \{(x_1, \dots, x_n) | R_n(x_1, \dots, x_n) < B\}$ regiunea de acceptare a lui H

$W_n^r = \{(x_1, \dots, x_n) | R_n(x_1, \dots, x_n) \geq A\}$ regiunea de respingere a lui H

$W_n = \{(x_1, \dots, x_n) | B < R_n(x_1, \dots, x_n) < A\}$ regiunea de continuare a testului,

T.S.R.P. se încheie aproape sigur, adică

$$\lim_{n \rightarrow \infty} P_0^\infty((X_1, \dots, X_n) \in W_n) = 0 \text{ pentru orice } \theta.$$

Constantele A și B se determină astfel încât probabilitățile de eroare să fie α , respectiv β și ele sînt $A = \frac{1-\beta}{\alpha}$, $B = \frac{\beta}{1-\alpha}$.

§ 1. Completitudine, suficiență

4.1. Se consideră următoarele familii de repartiții discrete

$$\mathfrak{E}_1 = \{P(X=1) = p, P(X=0) = 1-p; 0 < p < 1\}$$

$$\mathfrak{E}_2 = \{P(X=k) = 1/m, k = 1, \dots, m; m \in N^*\}.$$

Utilizînd variabilele aleatoare X_1, \dots, X_n independente, identic repartizate conform familiei \mathfrak{E}_1 (respectiv \mathfrak{E}_2), să se arate că statistica

$T_1 = \sum_{i=1}^n X_i$ este suficiență pentru \mathfrak{E}_1 , iar $T_2 = \max_{1 \leq i \leq n} X_i$ este suficiență pentru \mathfrak{E}_2 .

Soluție. În cazul familiei \mathfrak{E}_1 , dacă $T_1 = t$, atunci exact C_n^t dintre valorile observate (x_1, \dots, x_n) sînt egale cu unu. Obținem

$$P(X_1 = x_1, \dots, X_n = x_n | T_1 = t) = \frac{p^t(1-p)^{n-t}}{C_n^t p^t (1-p)^{n-t}} = \frac{1}{C_n^t},$$

care nu depinde de parametrul p .

În cazul familiei \mathfrak{E}_2 avem:

$$P(T_2 = t) = P(T_2 \leq t) - P(T_2 < t) = \frac{t^n}{m^n} - \frac{(t-1)^n}{m^n}.$$

Atunci

$$P(X_1 = x_1, \dots, X_n = x_n | T_2 = t) = \frac{\frac{1}{m^n}}{\frac{t^n}{m^n} - \frac{(t-1)^n}{m^n}} = \frac{1}{t^n - (t-1)^n},$$

care nu depinde de parametrul m .

4.2. Fie X_1, \dots, X_n variabile aleatoare independente, avînd aceeași repartiție de tip exponențial, cu parametru unidimensional:

$$p(x, \theta) = \exp \{a(\theta) b(x) + c(\theta) + d(x)\}, \theta \in R.$$

Să se determine o statistică suficientă pentru această familie de repartiții și să se particularizeze rezultatul pentru următoarele cazuri:

- $N(m, \sigma_0^2)$, σ_0^2 cunoscut, $m \in R$;
- $N(m_0, \sigma^2)$, m_0 cunoscut, $\sigma^2 > 0$;
- gama $\gamma(k_0, \theta)$, k_0 cunoscut, $\theta \in R$;
- binomială $B_m(\theta)$, $0 < \theta < 1$;
- Poisson π_0 , $\theta > 0$.

Soluție. $p(x_1, \dots, x_n; \theta) = \exp \left\{ a(\theta) \sum_{j=1}^n b(x_j) + nc(\theta) + \sum_{j=1}^n d(x_j) \right\}$

Considerăm statistica $T(X_1, \dots, X_n) = \sum_{j=1}^n b(X_j)$. Avem

$$p(x_1, \dots, x_n; \theta) = \exp \{ a(\theta) t(x_1, \dots, x_n) + nc(\theta) \} \exp \left\{ \sum_{j=1}^n d(x_j) \right\}.$$

Conform teoremei de factorizare t este o statistică suficientă pentru familia de repartiții de tip exponențial.

$$\begin{aligned} \text{a) } p(x; m, \sigma_0^2) &= \exp \left\{ -\frac{1}{2} \ln 2\pi - \frac{1}{2} \ln \sigma_0^2 - \frac{(x-m)^2}{2\sigma_0^2} \right\} = \\ &= \exp \left\{ \frac{mx}{\sigma_0^2} - \frac{m^2}{2\sigma_0^2} - \frac{x^2}{2\sigma_0^2} - \frac{1}{2} \ln \sigma_0^2 - \frac{1}{2} \ln 2\pi \right\}. \end{aligned}$$

$T(X_1, \dots, X_n) = \sum_{i=1}^n X_i$ este o statistică suficientă pentru m .

$$\begin{aligned} \text{b) } p(x; m_0; \sigma^2) &= \exp \left\{ -\frac{1}{2} \ln 2\pi - \frac{1}{2} \ln \sigma^2 - \frac{(x-m_0)^2}{2\sigma^2} \right\} = \\ &= \exp \left\{ \frac{2m_0x - x^2}{2\sigma^2} - \frac{m_0^2}{2\sigma^2} - \frac{1}{2} \ln \sigma^2 - \frac{1}{2} \ln 2\pi \right\}. \end{aligned}$$

$T(X_1, \dots, X_n) = 2m_0 \sum_{i=1}^n X_i - \sum_{i=1}^n X_i^2$ este o statistică suficientă pentru σ^2 .

$$\begin{aligned} \text{c) } p(x; \theta) &= \frac{1}{\Gamma(k_0)} \theta^{k_0} x^{k_0-1} e^{-\theta x} = \\ &= \exp \{ -\theta x + k_0 \ln \theta + (k_0 - 1) \ln x - \ln \Gamma(k_0) \}. \end{aligned}$$

$T(X_1, \dots, X_n) = \sum_{i=1}^n X_i$ este o statistică suficientă pentru θ .

$$\begin{aligned} \text{d) } P(x; \theta) &= C_m^x \theta^x (1 - \theta)^{m-x} = \\ &= \exp \{x(\ln \theta - \ln(1 - \theta)) + m \ln(1 - \theta) + \ln C_m^x\}. \end{aligned}$$

$T(X_1, \dots, X_n) = \sum_{i=1}^n X_i$ este o statistică suficientă pentru θ .

$$\text{e) } P(x; \theta) = \frac{\theta^x}{x!} e^{-\theta} = \exp \{x \ln \theta - \theta - \ln x!\}.$$

$T(X_1, \dots, X_n) = \sum_{i=1}^n X_i$ este o statistică suficientă pentru θ .

Observație. În practică, pentru cazurile a), c), d), e) se lucrează cu $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$,

iar pentru cazul b) se lucrează cu $S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - m_0)^2 = \frac{1}{n} (nm_0^2 - T)$.

4.3. Fie X_1, \dots, X_n variabile independente, identic repartizate uniform pe $(0, \theta)$ cu $\theta > 0$. Să se arate că statistica $T = \max_{1 \leq i \leq n} X_i$ este suficientă pentru θ .

Soluție.

$$p(x_1, \dots, x_n; \theta) = \begin{cases} 1/\theta^n, & \text{dacă } x_i \in (0, \theta), i = 1, \dots, n \\ 0 & \text{în rest} \end{cases}$$

$$\text{Fie } f(a, b) = \begin{cases} 1 & \text{dacă } b \geq a \\ 0 & \text{dacă } b < a \end{cases} \text{ și fie } T = \max_{1 \leq i \leq n} X_i, Z = \min_{1 \leq i \leq n} X_i.$$

Atunci

$$p(x_1, \dots, x_n; \theta) = \frac{f(t, \theta)}{\theta^n} f(0, z), \text{ pentru } T = t, Z = z.$$

Suficiența statisticii T rezultă din criteriul de factorizare.

4.4. Fie X_1, \dots, X_n independente, urmînd o repartiție de tip exponențial cu parametru multidimensional:

$$p(x, \theta) = \exp \left\{ \sum_{k=1}^m a_k(\theta) b_k(x) + c(\theta) + d(x) \right\} \text{ cu } \theta = (\theta_1 \dots \theta_r) \in R^r.$$

Să se determine o statistică suficientă pentru această familie.

Să se particularizeze rezultatul în cazul normal $N(\mu, \sigma^2)$ cu ambii parametri necunoscuți și în cazul normal $N(\mu, \gamma_0^2 \mu^2)$ cu coeficient de variație cunoscut.

Soluție. $p(x_1, \dots, x_n; \theta) = \exp \left\{ \sum_{k=1}^m a_k(\theta) S_k(x_1, \dots, x_n) + nc(\theta) + \sum_{i=1}^n d(x_i) \right\}$, cu $S_k(x_1, \dots, x_n) = \sum_{j=1}^n b_k(x_j)$, $k = 1, \dots, m$.

O statistică suficientă pentru familia considerată este

$$S(X_1, \dots, X_n) = (S_1(X_1, \dots, X_n), \dots, S_m(X_1, \dots, X_n)).$$

Observăm că m și r nu coincid în mod obligatoriu. Dacă între componentele lui θ există o relație oarecare, se poate ca $m < r$. Cel mai frecvent este cazul $m = r$ (Aici se înscrie și modelul normal cu ambii parametri necunoscuți).

Cazul $N(\mu, \sigma^2)$. $p(x; \mu, \sigma^2) = \exp \left\{ \frac{\mu}{\sigma^2} x - \frac{1}{2\sigma^2} x^2 - \frac{\mu^2}{2\sigma^2} - \frac{1}{2} \ln \sigma^2 - \frac{1}{2} \ln 2\pi \right\}$. O statistică suficientă pentru $\theta = (\mu, \sigma^2)$ este $S = (S_1, S_2)$

$$\text{cu } S_1(X_1, \dots, X_n) = \sum_{i=1}^n X_i, \quad S_2(X_1, \dots, X_n) = \sum_{i=1}^n X_i^2.$$

Cazul $N(\mu, \gamma_0^2 \mu^2)$, cu γ_0 cunoscut:

$$p(x; \mu) = \exp \left\{ \frac{x}{\gamma_0^2 \mu} - \frac{x^2}{2\gamma_0^2 \mu^2} - \frac{1}{2\gamma_0^2} - \frac{1}{2} \ln (2\pi \gamma_0^2 \mu^2) \right\}.$$

O statistică suficientă pentru μ este $S = (S_1, S_2)$ cu $S_1(X_1, \dots, X_n) = \sum_{i=1}^n X_i$, $S_2(X_1, \dots, X_n) = \sum_{i=1}^n X_i^2$ (sintem în cazul $m > r$).

Observație. În cazul repartiției normale, în practică se utilizează statistica suficientă bidimensională (\bar{X}, S^2) cu $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$,

$$S^2 = \sum_{i=1}^n (X_i - \bar{X})^2.$$

4.5. Se consideră modelul general liniar normal:

$$Y_j = \sum_{k=1}^q x_{jk} \beta_k + \varepsilon_j \quad j = 1, \dots, n,$$

cu $X = (x_{jk})_{j=1, \dots, n; k=1, \dots, q}$ matrice de componente cunoscute, $\text{rg } X = q \leq n$, ε_j , $j = 1, \dots, n$ variabile aleatoare independente, identic repartizate $N(0, \sigma^2)$, $\beta = (\beta_1, \dots, \beta_q)^t$ și σ^2 parametri necunoscuți. Să se determine o statistică suficientă pentru $\theta = (\beta, \sigma^2)$.

$$\text{Soluție. } p(y_1, \dots, y_n; \beta, \sigma^2) = \frac{1}{(2\pi)^{n/2} \sigma^n} \exp \left\{ -\frac{1}{2\sigma^2} (y - X\beta)'(y - X\beta) \right\}.$$

Pentru parametrul β luăm în considerație estimatorul care realizează condiția $\min_{\beta} (y - X\beta)'(y - X\beta)$, respectiv care verifică sistemul

$$(X'X)\beta = X'y$$

$$\hat{\beta} = (X'X)^{-1}X'y.$$

Pentru parametrul σ^2 luăm în considerație statistica

$$S^2 = (y - X\hat{\beta})'(y - X\hat{\beta}).$$

Atunci:

$$p(y_1, \dots, y_n; \beta, \sigma^2) = \frac{1}{(2\pi)^{n/2} \sigma^n} \exp \left\{ -\frac{S^2}{2\sigma^2} - \frac{(\hat{\beta} - \beta)'X'X(\hat{\beta} - \beta)}{2\sigma^2} \right\}.$$

Cum densitatea de repartiție depinde de y_1, \dots, y_n doar prin intermediul lui $\hat{\beta}$ și S^2 , rezultă în baza teoremei de factorizare că $(\hat{\beta}, S^2)$ formează o statistică suficientă pentru (β, σ^2) .

4.6. Pentru modelul regresiei liniare, care este un caz particular al modelului general liniar normal, să se determine o statistică suficientă pentru parametrii modelului.

Fie deci $Y_j = \beta x_j + \varepsilon_j$, $j = 1, \dots, n$, cu x_1, \dots, x_n valori scalare cunoscute, $\varepsilon_1, \dots, \varepsilon_n$ variabile aleatoare independente, identic repartizate $N(0, \sigma^2)$, β și σ^2 parametri necunoscuți. Să se determine o statistică suficientă pentru (β, σ^2) .

$$\text{Soluție. } p(y_1, \dots, y_n; \beta, \sigma^2) = \frac{1}{(2\pi)^{n/2} \sigma^n} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{j=1}^n (y_j - \beta x_j)^2 \right\}.$$

$\hat{\beta}$ se obține din condiția $\min_{\beta} \sum_{j=1}^n (y_j - \beta x_j)^2$:

$$\hat{\beta} = \frac{\sum_{j=1}^n x_j y_j}{\sum_{j=1}^n x_j^2}.$$

$$S^2 = \sum_{j=1}^n (y_j - \hat{\beta} x_j)^2.$$

$$p(y_1, \dots, y_n; \beta, \sigma^2) = \frac{1}{(2\pi)^{n/2} \sigma^n} \exp \left\{ -\frac{1}{2\sigma^2} \left(\sum_{j=1}^n (y_j - \hat{\beta} x_j)^2 + (\hat{\beta} - \beta)^2 \sum_{j=1}^n x_j^2 \right) \right\}.$$

Din teorema de factorizare rezultă că $(\hat{\beta}, S^2)$ formează o statistică suficientă pentru (β, σ^2) .

4.7. Pentru următorul model de clasificare bifactorială (care este un caz particular al modelului liniar normal) să se determine o statistică suficientă pentru parametrii modelului. Fie deci

$$Y_{jk} = \mu + \alpha_j + \beta_k + \varepsilon_{jk}, \quad j = 1, \dots, m_1, \quad k = 1, \dots, m_2$$

cu ε_{jk} variabile aleatoare independente, identic repartizate, $N(0, \sigma^2)$, iar $\mu, \alpha = (\alpha_1, \dots, \alpha_{m_1}), \beta = (\beta_1, \dots, \beta_{m_2})$ și σ^2 parametri necunoscuți care verifică condițiile $\sum_{j=1}^{m_1} \alpha_j = \sum_{k=1}^{m_2} \beta_k = 0$. Să se determine o statistică suficientă pentru $(\mu, \alpha, \beta, \sigma^2)$.

Soluție. $p(y_{11}, \dots, y_{m_1 m_2}; \mu, \alpha, \beta, \sigma^2) =$

$$= \frac{1}{(2\pi)^{m_1 m_2 / 2} \sigma^{m_1 m_2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{j=1}^{m_1} \sum_{k=1}^{m_2} (y_{jk} - \mu - \alpha_j - \beta_k)^2 \right\}.$$

$\hat{\mu}, \hat{\alpha}$ și $\hat{\beta}$ se obțin din condiția

$$\min_{\mu, \alpha, \beta} \sum_{j=1}^{m_1} \sum_{k=1}^{m_2} (y_{jk} - \mu - \alpha_j - \beta_k)^2.$$

$$\hat{\mu} = \bar{y}_{..} = \frac{1}{m_1 m_2} \sum_j \sum_k y_{jk}$$

$$\hat{\alpha}_j = \bar{y}_{.j} = \frac{1}{m_2} \sum_k y_{jk}; \quad \hat{\beta}_k = \bar{y}_{.k} = \frac{1}{m_1} \sum_j y_{jk}.$$

Atunci $S^2 = \sum_j \sum_k (y_{jk} - \bar{y}_{.j} - \bar{y}_{.k} + \bar{y}_{..})^2$, iar

$$p(y_{11}, \dots, y_{m_1 m_2}; \mu, \alpha, \beta, \sigma^2) = \frac{1}{(2\pi)^{m_1 m_2 / 2} \sigma^{m_1 m_2}} \exp \left\{ -\frac{1}{2\sigma^2} \left(S^2 + m_2 \sum_j (\bar{y}_{.j} - \bar{y}_{..} - \alpha_j)^2 + m_1 \sum_k (\bar{y}_{.k} - \bar{y}_{..} - \beta_k)^2 + m_1 m_2 (\bar{y}_{..} - \mu)^2 \right) \right\}.$$

Din teorema de factorizare rezultă că

$$(\bar{Y}_{..}, (\bar{Y}_{1.}, \dots, \bar{Y}_{m.}), (\bar{Y}_{.1}, \dots, \bar{Y}_{.m}), S^2)$$

formează o statistică suficientă pentru $(\mu, \alpha, \beta, \sigma^2)$.

4.8. Se consideră următorul model de analiză dispersională unifactorială: $X_{jk} = \mu_j + \varepsilon$, $k = 1, \dots, r$, $j = 1, \dots, m$, cu X_{jk} variabile aleatoare independente, ε variabilă aleatoare repartizată $N(0, \sigma^2)$. Să se determine o statistică suficientă pentru parametrii necunoscuți $(\mu_1, \dots, \mu_m, \sigma^2)$, $\mu_j \in R$, $j = 1, \dots, m$, $\sigma^2 > 0$.

Soluție. $p(x_{11}, \dots, x_{mr}; \mu, \sigma^2) =$

$$= (2\pi\sigma^2)^{-rm/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{j=1}^m \sum_{k=1}^r (x_{jk} - \mu_j)^2\right\} =$$

$$= (2\pi\sigma^2)^{-rm/2} \exp\left\{-\frac{r}{2\sigma^2} \sum_{j=1}^m (\bar{x}_j - \mu_j)^2\right\} \cdot$$

$$\cdot \exp\left\{-\frac{1}{2\sigma^2} \sum_{j=1}^m \sum_{k=1}^r (x_{jk} - \bar{x}_j)^2\right\}, \text{ cu } \bar{x}_j = \frac{1}{r} \sum_{k=1}^r x_{jk}.$$

Fie $S^2 = \sum_j \sum_k (x_{jk} - \bar{x}_j)^2$. Atunci

$$p(x_{11}, \dots, x_{mr}; \mu, \sigma^2) =$$

$$= (2\pi\sigma^2)^{-rm/2} \exp\left\{-\frac{r}{2\sigma^2} \sum_j (\bar{x}_j - \mu_j)^2\right\} \exp\left\{-\frac{1}{2\sigma^2} S^2\right\}.$$

Din teorema de factorizare rezultă că $(\bar{X}_{1.}, \dots, \bar{X}_{m.}, S^2)$ este o statistică suficientă pentru parametrii modelului considerat.

4.9. Se consideră următorul model de analiză dispersională cu factori aleatori:

$$X_{jk} = \mu + \eta_j + \varepsilon_{jk}, \quad j = 1, \dots, m, \quad k = 1, \dots, r,$$

unde η_j , ε_{jk} sînt variabile aleatoare independente în totalitate, η_j identic repartizați $N(0, \sigma_1^2)$ ($j = 1, \dots, m$), ε_{jk} identic repartizați $N(0, \sigma_2^2)$ ($j = 1, \dots, m$, $k = 1, \dots, r$), iar μ , σ_1^2 și σ_2^2 sînt parametri necunoscuți. Fie $\bar{X}_j = \frac{1}{r} \sum_{k=1}^r X_{jk}$, $\bar{X}_{..} = \frac{1}{m} \sum_{j=1}^m \bar{X}_j$.

Să se arate că

$$(S_1, S_2, S_3) = \left(\bar{X}_{..}, \sum_{j=1}^m (\bar{X}_j - \bar{X}_{..})^2, \sum_{j=1}^m \sum_{k=1}^r (X_{jk} - \bar{X}_j)^2 \right)$$

este o statistică suficientă pentru $(\mu, \sigma_1^2, \sigma_2^2)$.

Soluție. Vectorii $X_j = (X_{j1}, \dots, X_{jr})^t$, $j = 1, \dots, m$ sînt independenți, identic repartizați $N(r; (\mu, \dots, \mu)^t, \Sigma)$ cu $\Sigma = \sigma_3^2 I + \sigma_1^2 J$, unde I este matricea unitate, iar J este matricea cu toate cele r^2 elemente egale cu unu. Atunci

$$\Sigma^{-1} = \frac{1}{\sigma_3^2} \left\{ I - \frac{\sigma_1^2}{\sigma_3^2 + r\sigma_1^2} J \right\}.$$

Notînd $\frac{\sigma_1^2}{\sigma_3^2 + r\sigma_1^2}$ cu λ , obținem $\Sigma^{-1} = \frac{1}{\sigma_3^2} (I - \lambda J)$.

Deci X_j , $j = 1, \dots, m$, au fiecare densitatea de repartiție

$$p(x_j; \mu, \Sigma) = \frac{1}{(2\pi)^{r/2} (\det \Sigma)^{1/2}} \exp \left\{ -\frac{1}{2} (x_j - \mu)^t \Sigma^{-1} (x_j - \mu) \right\}.$$

Atunci

$$\begin{aligned} \ln p(x_1, \dots, x_m; \mu, \Sigma) &= K - \frac{m}{2} \ln (\det \Sigma) - \\ &- \frac{1}{2} \sum_{j=1}^m (x_j - \mu)^t \Sigma^{-1} (x_j - \mu) = K - \frac{m}{2} \ln (\det \Sigma) - \\ &- \frac{1}{2\sigma_3^2} \left\{ \sum_j x_j^t x_j - 2\mu 1^t \sum_j x_j + m\mu^2 1^t \cdot 1 - \right. \\ &\left. - \lambda \left\{ \sum_j x_j^t J x_j - 2\mu 1^t J \sum_j x_j + \mu^2 1^t J 1 \right\} \right\}, \end{aligned}$$

unde am notat vectorul μ prin scalarul μ înmulțit cu vectorul 1 , $\mu = \mu \cdot 1$, iar $J = 1 \cdot 1^t$.

$$\begin{aligned} \ln p(x_1, \dots, x_m; \mu, \Sigma) &= K - \frac{m}{2} \ln (\det \Sigma) - \\ &- \frac{1}{2\sigma_3^2} \left\{ \sum_j \sum_k x_{jk}^2 - 2\mu r m \bar{x}_{..} + r m \mu^2 - \lambda \left(r \sum_j \bar{x}_j^2 - 2\mu r m \bar{x}_{..} + r^2 \mu^2 \right) \right\} = \\ &= K - \frac{m}{2} \ln (\det \Sigma) - \frac{1}{2\sigma_3^2} \{ S_3 + r S_2 (1 - \lambda) + \\ &+ r m S_1^2 (1 - \lambda) - 2\mu r m S_1 (1 - \lambda) + r m \mu^2 - \lambda r^2 \mu^2 \}. \end{aligned}$$

Deci $\ln p(x_1, \dots, x_m; \theta) = K - \frac{m}{2} \ln(\det \Sigma) - \frac{1}{2\sigma_1^2} (r m \mu^2 - \lambda \sigma^2 \mu^2) - \frac{1}{2\sigma_2^2} \{S_3 + r(1 - \lambda)(S_2 + mS_1^2 - 2\mu m S_1)\}$.

Din teorema de factorizare rezultă că (S_1, S_2, S_3) este o statistică suficientă pentru parametrul $\theta = (\mu, \sigma_1^2, \sigma_2^2)$.

4.10. Se consideră familia de probabilități $\{P_\theta(\cdot), \theta \in \Theta\}$ pe (E, \mathfrak{A}) . Să se demonstreze că dacă t este o statistică completă pentru θ , atunci orice statistică t^* care depinde de x doar prin intermediul lui t este de asemenea completă.

Soluție. Fie $t^* = g(t)$. Presupunem că există o funcție h care satisface condițiile: $h(t^*)$ măsurabilă,

$$M_\theta(h(t^*)) = 0 \text{ pentru orice } \theta \in \Theta.$$

Deci $h(g(t))$ este măsurabilă și $M_\theta(h(g(t))) = 0$ pentru orice $\theta \in \Theta$. Din completitudinea lui t rezultă $h(g(t)) = 0$, P_θ — aproape sigur pentru orice θ , de unde $h(t^*) = 0$, P_θ — aproape sigur pentru orice θ .

4.11. Să se demonstreze completitudinea următoarelor familii de repartiții de probabilitate:

- $\{P_\theta = B_n(\theta), \theta \in (0, 1)\}$ (binomială);
- $\{P_\theta = \pi_\theta, \theta \in (0, \infty)\}$ (Poisson);
- $\{P_\theta = N(\theta, 1), \theta \in (-\infty, \infty)\}$ (normală);
- $\{P_\theta = \gamma\left(\theta, \frac{1}{2}\right), \theta \in N\}$ (gama).

Soluție. a) $P_\theta(x) = C_n^x \theta^x (1 - \theta)^{n-x}$, $x = 0, 1, \dots, n$.
Fie $f(x)$ măsurabilă, așa încât

$$\sum_{x=0}^n f(x) C_n^x \theta^x (1 - \theta)^{n-x} \equiv 0 \text{ pentru orice } \theta \in (0, 1).$$

Notăm $a(x) = f(x) \cdot C_n^x$ și $\varphi = \frac{\theta}{1 - \theta}$. Avem atunci $\sum_{x=0}^n a(x) \varphi^x \equiv 0$.

Am obținut un polinom de gradul n în variabila φ , care este identic nul. Rezultă $a(x) = 0$ pentru $x = 0, 1, \dots, n$, deci $f(x) = 0$ pentru $x = 0, 1, \dots, n$.

b) $P_\theta(x) = c^{-\theta} \frac{\theta^x}{x!}$, $x = 0, 1, 2, \dots$

Fie $f(x)$ măsurabilă, care satisface relația

$$e^{-\theta} \sum_{x=0}^{\infty} f(x) \frac{\theta^x}{x!} = 0 \text{ pentru } \theta > 0.$$

Din proprietățile seriei de puteri rezultă $f(x) = 0$ a.s.

$$c) p_{\theta}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2}}, \quad x \in (-\infty, \infty).$$

Considerăm relația

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-\frac{(x-\theta)^2}{2}} dx = 0 \text{ pentru orice } \theta \in (-\infty, \infty).$$

Rezultă $\int_{-\infty}^{\infty} f(x) e^{-\frac{x^2}{2}} e^{\theta x} dx \equiv 0$. Deci transformata Laplace a funcției

$f(x) e^{-\frac{x^2}{2}}$ este identic nulă. Din unicitatea transformatei Laplace rezultă $f(x) e^{-\frac{x^2}{2}} = 0$ aproape sigur, de unde $f(x) = 0$ aproape sigur.

d) Rezultă din unicitatea transformatei Laplace.

4.12. Să se arate că familia repartițiilor uniforme pe $(\theta, \theta + 1)$, cu $\theta \in (-\infty, \infty)$ nu este completă.

$$\text{Soluție. Considerăm } f(x) = \begin{cases} 1, & n < x \leq \frac{1}{2} + n \\ -1 & \frac{1}{2} + n < x < n + 1, \quad n \in \mathbf{Z}. \end{cases}$$

Dacă $0 < n \leq \theta + \frac{1}{2}$, atunci

$$\begin{aligned} \int_0^{\theta+1} f(x) dx &= \int_0^n f(x) dx + \int_n^{n+\frac{1}{2}} f(x) dx + \int_{n+\frac{1}{2}}^{\theta+1} f(x) dx = \\ &= -(n - \theta) + \frac{1}{2} - \left(\theta + 1 - n - \frac{1}{2} \right) = 0. \end{aligned}$$

Dacă $\theta + \frac{1}{2} < n \leq \theta + 1$, obținem din nou $\int_0^{\theta+1} f(x) dx = 0$.

Deci $M_0(f(X)) = 0$ pentru orice $\theta \in \mathbf{R}$, iar $f \neq 0$.

4.13. Fie X o variabilă aleatoare cu densitatea de repartiție

$$p_{\theta}(x) = \begin{cases} (\theta + 1) x^{\theta} & 0 \leq x \leq 1 \\ 0 & \text{în rest} \end{cases}$$

cu parametrul $\theta \in N$. Să se arate că $t(x) = x$ este o statistică completă.

Soluție. Vom arăta că dacă $h(x)$ este o funcție măsurabilă pe $[0, 1]$, așa încît $M_\theta(h) = \int_0^1 h(x)(\theta + 1)x^\theta dx = 0$ pentru $\theta = 0, 1, 2, \dots$ atunci $h(x) = 0$ aproape sigur.

$$\text{Fie } H(x) = \int_0^x h(s) ds, \quad 0 \leq x \leq 1.$$

$$\text{Avem deci } \int_0^1 x^\theta dH(x) = 0, \quad \theta = 0, 1, 2, \dots$$

$$\text{Dar } \int_0^1 x^\theta dH(x) = \int_0^1 H(x) dx^\theta + H(1) \cdot 1 - H(0) \cdot 0.$$

Cum $H(1) = M_0(h) = 0$, rezultă

$$0 = \int_0^1 H(x) dx^\theta = \theta \int_0^1 H(x) x^{\theta-1} dx, \quad \theta = 1, 2, \dots$$

Fie $K(x) = \int_0^x H(s) ds$. Procedînd ca mai sus, obținem

$$\int_0^1 x^\theta K(x) dx = 0, \quad \theta = 0, 1, 2, \dots$$

Cum $H(s)$ este diferența a două funcții monotone, K este continuă pe $[0, 1]$.

Pentru orice polinom $p(x) = a_0 x^r + a_1 x^{r-1} + \dots + a_r$, avem

$$\int_0^1 p(x) K(x) dx = 0.$$

Cum $K(x)$ este continuă, pentru orice $\varepsilon > 0$ există un polinom $p(x)$ așa încît $|p(x) - K(x)| \leq \varepsilon$ pentru $x \in [0, 1]$.

$$\left| \int_0^1 K^2(x) dx \right| \leq \int_0^1 |K(x)| \cdot |K(x) - p(x)| dx \leq \varepsilon \int_0^1 |K(x)| dx < \infty.$$

Cum ε este arbitrar de mic, obținem $\int_0^{\varepsilon} K^2(x) dx = 0$, deci

$$K^2(x) = K(x) = 0 \text{ aproape sigur.}$$

Dar $K(x) = \int_0^x H(s) ds = 0$ a.s. implică $H(x) = \int_0^x h(s) ds = 0$ aproape sigur și deci $h(x) = 0$ a.s.

4.14. Să se arate că familia de probabilități induse de variabila aleatoare X definită prin densitățile de repartiție

$$p_{\theta}(x) = \begin{cases} \theta e^{-\theta x}, & x > 0 \\ 0 & x \leq 0, \theta \in N \end{cases}$$

este completă.

Soluție. Fie $g(x) = e^{-x}$. Densitatea de repartiție pentru g este

$$\tilde{p}_{\theta}(g) = \begin{cases} \theta g^{\theta-1}, & 0 \leq g \leq 1 \\ 0, & \text{în rest} \end{cases}$$

Din problema 4.4. știm că pentru familia de repartiții definită de $\{\tilde{p}_{\theta}(g), \theta \in N\}$ statistica $t(g) = g$ este completă.

Cum $x = -\log g$, folosind rezultatul din problema 4.1, obținem faptul că variabila aleatoare X este completă, deci familia considerată este completă.

4.15. Fie X_1, X_2, \dots, X_n independente, identic repartizate cu $P_m(X_1 = k) = \frac{1}{m}$, $k = 1, 2, \dots, m$, $m \in N^*$. Să se arate că $t = \max_{1 \leq i \leq n} X_i$ este completă.

Soluție. $P_m(t = y) = \frac{y^n}{m^n} - \frac{(y-1)^n}{m^n}$. Presupunem că este adevărată relația

$$\frac{1}{m^n} \sum_{y=1}^m f(y) [y^n - (y-1)^n] = 0 \text{ pentru } m = 1, 2, \dots$$

Pentru $m = 1$ obținem $f(1) = 0$.

Pentru $m = 2$ obținem $f(1) = 0$, $f(2) = 0$ etc.

Deci $f(y) = 0$ pentru $y = 1, 2, \dots, m$, oricare ar fi $m \in N^*$.

4.16. Fie X_1, \dots, X_n variabile aleatoare independente, identic repartizate, de tip exponențial cu parametru unidimensional:

$$p(x; \theta) = \exp \{a(\theta) b(x) + c(\theta) + d(x)\}, \quad \theta \in R.$$

Să se arate că statistica suficientă construită în problema 4.2,

$$S = \sum_{i=1}^n b(X_i)$$

este completă pentru familia considerată.

Soluție.

$$p(x_1, \dots, x_n, s; \theta) = \exp \left\{ a(\theta)s + nc(\theta) + \sum_{i=1}^n d(x_i) \right\},$$

cu $s = \sum_{i=1}^n b(X_i)$.

Atunci S are densitatea de repartiție

$$p(s; \theta) = \exp \{a(\theta)s + nc(\theta)\} \int_{R^n} \exp \left(\sum_{i=1}^n d(x_i) \right) dx_1 \dots dx_n.$$

Fie f cu proprietatea că $M_0(f(S)) = 0$, adică

$$\exp(nc(\theta)) \int_{R^n} \exp \left(\sum_{i=1}^n d(x_i) \right) dx_1 \dots dx_n \int_R f(s) \exp(s a(\theta)) ds = 0$$

Rezultă că $f(s) = 0$ aproape sigur pentru orice θ , deci S este completă.

4.17. Fie X_1, \dots, X_n variabile independente, identic repartizate normal, cu coeficient de variație cunoscut $N(\mu, \gamma_0^2 \mu^2)$. Să se arate

că statistica suficientă $S = \left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2 \right)$ construită în problema

4.4 nu este completă.

Soluție.

$$M_\mu \left(\frac{n + \gamma_0^2}{1 + \gamma_0^2} \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2 \right) = 0,$$

iar $f(S) = \frac{n + \gamma_0^2}{1 + \gamma_0^2} \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2$ este diferită de zero.

Lehman a demonstrat că o statistică S este suficientă pentru o familie de tip exponențial cu parametrul θ dacă și numai dacă $\dim S = \dim \theta$.

§2. Selecție, concordanța repartiției de selecție cu o repartiție teoretică

4.18. Se fac 100 observații independente asupra caracteristicii aleatoare X , valorile obținute fiind date în tabelul următor:

Valori observate x_j	0,20	0,34	0,38	0,40	0,48	0,50	0,55	0,59	0,60	0,70
Frecvențe absolute n_j	9	11	8	10	12	10	9	10	11	10

Se face ipoteza că X urmează repartiția uniformă pe intervalul $(0,1)$. Să se verifice această ipoteză utilizând testul lui Kolmogorov.

Soluție. Funcția de repartiție teoretică este $F(x) = x$ pentru $x \in (0, 1)$.

Elementele de calcul necesare aplicării testului lui Kolmogorov sînt prezentate în tabelul următor:

x_i	0,20	0,34	0,38	0,40	0,48	0,50	0,55	0,59	0,60	0,70
$F(x_i)$	0,20	0,34	0,38	0,40	0,48	0,50	0,55	0,59	0,60	0,70
$F_n(x_i)$	0,09	0,20	0,28	0,38	0,50	0,60	0,69	0,79	0,90	1,00
$ F_n(x_i) - F(x_i) $	0,11	0,14	0,10	0,02	0,02	0,10	0,14	0,20	0,30	0,30

$$K_n = \sqrt{n} \max_{1 \leq i \leq n} |F_n(x_i) - F(x_i)| = 10 \cdot 0,30 = 3.$$

Pentru $\alpha = 0,05$, cuantila de rang 0,95 a repartiției lui Kolmogorov este $k_{0,95} = 3,81$.

Cum $K^n < k_{0,95}$, decidem să acceptăm ipoteza că X urmează repartiția $U_{0,1}$, la pragul de semnificație $\alpha = 0,05$.

4.19. Se fac 15 observații independente asupra unei caracteristici aleatoare X , valorile observate fiind date în tabelul de mai jos:

Valori observate x_i	16	19	22	23	24	26	27	30	32	33	36	39
Frecvențe absolute n_i	1	2	2	1	2	1	1	1	1	1	1	1

Să se verifice normalitatea datelor, utilizând testul Lilliefors.

Soluție. Estimăm întâi parametrii repartiției normale prin

$$\hat{m} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = 26,13,$$

$$\hat{\sigma} = s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = 41,18, \quad s \simeq 6,4.$$

Facem deci ipoteza că datele observate urmează o repartiție $N(26,13; 41,18)$. Calculele necesare aplicării testului Lilliefors sînt date în tabelul următor, în care $\mathcal{O}(z_i)$ reprezintă valoarea funcției de repartiție normale standard $N(0, 1)$ în punctul z_i .

x_i	$x_i - \bar{x}$	$z_i = \frac{x_i - \bar{x}}{s}$	$\mathcal{O}(z_i)$	$F_n(z_i)$	$ F_n(z_i) - \mathcal{O}(z_i) $
16	-10,13	-0,58	0,2810	0,0667	0,2143
19	-7,13	-0,47	0,3192	0,2000	0,1192
22	-4,13	-0,27	0,3936	0,3333	0,0603
23	-3,13	-0,21	0,4168	0,4000	0,0168
24	-2,13	-0,14	0,4443	0,5333	0,0890
26	-0,13	0	0,5000	0,6000	0,1000
27	0,87	0,06	0,5239	0,6667	0,1428
30	3,87	0,26	0,6026	0,7333	0,1307
32	5,87	0,39	0,6517	0,8000	0,1483
33	6,87	0,49	0,6879	0,8667	0,1788
36	9,87	0,66	0,7454	0,9333	0,1879
39	12,87	0,86	0,8051	1,0000	0,1949

$$D_{max} = \max_{1 \leq i \leq 15} |F_n(z_i) - \mathcal{O}(z_i)| = 0,2143.$$

Utilizând Anexa IV găsim $D_{15; 0,95} = 0,220$.
 Cum $D_{max} < D_{15; 0,95}$ decidem să acceptăm ipoteza de normalitate
 a datelor, la pragul de semnificație $\alpha = 0,05$.

4.20. Se fac 100 observații independente asupra unei caracteris-
 tici aleatoare X , valorile observate fiind date în tabelul de mai jos :

Valori observate x_i	16	19	22	23	24	27	30	32	33	36
Frecvențe abso- lute n_i	1	9	18	10	15	18	12	14	2	1

Să se verifice utilizând testul χ^2 concordanța dintre repartiția de
 selecție a datelor și repartiția $N(26; 49)$.

Soluție. Stabilim numărul intervalelor de grupare a datelor utili-
 zînd recomandarea STAS 7122-65.

$$k \simeq 1 + 3,322 \lg n \simeq 8.$$

Calculule necesare aplicării testului χ^2 sînt date în tabelul de mai
 jos :

Intervale pt. X (a_j, a_{j+1}]	Frecv. abso- lută pt. (a_j, a_{j+1}]	Intervale pt. $t = \frac{X-26}{7}$ (b_j, b_{j+1}]	$p_j = \Phi(b_{j+1}) - \Phi(b_j)$	$\frac{(n_j - np_j)^2}{np_j}$
16 < X < 19	10	-1,43 < t < -1	0,1587 - 0,0764 = 0,0823	0,38
19 < X < 22	8	-1 < t < -0,57	0,2843 - 0,1587 = 0,1256	2,35
22 < X < 23	10	-0,57 < t < -0,43	0,3336 - 0,2843 = 0,0493	5,21
23 < X < 24	15	-0,43 < t < -0,28	0,3897 - 0,3336 = 0,0561	15,72
24 < X < 27	18	-0,28 < t < 0,14	0,5557 - 0,3897 = 0,166	0,12
27 < X < 30	12	0,14 < t < 0,57	0,7157 - 0,5557 = 0,16	1
30 < X < 32	14	0,57 < t < 0,86	0,8051 - 0,7157 = 0,0894	2,86
32 < X < 36	3	0,86 < t < 1,43	0,9326 - 0,8051 = 0,1185	6,61

$$H_n^2 = \sum_{j=1}^8 \frac{(n_j - np_j)^2}{np_j} = 34,25$$

Numărul gradelor de libertate este $k - 1 = 7$. Din Anexa II determinăm cuantila $h_{7; 0,95}^2 = 14,067$.

Cum $H_n^2 > h_{7; 0,95}^2$, decidem să respingem ipoteza de normalitate $N(26; 49)$ a datelor considerate, pentru $\alpha = 0,05$.

4.21. Într-o operație de control al funcționării unei mașini automate se prelevează la intervale de o jumătate de oră câte 20 de piese care sînt verificate. După prelevarea a 100 eşantioane de control, datele obținute se prezintă astfel:

Nr. piese defecte x_i	0	1	2	3	4	5	6	8
Nr. de eşantioane n_i	14	25	27	23	7	2	1	1

Utilizînd testul χ^2 să se verifice concordanța repartiției de selecție cu o repartiție binomială.

Soluție. Cum volumul fiecărui eşantion a fost $n = 20$, iar valoarea medie a variabilei „număr de piese defecte” este $\bar{x} = \frac{1}{100} \sum_{i=1}^{100} x_i = 2$, rezultă valoarea estimată a parametrului p egală cu $\tilde{p} = \frac{\bar{x}}{n} = \frac{2}{20} = 0,1$.

Vom face ipoteza că datele considerate urmează o repartiție $B_{20}(0,1)$.

Calculule necesare aplicării testului χ^2 sînt date în tabelul următor:

Intervale pt. X $(a_j, a_{j+1}]$	n_j	p_j	$\frac{(n_j - np_j)^2}{np_j}$
$X < 0$	14	0,122	0,27
$0 < X < 1$	25	0,270	0,15
$1 < X < 2$	27	0,285	0,08
$2 < X < 3$	23	0,190	0,84
$3 < X < 4$	7	0,090	0,44
$4 < X < 8$	4	0,043	0,20

$$H_n^2 = \sum_{j=1}^6 \frac{(n_j - np_j)^2}{np_j} = 1,80.$$

Numărul gradelor de libertate este $k - 2 = 4$

Pentru $\alpha = 0,05$ determinăm din anexa II cuantila $h_4^2; 0,95 = 9,488$.

Cum $H_n^2 < h_4^2; 0,95$, decidem să acceptăm ipoteza că variabila „număr de piese defecte” urmează repartiția $B_{20}(0,1)$.

4.22. Se fac 100 observații independente asupra unei caracteristici aleatoare X , valorile observate fiind date în tabelul de mai jos:

Valori observate x_i	85	86	87	88	89	90	91	92	93	94
Frecvențe absolute n_i	1	2	6	14	21	23	16	10	5	2

Să se verifice normalitatea datelor, utilizând testul χ^2

Soluție. Cum repartiția teoretică este incomplet precizată (normală, cu parametri neprecizați), vom construi întâi estimatori nedepășați pentru cei doi parametri

$$\hat{m} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = 89,79$$

$$\hat{\sigma}^2 = s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = 3,166; \quad s = 1,78.$$

Vom verifica deci ipoteza că X urmează o repartiție $N(89,79; 3,166)$.

Conform recomandării STASS 7122-65, numărul de intervale de grupare este

$$k \simeq 1 + 3,322 \lg n \simeq 8.$$

Calculule necesare aplicării testului χ^2 sînt date în tabelul următor:

Intervale pt. X $(a_j, a_{j+1}]$	Frecv. absolută pt. $(a_j, a_{j+1}]$	Intervale pt. $t = \frac{X - \bar{x}}{s}$ $(b_j, b_{j+1}]$	$p_j =$ $\Phi(b_{j+1}) - \Phi(b_j)$	$\frac{(n_j - np_j)^2}{np_j}$
$85 < X < 86$	3	$-2,7 < t < -2,13$	0,032	0,01
$86 < X < 87$	6	$-2,13 < t < -1,57$	0,067	0,07
$87 < X < 88$	14	$-1,57 < t < -1$	0,134	0,03
$88 < X < 89$	21	$-1 < t < -0,44$	0,200	0,05
$89 < X < 90$	23	$-0,44 < t < 0,12$	0,219	0,06
$90 < X < 91$	16	$0,12 < t < 0,67$	0,177	0,17
$91 < X < 92$	10	$0,67 < t < 1,24$	0,107	0,05
$92 < X < 94$	7	$1,24 < t < 2,36$	0,064	0,06

$$H_n^2 = \sum_{j=1}^8 \frac{(n_j - np_j)^2}{np_j} = 0,50.$$

Numărul gradelor de libertate este $k - 3 = 5$. Din anexa II determinăm cuantila $h_{5; 0,95}^2 = 11,070$ pentru $\alpha = 0,05$.

Cum $H_n^2 < h_{5; 0,95}^2$, decidem să acceptăm ipoteza de normalitate a datelor, la pragul de semnificație $\alpha = 0,05$.

§ 3. Estimarea parametrilor

4.23. Fie X_1, \dots, X_n variabile aleatoare independente, identic repartizate uniform pe mulțimea $\{1, \dots, m\}$. Utilizând teorema Rao—Blackwell, să se găsească un estimator nedeplasat, de dispersie minimă pentru parametrul m , $m \in N^*$.

Soluție. $P_m(X_1 = k) = \frac{1}{m}$, $k = 1, \dots, m$ $M_m(X_1) = \sum_{k=1}^m \frac{k}{m} = \frac{m+1}{2}$

Observăm că $M_m(2X_1 - 1) = m$, deci $\tilde{g}(X_1) = 2X_1 - 1$ este un estimator nedeplasat pentru m .

$\max_{1 \leq i \leq n} X_i$ este o statistică suficientă și completă pentru m .

$$P_m(t = y) = \frac{y^n}{m^n} - \frac{(y-1)^n}{m^n}$$

$$P_m(X_1 = x_1, \dots, X_n = x_n | t = y) = \frac{1}{y^n - (y-1)^n}$$

$$P_m(X_1 = x_1 | t = y) = \begin{cases} \frac{y^{n-1} - (y-1)^{n-1}}{y^n - (y-1)^n}, & x_1 = 1, \dots, (y-1) \\ \frac{y^{n-1}}{y^n - (y-1)^n}, & x_1 = y \end{cases}$$

$$M_m(\tilde{g} | t = y) = \frac{y^{n-1} - (y-1)^{n-1}}{y^n - (y-1)^n} \sum_{x=1}^{y-1} (2x-1) + \frac{y^{n-1}}{y^n - (y-1)^n} (2y-1) = \frac{y^{n+1} - (y-1)^{n+1}}{y^n - (y-1)^n}.$$

Deci

$$h(t(X)) = \frac{(t(X))^{n+1} - (t(X) - 1)^{n+1}}{(t(X))^n - (t(X) - 1)^n}$$

este o estimatie nedepasata, de dispersie minima a lui m .

4.24. Fie X_1, \dots, X_n variabile aleatoare independente, identic repartizate gama $\gamma(2, \theta)$. Sa se determine, utilizand teorema Rao-Blackwell o estimatie nedepasata, de dispersie minima pentru θ .

Solutie. $p(x, \theta) = \theta^2 x e^{-\theta x}$, $x > 0$, $\theta > 0$.

$$p(x_1, \dots, x_n; \theta) = \theta^{2n} \exp \left\{ -\theta \sum_{i=1}^n x_i \right\} \prod_{i=1}^n x_i.$$

$t(X_1, \dots, X_n) = \sum_{i=1}^n X_i$ este o statistica suficienta pentru θ .

$\tilde{g}(X_1, \dots, X_n) = \frac{1}{X_1}$ este un estimator nedepasat pentru θ , caci

$$M_{\theta}(\tilde{g}) = \theta.$$

In virtutea teoremei Rao-Blackwell, un estimator nedepasat de dispersie minima este $M_{\theta}(\tilde{g}|t)$. Pentru a-l evalua, trebuie sa determinam densitatea de repartitie conditionata $p_{\theta}(x_1|t)$.

Consideram urmatoarea transformare de coordonate, cu jacobianul egal cu unu:

$$\begin{cases} x_i = x_i, & i = 1, \dots, n-1 \\ t = \sum_{i=1}^n x_i \end{cases}$$

Atunci $p_{\theta}(x_1, \dots, x_{n-1}, t) = \theta^{2n} e^{-\theta t} \left(t - \sum_{i=1}^{n-1} x_i \right) \prod_{i=1}^{n-1} x_i$, de unde rezultă

$$p_{\theta}(x_1, t) = \theta^{2n} e^{-\theta t} x_1 (t - x_1)^{2n-3} \cdot \frac{1}{(2n-3)!} \text{ pentru } t > 0, 0 \leq x_1 \leq t.$$

$$p_{\theta}(x_1|t) = \frac{p_{\theta}(x_1, t)}{p_{\theta}(t)} = (2n-1)(2n-2) \frac{x_1}{t^2} \left(1 - \frac{x_1}{t} \right)^{2n-3}, 0 \leq x_1 \leq t$$

Atunci estimatorul cautat este

$$M_{\theta} \left(\frac{1}{X_1} \middle| t \right) = \int_0^t \frac{1}{x_1} (2n-1)(2n-2) \frac{x_1}{t^2} \left(1 - \frac{x_1}{t} \right)^{2n-3} dx_1 = \frac{2n-1}{t}.$$

Deci $(2n - 1) / \sum_{i=1}^n X_i$ este o estimatie nedeplasata, de dispersie minima. Dispersia acestei estimatii este

$$D_0^2 \left(\frac{2n - 1}{\sum_{i=1}^n X_i} \right) = \frac{\theta^2}{2(n - 1)}.$$

Pe de alta parte, informatia Fischer este

$$I(\theta) = nM(U^2(\theta)) = \frac{2n}{\theta^2}, \text{ unde } U(\theta) = \frac{\partial \ln p(X, \theta)}{\partial \theta} = \frac{2}{\theta} - X.$$

Atunci marginea inferioara a dispersiei unui estimator nedeplasat este $1/I(\theta) = \theta^2/2n$. Deci

$$D_0^2 \left(\frac{2n - 1}{\sum_{i=1}^n X_i} \right) > \frac{1}{I(\theta)}.$$

Pentru n suficient de mare, estimatorul construit este aproximativ eficient, dar pentru $n = 2$ eficienta sa este de 0,5.

4.25. Fie X_1, \dots, X_n variabile aleatoare independente, identic repartizate exponential, cu densitatea de repartitie a lui X_i ($i = 1, \dots, n$) data sub una din formele:

$$a) p(x, \mu) = \frac{1}{\mu} e^{-x/\mu}, \quad x \geq 0, \quad \mu > 0$$

$$b) p(x, \rho) = \rho e^{-\rho x}, \quad x \geq 0, \quad \rho > 0.$$

Sa se determine o estimatie nedeplasata, de dispersie minima pentru parametrul repartitiei considerate.

Solutie.

$$a) U(\mu) = \frac{\partial \ln p(X, \mu)}{\partial \mu} = -\frac{1}{\mu} + \frac{X}{\mu^2}$$

$$I(\mu) = nM_\mu(U^2(\mu)) = \frac{n}{\mu^3}.$$

Deci marginea inferioara a dispersiei unui estimator nedeplasat pentru μ este $\frac{\mu^3}{n}$.

$$p(x_1, \dots, x_n; \mu) = \frac{1}{\mu^n} \exp \left\{ -\frac{1}{\mu} \sum_{i=1}^n x_i \right\} = \frac{1}{\mu^n} \exp \left\{ -\frac{n}{\mu} \bar{x} \right\}.$$

$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$; este o statistică suficientă pentru μ , iar

$$M_{\mu}(\bar{X}) = \mu, \quad D_{\mu}^2(\bar{X}) = \frac{D_{\mu}^2(X)}{n} = \frac{\mu^2}{n}.$$

Deci \bar{X} este estimatorul nedepasat care realizează egalitate în inegalitatea Rao — Cramer (estimator eficient).

$$b) \quad U(\rho) = \frac{\partial \ln p(X, \rho)}{\partial \rho} = \frac{1}{\rho} - X$$

$$I(\rho) = nM(U^2(\rho)) = \frac{n}{\rho^2}.$$

Deci marginea inferioară a dispersiei unui estimator nedepasat pentru ρ este ρ^2/n .

Vom studia estimațiile de forma $c / \sum_{j=1}^n X_j$, constanta c urmînd să fie aleasă așa încît estimația să fie nedepasată.

Variabila aleatoare $S = \sum_{j=1}^n X_j$ are densitatea de repartiție $p(s; \rho) = \frac{\rho(\rho s)^{n-1} e^{-\rho s}}{(n-1)!}$. Atunci $M_{\rho}(1/S) = \rho/(n-1)$. Vom alege deci $c = n-1$, pentru ca să obținem $M_{\rho}\left(\frac{n-1}{\sum_{j=1}^n X_j}\right) = \rho$. Avem și $D_{\rho}^2\left(\frac{n-1}{\sum_{j=1}^n X_j}\right) = \frac{\rho^2}{n-2} > \frac{1}{I(\rho)}$.

Am obținut deci $S = \sum_{j=1}^n X_j$, o statistică suficientă pentru ρ și $t = \frac{n-1}{\sum_{j=1}^n X_j}$ o estimație nedepasată pentru ρ . În virtutea teoremei

Rao—Blackwell, t este chiar estimatorul de dispersie minimă. Pentru n suficient de mare t este aproximativ eficient, dar pentru $n = 3$ eficiența sa este doar de 0,33.

4.26. Fie X_1, \dots, X_n independente, identic repartizate Poisson de parametru $\mu(\pi_{\mu})$. Să se determine un estimator nedepasat, de dispersie minimă pentru $\theta = e^{-\mu}$.

Soluție. O primă modalitate de estimare a lui θ ne este oferită de estimația nedepășată, de dispersie minimă pentru μ care era $\bar{X} = \frac{1}{n} \sum_{j=1}^n X_j$, cu $M_{\mu}(\bar{X}) = \mu$, $D_{\mu}^2(\bar{X}) = \mu^2/n$.

Dar $e^{-\bar{X}}$ este un estimator deplasat al lui θ . Într-adevăr, dacă $(D_{\mu}^2(\bar{X}))^{1/2}$ nu este prea mare, putem considera dezvoltarea Taylor $g(\bar{X}) \simeq g(\mu) + (\bar{X} - \mu)g'(\mu) + \frac{1}{2}(\bar{X} - \mu)^2g''(\mu)$, unde $g(\mu) = e^{-\mu}$.

Mediind, obținem:

$$M_{\mu}(g(\bar{X})) \simeq g(\mu) + \frac{1}{2} D_{\mu}^2(\bar{X}) g''(\mu), \text{ sau } M_{\mu}(e^{-\bar{X}}) \simeq e^{-\mu} \left(1 + \frac{\mu^2}{2n} \right).$$

Deci $e^{-\bar{X}}$ nu este estimatorul căutat. Pentru rezolvarea problemei vom face din nou apel la teorema Rao-Blackwell.

$$P(x; \mu) = \frac{\mu^x}{x!} e^{-\mu} = \theta \frac{1}{x!} \left(\ln \frac{1}{\theta} \right)^x$$

$$P(x_1, \dots, x_n; \theta) = \theta^n \frac{1}{\sum_{j=1}^n x_j!} \left(\ln \frac{1}{\theta} \right)^{\sum_{j=1}^n x_j}$$

$S = \sum_{j=1}^n X_j$ este o statistică suficientă pentru θ . Ea este repartizată Poisson de parametru $n\mu$ ($\pi_{n\mu}$).

Considerăm acum o estimație nedepășată a lui θ :

$$t = \begin{cases} 1, & \text{dacă } X_1 = 0 \\ 0, & \text{dacă } X_1 \neq 0. \end{cases} \quad M_{\theta}(t) = \theta.$$

Repartiția lui t condiționată de $S = s$ este binomială $B\left(s, \frac{1}{n}\right)$.

$$\text{Atunci } M(t|S) = P(X_1 = 0|S) = \left(1 - \frac{1}{n}\right)^S = \left(1 - \frac{1}{n}\right)^{n\bar{X}}$$

este o estimație nedepășată, de dispersie minimă pentru θ .

4.27. Se alege o selecție n -dimensională dintr-o populație împărțită în trei clase. Probabilitățile teoretice ale celor trei clase sînt $\theta_1, \theta_2, 1 - \theta_1 - \theta_2$. Să se determine matricea informațională Fischer corespunzătoare și estimatorii eficienți ai parametrilor θ_1 și θ_2 .

Soluție. Notăm parametrul bidimensional $\theta = (\theta_1, \theta_2)$, $\theta_1 \in (0,1)$, $\theta_2 \in (0,1)$ și $x = (x_1, x_2, x_3)$ vectorul ce dă numărul observațiilor din cele trei clase, $x_1 + x_2 + x_3 = n$.

$$P(x; \theta) = \frac{n!}{x_1! x_2! x_3!} \theta_1^{x_1} \theta_2^{x_2} (1 - \theta_1 - \theta_2)^{x_3}.$$

Deci $X = (X_1, X_2, X_3)$ urmează o repartiție multinomială și $M(X_1) = n\theta_1$, $M(X_2) = n\theta_2$, $M(X_3) = n(1 - \theta_1 - \theta_2)$.

$$I_{11}(\theta) = M_\theta \left(-\frac{\partial^2 \ln P(X; \theta)}{\partial \theta_1^2} \right) = n \left(\frac{1}{\theta_1} + \frac{1}{1 - \theta_1 - \theta_2} \right)$$

$$I_{12}(\theta) = I_{21}(\theta) = M_\theta \left(-\frac{\partial^2 \ln P(X; \theta)}{\partial \theta_1 \partial \theta_2} \right) = \frac{n}{1 - \theta_1 - \theta_2}$$

$$I_{22}(\theta) = M_\theta \left(-\frac{\partial^2 \ln P(X; \theta)}{\partial \theta_2^2} \right) = n \left(\frac{1}{\theta_2} + \frac{1}{1 - \theta_1 - \theta_2} \right)$$

$$I(\theta) = n \begin{pmatrix} \frac{1}{\theta_1} + \frac{1}{1 - \theta_1 - \theta_2} & \frac{1}{1 - \theta_1 - \theta_2} \\ \frac{1}{1 - \theta_1 - \theta_2} & \frac{1}{\theta_2} + \frac{1}{1 - \theta_1 - \theta_2} \end{pmatrix}$$

$$(I(\theta))^{-1} = \frac{1}{n} \begin{pmatrix} \theta_1(1 - \theta_1) & -\theta_1\theta_2 \\ -\theta_1\theta_2 & \theta_2(1 - \theta_2) \end{pmatrix}$$

Estimatorii eficienți ai parametrilor θ_1, θ_2 sînt:

$$t_1(X) = X_1/n, \quad M_0(t_1) = \theta_1, \quad D_0^2(t_1) = \theta_1(1 - \theta_1)/n$$

$$t_2(X) = X_2/n, \quad M_0(t_2) = \theta_2, \quad D_0^2(t_2) = \theta_2(1 - \theta_2)/n.$$

4.28. Să se construiască estimatorii nedepasați, de dispersie minimă pentru parametrii repartiției normale $N(\mu, \tau)$, $\mu \in \mathbb{R}$, $\tau \in (0, \infty)$ și să se evalueze matricea informațională Fischer.

Soluție. Fie X_1, \dots, X_n variabile independente, identic repartizate $N(\mu, \tau)$ S-a arătat în problema 4.4 că $\left(\bar{X}, \sum_{i=1}^n (X_i - \bar{X})^2 \right)$ formează o statistică suficientă pentru (μ, τ) . Pe baza acestei statistici construim estimatorii nedepasați pentru μ și τ dați de \bar{X} și $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$. Datorită completitudinii familiei repartițiilor normale, \bar{X} și S^2 sînt estimatori de dispersie minimă, pentru $\theta = (\mu, \tau)$.

Matricea informațională Fisher este

$$I(\theta) = ||M(-U_{ij}(\theta))||_{i,j=1,2}, \text{ unde}$$

$$U_{ij}(\theta) = \frac{\partial^2 \ln p(X_1, \dots, X_n; \theta)}{\partial \theta_i \partial \theta_j} \text{ cu } \theta_1 = \mu, \theta_2 = \tau, \text{ iar}$$

$$\ln p(X_1, \dots, X_n; \theta) = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \tau - \frac{1}{2\tau} \sum_{i=1}^n (X_i - \mu)^2.$$

Dintr-un calcul simplu rezultă

$$I(\theta) = \begin{pmatrix} \frac{n}{\tau} & 0 \\ 0 & \frac{n}{2\tau^2} \end{pmatrix}.$$

Atunci

$$I^{-1}(\theta) = \begin{pmatrix} \tau & 0 \\ 0 & \frac{2\tau^2}{n} \end{pmatrix}.$$

Observăm că $D^2(\bar{X}) = \frac{\tau}{n}$, deci \bar{X} este un estimator eficient, dar $D^2(S^2) > \frac{2\tau^2}{n}$.

4.29. Să se determine o estimatie nedeplasată pentru $\theta = \log \sigma$ în cazul repartiției normale $N(\mu, \sigma^2)$ cu μ și σ^2 necunoscute.

Soluție. Fie X_1, \dots, X_n variabile independente, identic repartizate $N(\mu, \sigma^2)$. $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ este o estimatie nedeplasată pentru σ^2 . $\frac{(n-1)S^2}{\sigma^2}$ este repartizată χ^2 cu $(n-1)$ grade de libertate.

$$\text{Atunci } M\left(\frac{1}{2} \log S^2\right) = \log \sigma - \frac{1}{2} \log(n-1) +$$

$$+ \int_0^{\infty} \left(\frac{1}{2} \log x\right) \frac{\frac{1}{2} \left(\frac{1}{2} x\right)^{\frac{n-1}{2}-1} e^{-x/2}}{\Gamma\left(\frac{n-1}{2}\right)} dx = \log \sigma - \frac{1}{2} \log(n-1) + \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}.$$

Deci o estimatie nedepasata pentru $\theta = \log \sigma$ este

$$t = \frac{1}{2} \log S^2 + \frac{1}{2} \log (n-1) - \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}.$$

Funcția $\Psi(m) = \frac{\Gamma'(m)}{\Gamma(m)}$ are valorile tabelate.

4.30. Fie X_1, \dots, X_n variabile independente, identic repartizate exponențial, cu densitatea de repartiție

$$p(x; \rho) = \rho e^{-\rho x}, \quad x \geq 0, \quad \rho > 0.$$

Pentru $z \in R_+$ fixat, notăm prin $F_\rho(z)$ valoarea funcției de repartiție $F_\rho(z) = P_\rho(X < z)$. Să se arate că estimatia nedepasată, de dispersie minimă pentru $F_\rho(z)$ este

$$t = 1 - \left\{ 1 - \frac{\min(z, S)}{S} \right\}^n \text{ cu } S = \sum_{i=1}^n X_i.$$

Soluție. $S = \sum_{j=1}^n X_j$ este o statistică suficientă pentru ρ .

$V = \begin{cases} 1, & \text{dacă } X_1 < z \\ 0, & \text{dacă } X_1 \geq z \end{cases}$ este o estimatie nedepasată a lui $F_\rho(z)$.

Atunci, conform teoremei Rao-Blackwell, $t = M(V|S)$ este o estimatie nedepasată, de dispersie minimă pentru $F_\rho(z)$.

Densitatea de repartiție condiționată este:

$$\begin{aligned} p(x_1|s; \rho) &= \frac{p_{X_1}(x_1; \rho) p_{S-X_1}(s-x_1; \rho)}{p_S(s; \rho)} = \\ &= \frac{\rho e^{-\rho x_1} \rho^n (s-x_1)^{n-1} e^{-\rho(s-x_1)}}{\rho^{n+1} s^n e^{-\rho s}} = \frac{(n-1)!}{n!} \left(1 - \frac{x_1}{s}\right)^{n-1} \text{ pentru } 0 \leq x_1 \leq s. \end{aligned}$$

Atunci

$$t = M(V|S) = P(V=1|S) = \begin{cases} 1 - \left(1 - \frac{z}{s}\right)^n, & z < S \\ 1, & z \geq S \end{cases}$$

$$\text{adică } t = 1 - \left\{ 1 - \frac{\min(z, S)}{S} \right\}^n.$$

4.31. Fie X_1, \dots, X_n variabile independente, cu aceeași repartiție de tip exponențial:

$$p(x, \theta) = \exp \{a(\theta) b(x) + c(\theta) + d(x)\}, \quad \theta \in R.$$

Să se arate că $t = \frac{1}{n} \sum_{j=1}^n b(X_j)$ este o estimatie de dispersie minimă pentru θ .

Soluție.
$$U(\theta) = \frac{\partial \ln p(X_1, \dots, X_n; \theta)}{\partial \theta}.$$

Dar
$$p(x_1, \dots, x_n; \theta) = \exp \left\{ a(\theta) \sum_{j=1}^n b(x_j) + nc(\theta) + \sum_{j=1}^n d(x_j) \right\}.$$

Atunci
$$U(\theta) = a'(\theta) \sum_{j=1}^n b(X_j) + nc'(\theta) = na'(\theta)t + nc'(\theta)$$

$$t = \frac{1}{na'(\theta)} U(\theta) - \frac{c'(\theta)}{a'(\theta)}.$$

t apare deci ca o funcție liniară de $U(\theta)$, deci pentru t se obține egalitate în inegalitatea Rao-Cramer. Nu putem însă preciza dacă t este o estimatie nedepasată sau nu pentru θ .

4.32. Fie X_1, \dots, X_n variabile independente, identic repartizate log-normal (adică $Y_i = \ln X_i$ sînt repartizate $N(\lambda, \tau^2)$). Să se evalueze deplasarea estimatiei

$$t = \exp \left\{ \bar{Y} + \frac{1}{2} S_Y^2 \right\}$$

pentru parametrul „media repartiției log-normale” (s-a folosit notația

$$\bar{Y} = \frac{1}{n} \sum_{j=1}^n \ln X_j, \quad S_Y^2 = \frac{1}{n-1} \sum_{j=1}^n (\ln X_j - \bar{Y})^2.$$

Soluție. Dacă $Y = \ln X$ este repartizată $N(\lambda, \tau^2)$, atunci

$$M(X) = M(e^Y) = \exp \left(\lambda + \frac{\tau^2}{2} \right).$$

Deci parametrul ce trebuie estimat este $\theta = \exp \left(\lambda + \frac{\tau^2}{2} \right)$. Y_1, \dots, Y_n sînt variabile independente, identic repartizate $N(\lambda, \tau^2)$. \bar{Y} și S_Y^2

sînt estimății nedepasate, de dispersie minimă pentru λ și τ^2 . O estimăție a lui θ poate fi deci

$$t = \exp\left(\bar{Y} + \frac{1}{2} S_Y^2\right).$$

\bar{Y} urmează o repartiție $N\left(\lambda, \frac{\tau^2}{n}\right)$, iar $\frac{(n-1) S_Y^2}{\tau^2}$ este repartizată χ^2 cu $n-1$ grade de libertate. Atunci

$$M(e^{\bar{Y}}) = \exp\left(\lambda + \frac{1}{2} \frac{\tau^2}{n}\right),$$

$$M\left(\exp\left(\frac{1}{2} S_Y^2\right)\right) = \left(1 - \frac{\tau^2}{n-1}\right)^{-\frac{n-1}{2}}.$$

$$\begin{aligned} \text{Deci } M(t) &= \exp\left(\lambda + \frac{\tau^2}{2n} - \frac{n-1}{2} \ln\left(1 - \frac{\tau^2}{n-1}\right)\right) = \\ &= \exp\left(\lambda + \frac{\tau^2}{2}\right) \exp\left(\frac{\tau^2}{2n} + \frac{\tau^4}{4n} + o\left(\frac{1}{n^2}\right)\right). \end{aligned}$$

Deplasarea de ordinul $1/n$ se înlătură trecînd la estimăția

$$\tilde{t} = t \exp\left(-\frac{1}{n} \left(\frac{S_Y^2}{2} + \frac{S_Y^4}{4}\right)\right).$$

4.33. Să se determine estimăția de verosimilitate maximă (E.V.M.) pentru parametrii următoarelor repartiții:

- $N(\mu, \sigma^2)$ în cazul $\mu \in R$, $\sigma^2 > 0$ și în cazul $\mu \geq 0$, $\sigma^2 > 0$;
- exponențială c_μ , $\mu > 0$;
- gama $\gamma(k_0, \beta)$, $\beta > 0$;
- binomială $B_m(\theta)$, $0 < \theta < 1$;
- Poisson π_0 , $\theta > 0$.

Soluție. a) Fie X_1, \dots, X_n variabile independente, identic repartizate $N(\mu, \sigma^2)$.

$$\ln L(x_1, \dots, x_n; \mu, \sigma^2) = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2.$$

Se vede că funcția de log-verosimilitate este derivabilă în raport cu μ și σ^2 . Sistemul de verosimilitate maximă

$$\frac{\partial \ln L(x_1, \dots, x_n; \mu, \sigma^2)}{\partial \mu} = 0, \quad \frac{\partial \ln L(x_1, \dots, x_n; \mu, \sigma^2)}{\partial \sigma^2} = 0$$

are soluția unică $\hat{\mu} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$.

În cazul $\mu \in R$, $\sigma^2 > 0$, $(\hat{\mu}, \hat{\sigma}^2)$ dau un maxim global al funcției de verosimilitate și deci ele reprezintă E.V.M. pentru μ și σ^2 . În acest caz $\hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i$ este o estimatie eficientă a lui μ . Însă

$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ are media egală cu $(\sigma^2 - \frac{1}{n} \sigma^2)$, deci este o estimatie deplasată. Marginea inferioară a dispersiei unei estimatii a lui σ^2 este, din inegalitatea Rao-Cramer, $(1 - \frac{1}{n}) \frac{2\sigma^4}{n}$, iar $D^2(\hat{\sigma}^2) = \frac{n-1}{n^2} 2\sigma^4$, deci $\hat{\sigma}^2$ este estimatie de dispersie minimă a lui σ^2 .

În cazul $\mu \geq 0$, $\sigma^2 < 0$, $\hat{\mu}$ este E.V.M. doar dacă $\bar{x} \geq 0$. Deci E.V.M. este în acest caz $\hat{\mu} = \max(0, \bar{X})$, $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \hat{\mu})^2$.

b) Fie X_1, \dots, X_n variabile independente, identic repartizate e_μ .

$$\ln L(x_1, \dots, x_n; \mu) = -n \ln \mu - \frac{1}{\mu} \sum_{i=1}^n x_i.$$

Ecuatia de verosimilitate maximă $\frac{\partial \ln L(x_1, \dots, x_n; \mu)}{\partial \mu} = 0$ are soluția unică $\hat{\mu} = \bar{x}$. S-a arătat în problema 4.25 că \bar{X} este o estimatie eficientă pentru μ .

c) Fie X_1, \dots, X_n variabile independente, identic repartizate $\gamma(k_0, \beta)$, cu k_0 cunoscut, $\beta > 0$.

$$\begin{aligned} \ln L(x_1, \dots, x_n; \beta) &= -n(k_0 + 1) \ln \beta - \\ &- n \ln \Gamma(k_0 + 1) + (k_0 + 1) \sum_{i=1}^n \ln x_i - \frac{1}{\beta} \sum_{i=1}^n x_i. \end{aligned}$$

Ecuatia de verosimilitate maximă $\frac{\partial \ln L(x_1, \dots, x_n; \beta)}{\partial \beta} = 0$ are soluția unică $\hat{\beta} = \frac{\bar{x}}{k_0 + 1}$.

Estimația $\hat{\beta}(X_1, \dots, X_n) = \frac{1}{k_0 + 1} \bar{X}$ are $M_{\beta}(\hat{\beta}) = \beta$, $D_{\beta}^2(\hat{\beta}) = \frac{\beta^2}{n(k_0 + 1)}$.

Pe de altă parte, marginea inferioară a dispersiei unei estimații nedepășite a lui β este

$$\frac{1}{ni(\beta)} = \frac{1}{nM_{\beta}\left(\left[\frac{\partial \ln p(X, \beta)}{\partial \beta}\right]^2\right)} = \frac{\beta^2}{n(k_0 + 1)}$$

Deci E.V.M. $\hat{\beta}$ este eficientă.

d) Fie X_1, \dots, X_n variabile independente, identic repartizate $B_m(\theta)$, $0 < \theta < 1$.

$$L(x_1, \dots, x_n; \theta) = \left(\prod_{i=1}^n C_m^{x_i}\right) \theta^{\sum_{i=1}^n x_i} (1 - \theta)^{nm - \sum_{i=1}^n x_i}$$

Ecuția de verosimilitate maximă $\frac{\partial \ln L(x_1, \dots, x_n; \theta)}{\partial \theta} = 0$ are soluția unică $\hat{\theta} = \frac{1}{m} \bar{x}$, care realizează un maxim al funcției de verosimilitate. Estimația $\hat{\theta} = \frac{1}{m} \bar{X}$ are $M_{\theta}(\hat{\theta}) = \theta$, $D_{\theta}^2(\hat{\theta}) = \frac{\theta(1 - \theta)}{nm}$.

Pe de altă parte, marginea inferioară a dispersiei unei estimații nedepășite pentru θ este

$$\frac{1}{ni(\theta)} = \frac{1}{nM_{\theta}\left(\left[\frac{\partial \ln P(X; \theta)}{\partial \theta}\right]^2\right)} = \frac{\theta(1 - \theta)}{nm}$$

Deci E.V.M. $\hat{\theta}$ este o estimație eficientă a lui θ .

e) Fie X_1, \dots, X_n variabile independente, identic repartizate π_{θ} .

$$\ln L(x_1, \dots, x_n; \theta) = \sum_{i=1}^n x_i \ln \theta - n\theta - \sum_{i=1}^n \ln(x_i!)$$

Ecuția de verosimilitate maximă are soluția unică $\hat{\theta} = \bar{x}$, care realizează un maxim al funcției de verosimilitate. $\theta(X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^n X_i$ are $M_{\theta}(\hat{\theta}) = \theta$, $D_{\theta}^2(\hat{\theta}) = \frac{\theta}{n}$ și este o estimație eficientă a lui θ .

4.34. Să se construiască E.V.M. pentru parametrii repartiției normale s-dimensionale, pe baza unei selecții n-dimensionale, $n > s$.

Soluție. Fie X_1, \dots, X_n vectori aleatori independenți, identic repartizați normal $N(s; \mu, \Sigma)$, $n > s$.

$$p(x_1, \dots, x_n; \mu, \Sigma) = \frac{1}{(2\pi)^{ns/2} (\det \Sigma)^{n/2}} \cdot \exp \left\{ -\frac{1}{2} \sum_{i=1}^n ((x_i - \mu)^t \Sigma^{-1} (x_i - \mu)) \right\}.$$

Ne ocupăm întâi de estimarea lui μ și a lui Σ^{-1} , pentru că ele apar în forma densității de repartiție. Notăm $\Sigma^{-1} = \psi$.

Funcția de log-verosimilitate este

$$\ln L(x_1, \dots, x_n; \mu, \psi) = -\frac{ns}{2} \ln 2\pi + \frac{n}{2} \ln (\det \psi) - \frac{1}{2} \sum_{i=1}^n (x_i - \mu)^t \psi (x_i - \mu).$$

Fie media de selecție $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = (\bar{x}_1, \dots, \bar{x}_s)^t$ cu $\bar{x}_k = \frac{1}{n} \sum_{i=1}^n x_{ki}$ și fie matricea

$$A = \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})^t = \left\| \sum_{i=1}^n (x_{ki} - \bar{x}_k)(x_{ji} - \bar{x}_j) \right\|_{k,j=1, \dots, s}$$

Dintr-un calcul direct se obține

$$\sum_{i=1}^n (x_i - \mu)(x_i - \mu)^t = A + n(\bar{x} - \mu)(\bar{x} - \mu)^t.$$

Notînd $\text{tr } CD = \sum_{i,j} c_{ij}d_{ji} = \text{tr } DC$, putem scrie

$$\begin{aligned} \sum_{i=1}^n (x_i - \mu)^t \psi (x_i - \mu) &= \text{tr} \sum_{i=1}^n (x_i - \mu)^t \psi (x_i - \mu) = \\ &= \text{tr} \sum_{i=1}^n \psi (x_i - \mu)(x_i - \mu)^t = \text{tr } \psi A + \text{tr } \psi n(\bar{x} - \mu)(\bar{x} - \mu)^t = \\ &= \text{tr } \psi A + n(\bar{x} - \mu)^t \psi (\bar{x} - \mu). \end{aligned}$$

Atunci $\ln L(x_1, \dots, x_n; \mu, \psi) = -\frac{ns}{2} \ln 2\pi + \frac{n}{2} \ln(\det \psi) - \frac{1}{2} \text{tr} \psi A - \frac{1}{2} n(\bar{x} - \mu)^t \psi (\bar{x} - \mu)$.

Cum ψ este pozitiv semidefinită, $n(\bar{x} - \mu)^t \psi (\bar{x} - \mu) \geq 0$ și este egală cu zero doar pentru $\hat{\mu} = \bar{x}$.

Pentru a minimiza suma $\frac{n}{2} \ln(\det \psi) - \frac{1}{2} \text{tr} \psi A$ procedăm astfel: egalăm cu zero derivatele parțiale ale lui

$$f(\psi) = \frac{n}{2} \ln(\det \psi) - (\text{tr} \psi A) \frac{1}{2}$$

în raport cu cele s^2 elemente ale lui ψ . Obținem $\hat{\psi} = nA^{-1}$.

Deci E.V.M. pentru μ este $\hat{\mu} = \bar{X}$, iar pentru ψ este $\hat{\psi} = nA^{-1}$. Atunci E.V.M. a lui $\Sigma = \psi^{-1}$ este

$$\hat{\Sigma} = \frac{1}{n} A = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})^t = \frac{1}{n} \sum_{i=1}^n x_i x_i^t - \bar{x} \bar{x}^t.$$

4.35. Presupunem că probabilitatea $\pi(s)$ ca un individ să răspundă la nivelul s al unui stimul este aproximativ egală cu

$$\pi(s) = \Phi\left(\frac{s - \mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi}} \int_0^{\frac{s - \mu}{\sigma}} e^{-z^2/2} dz.$$

Nivelul s_i ($i = 1, \dots, r$) al stimulului se aplică la n_i indivizi, dintre care m_i răspund. Să se determine E.V.M. pentru parametrul bidimensional $\theta = (\mu, \sigma)^t$, $\mu \in R$, $\sigma > 0$.

Soluție. Notăm cu $x = (m_1, m_2, \dots, m_r)$ vectorul care dă numărul de răspunsuri pentru cele r nivele ale stimulului aplicate, iar cu π_i valorile $\pi(s_i)$.

În ipoteza că indivizii răspund în mod independent, funcția de verosimilitate este

$$L(x; \theta) = \prod_{i=1}^r C_{n_i}^{m_i} (\pi_i)^{m_i} (1 - \pi_i)^{n_i - m_i}.$$

$$\text{Sistemul de verosimilitate maximă} \left\{ \begin{array}{l} \frac{\partial \log L(x; \theta)}{\partial \mu} = 0 \\ \frac{\partial \log L(x; \theta)}{\partial \sigma} = 0 \end{array} \right. \text{ revine la}$$

$$\left\{ \begin{array}{l} \sum_{i=1}^r \frac{m_i - n_i \pi_i}{\pi_i(1 - \pi_i)} \frac{\partial \pi_i}{\partial \mu} = 0, \\ \sum_{i=1}^r \frac{m_i - n_i \pi_i}{\pi_i(1 - \pi_i)} \frac{\partial \pi_i}{\partial \sigma} = 0. \end{array} \right.$$

Acest sistem poate fi rezolvat doar prin metode numerice. Vom căuta întâi o aproximație inițială a soluției, $(\mu_0, \sigma_0)^t = \theta_0$.

Fie Φ^{-1} inversa funcției monoton crescătoare Φ . Dacă valorile π_i , $i = 1, \dots, r$ ar fi cunoscute, atunci punctele $(s_i, \Phi^{-1}(\pi_i))$ $i = 1, \dots, r$ s-ar dispune după dreapta $\Phi^{-1}(\pi) = (s - \mu)/\sigma$. Nu cunoaștem însă decât valorile m_i/n_i , $i = 1, \dots, r$, care estimează pe π_i , $i = 1, \dots, r$.

Vom figura punctele $A_i \left(s_i, \Phi^{-1} \left(\frac{m_i}{n_i} \right) \right)$ $i = 1, \dots, r$ și dacă ipoteza de normalitate este corectă, acestea se vor dispune aproximativ după o dreaptă. Scriem ecuația acestei drepte sub forma $\Phi^{-1}(\pi) = \frac{s - \mu_0}{\sigma_0}$ și luăm $\theta_0 = (\mu_0, \sigma_0)^t$ drept aproximație inițială a estimației de verosimilitate maximă. Deci figurarea punctelor A_i ne oferă atât o verificare a normalității datelor, cât și o aproximație inițială a E.V.M.

Notînd cu $\hat{\theta}$ estimația de verosimilitate maximă, din formula lui Taylor obținem:

$$0 = D_0(\ln p(x, \hat{\theta})) \simeq D_0(\ln p(x, \theta_0)) + (\hat{\theta} - \theta_0) D_0^2(\ln p(x, \theta_0)),$$

unde am notat prin D_0 operatorul $\left(\frac{\partial}{\partial \mu}, \frac{\partial}{\partial \sigma} \right)^t$ iar prin D_0^2 operatorul $\left\| \frac{\partial^2}{\partial \mu \partial \sigma} \right\|$.

$$\text{Atunci } \hat{\theta} \simeq \theta_0 - [D_0^2(\ln p(x, \theta_0))]^{-1} D_0(\ln p(x, \theta_0)).$$

Notăm cu $\theta_1 = (\mu_1, \sigma_1)^t$ noua soluție obținută. Procedeeul se reia, obținîndu-se șirul de estimații $\theta_n = (\mu_n, \sigma_n)^t$ care converge la $\hat{\theta}$ (metoda lui Newton).

4.36. Fie următorul model de regresie liniară:

$$X_i = \beta_0 + \beta_1 a_i + \varepsilon_i, \quad i = 1, \dots, n,$$

cu a_1, \dots, a_n scalari cunoscuți, nu toți egali între ei, $\varepsilon_1, \dots, \varepsilon_n$ variabile aleatoare „eroare” necorelate, cu $D^2(\varepsilon_i) = \sigma^2, i = 1, \dots, n$, iar β_0, β_1 parametri necunoscuți.

Să se determine estimățiile lui β_0 și β_1 prin metoda celor mai mici pătrate și să se calculeze dispersiile și covarianța estimatorilor obținuți.

Soluție. $\hat{\beta}_0$ și $\hat{\beta}_1$ se obțin din condiția $\min \sum_{i=1}^n (x_i - \beta_0 - \beta_1 a_i)^2$.

Egalând cu zero derivatele parțiale ale sumei de pătrate în raport cu β_0 și β_1 se obțin estimățiile

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (a_i - \bar{a}) X_i}{\sum_{i=1}^n (a_i - \bar{a})^2}, \quad \hat{\beta}_0 = \bar{X} - \bar{a} \hat{\beta}_1,$$

$$\text{cu } \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \text{ și } \bar{a} = \frac{1}{n} \sum_{i=1}^n a_i.$$

$$D^2(\hat{\beta}_1) = \frac{1}{\left(\sum_{i=1}^n (a_i - \bar{a})^2 \right)^2} \sum_{i=1}^n (a_i - \bar{a})^2 D^2(X_i) = \frac{\sigma^2}{\sum_{i=1}^n (a_i - \bar{a})^2}.$$

$$D^2(\hat{\beta}_0) = D^2(\bar{X}) + \bar{a}^2 D^2(\hat{\beta}_1) = \sigma^2 \left(\frac{1}{n} + \frac{\bar{a}^2}{\sum_{i=1}^n (a_i - \bar{a})^2} \right).$$

$$\text{cov}(\hat{\beta}_0, \hat{\beta}_1) = \text{cov}(\bar{X}, \hat{\beta}_1) - \bar{a} D^2(\hat{\beta}_1).$$

Cum $\text{cov}(\varepsilon_i, \varepsilon_j) = 0$ pentru $i \neq j$, rezultă $\text{cov}(X_i, X_j) = 0$ pentru $i \neq j$. Deci

$$\begin{aligned} \text{cov}(\bar{X}, \hat{\beta}_1) &= \text{cov} \left(\frac{1}{n} \sum_{i=1}^n X_i, \frac{1}{\sum_{j=1}^n (a_j - \bar{a})^2} \sum_{j=1}^n (a_j - \bar{a}) X_j \right) = \\ &= \frac{1}{n \sum_{j=1}^n (a_j - \bar{a})^2} \left(\sum_{i,j} (a_j - \bar{a}) \text{cov}(X_i, X_j) + \sigma^2 \sum_{j=1}^n (a_j - \bar{a}) \right) = 0. \end{aligned}$$

$$\text{Atunci } \text{cov}(\hat{\beta}_0, \hat{\beta}_1) = -\bar{a}^2 D^2(\hat{\beta}_1).$$

4.37. Fie X o variabilă aleatoare repartizată $B_n(p)$. Să se găsească o estimatie minimax pentru p , cind funcția de pierdere este eroarea pătratică.

Soluție. $P(X = x) = C_n^x p^x (1-p)^{n-x}$, $0 < p < 1$; $w(p, \delta(X)) = (\delta(X) - p)^2$.

Presupunem că $\delta(X)$ este o estimatie liniară de forma $\delta(X) = aX + b$, cu a și b constante necunoscute.

$$\begin{aligned} R_\delta(p) &= M_p(aX + b - p)^2 = M_p(((X - np) + b + (an - 1)p)^2) = \\ &= anpq + (b + (an - 1)p)^2 = \\ &= [(an - 1)^2 - a^2n] p^2 + [a^2n + 2b(an - 1)] p + b^2. \end{aligned}$$

Pentru a determina valorile lui a și b pentru care riscul este constant rezolvăm sistemul

$$\begin{cases} (an - 1)^2 - a^2n = 0 \\ a^2n + 2b(an - 1) = 0. \end{cases}$$

$$\text{Se obține } a_1 = \frac{1}{\sqrt{n}(1 + \sqrt{n})} \text{ sau } a_2 = \frac{1}{\sqrt{n}(1 - \sqrt{n})}$$

$$b_1 = \frac{1}{2(1 + \sqrt{n})} \text{ sau } b_2 = \frac{1}{2(2n - 3\sqrt{n} - 1)}$$

(a_2, b_2) se elimină, căci estimatia corespunzătoare poate lua valori în afara intervalului $(0, 1)$

Estimatiei lui p are deci forma

$$\delta(x) = \frac{x}{\sqrt{n}(1 + \sqrt{n})} + \frac{1}{2(1 + \sqrt{n})}$$

Dacă se poate găsi o repartitie a priori $\pi(p)$ pentru care estimatie Bayes are riscul constant, atunci obținem estimatie minimax.

Considerăm $\pi(p) = cp^{d-1}(1-p)^{\beta-1}$, repartitia beta cu $d > 0$, $\beta > 0$ și c convenabil aleasă. Atunci

$$\begin{aligned} \delta_\pi(x) &= \frac{\int_0^1 p C_n^x p^x (1-p)^{n-x} \pi(p) dp}{\int_0^1 C_n^x p^x (1-p)^{n-x} \pi(p) dp} = \frac{\int_0^1 p^{d+x} (1-p)^{n+\beta-x-1} dp}{\int_0^1 p^{d+x-1} (1-p)^{n+\beta-x-1} dp} = \\ &= \frac{\Gamma(d+x+1) \Gamma(n+\beta+x) \Gamma(d+\beta+n)}{\Gamma(x+d) \Gamma(n+\beta-x) \Gamma(d+\beta+n+1)} = \frac{x}{n+d+\beta} + \frac{d}{n+d+\beta}. \end{aligned}$$

Impunem condiția ca

$$a = \frac{1}{n+d+\beta} = \frac{1}{\sqrt{n}(1+\sqrt{n})}; \quad b = \frac{d}{n+d+\beta} = \frac{1}{2(1+\sqrt{n})},$$

de unde obținem $d = \beta = \frac{\sqrt{n}}{2}$. Atunci estimăția

$$\delta = \frac{\sqrt{n}}{n(1+\sqrt{n})} X + \frac{1}{2(1+\sqrt{n})}$$

este o estimăție minimax și are funcția de risc constantă, egală cu $\frac{1}{4(1+\sqrt{n})^2}$.

4.38. Fie X_1, X_2 variabile independente, identic repartizate $N(\theta, 1)$, $\theta \in R$. Pentru parametrul θ se consideră repartiția a priori $N(0, 1)$. Să se determine estimăția Bayes a lui θ , bazată pe observațiile (x_1, x_2) pentru funcția de pierdere eroarea pătratică.

Soluție. $p_0(x_1, x_2) = \frac{1}{2\pi} \exp\left\{-\frac{1}{2}((x_1 - \theta)^2 + (x_2 - \theta)^2)\right\}.$

$$p(\theta) = \frac{1}{\sqrt{2\pi}} \exp(-\theta^2/2).$$

$$\begin{aligned} p(\theta | x_1, x_2) &= \frac{(2\pi)^{-3/2} \exp\left\{-\frac{1}{2}((x_1 - \theta)^2 + (x_2 - \theta)^2 + \theta^2)\right\}}{(2\pi)^{-3/2} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}((x_1 - \theta)^2 + (x_2 - \theta)^2 + \theta^2)\right\} d\theta} = \\ &= \frac{1}{\sqrt{2\pi} \frac{1}{\sqrt{3}}} \exp\left\{-\frac{1}{2} \frac{1}{3} \left(\theta - \frac{x_1 + x_2}{3}\right)^2\right\}. \end{aligned}$$

Deci repartiția a posteriori a parametrului θ este $N\left(\frac{x_1 + x_2}{3}; \frac{1}{3}\right)$.

Atunci $\theta^*(X_1, X_2) = M(\theta | X_1, X_2) = \frac{X_1 + X_2}{3}$.

4.39. Fie R variabila aleatoare ce dă numărul de succese în n probe bernoulliene, cu probabilitatea de succes egală cu θ , $0 < \theta < 1$. Parametrul θ se consideră repartizat a priori uniform pe $(0, 1)$. Să se

determine estimatiya Bayes a lui θ , pentru functia de pierdere croarea pătratică.

Soluție. $P_0(R = r) = C_n^r \theta^r (1 - \theta)^{n-r}$. Densitatea de repartiție a priori a lui θ este $p(\theta) = \begin{cases} 1, & 0 < \theta < 1 \\ 0, & \text{în rest} \end{cases}$

$$p(\theta|r) = \frac{\theta^r (1 - \theta)^{n-r}}{\int_0^1 \theta^r (1 - \theta)^{n-r} d\theta} = \frac{1}{B(r+1, n-r+1)} \theta^r (1 - \theta)^{n-r}.$$

Estimatiya Bayes pentru pierderea pătratică este valoarea medie a posteriori a lui θ .

$$\theta^* = \frac{1}{B(r+1, n-r+1)} \int_0^1 \theta^{r+1} (1 - \theta)^{n-r} d\theta = \frac{r+1}{n+2}.$$

Se observă că dispersia repartiției a posteriori a lui θ este egală cu $\frac{(r+1)(n-r+1)}{(n+2)^2(n+3)}$. Pentru n mare, repartiția a posteriori este puternic centrată în jurul valorii θ^* , în timp ce repartiția a priori avea maximum de difuzie.

4.40. Fic X_1, \dots, X_n variabile independente, identic repartizate $N(m, \sigma^2)$, $m \in R$, $\sigma^2 > 0$. Să se construiască intervale de estimare pentru m , respectiv pentru σ^2 , cu coeficient de încredere egal cu $(1 - \alpha)$.

Soluție. a) Interval de estimare pentru m :

Funcția $g(x_1, \dots, x_n; m) = \frac{\sqrt{n}(\bar{x} - m)}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}}$ are următoarele

proprietăți: este măsurabilă în (x_1, \dots, x_n) , este monoton descrescătoare în m , iar repartiția lui $g(X_1, \dots, X_n; m)$ nu depinde de $\theta = (m, \sigma^2)$, și anume:

$\frac{\sqrt{n}(\bar{X} - m)}{\sigma}$ este repartizat $N(0, 1)$.

$$\frac{\frac{\sqrt{n}(\bar{X} - m)}{\sigma}}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n \left[\frac{X_i - m}{\sigma} - \frac{\bar{X} - m}{\sigma} \right]^2}} = \frac{\sqrt{n}(\bar{X} - m)}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}} = \frac{\sqrt{n}(\bar{X} - m)}{\sqrt{S^2}}$$

este repartizată Student cu $(n - 1)$ grade de libertate, adică $g(X_1, \dots, X_n; m) \sim t(n - 1)$.

Repartiția $t(r)$ este dată de densitatea de repartiție

$$\frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{r\pi} \Gamma\left(\frac{r}{2}\right)} \left(1 + \frac{x^2}{r}\right)^{-\frac{r+1}{2}}, \quad x \in R.$$

Pentru $\alpha > 0$ fixat, fie $t_{n-1; \alpha}$ cuantila de rang α a repartiției $t(n - 1)$. Atunci

$$P\left(t_{n-1; \alpha/2} \leq \frac{\sqrt{n}(\bar{X} - m)}{S} \leq t_{n-1; 1-\alpha/2}\right) = 1 - \alpha.$$

$$\left[\bar{x} - \frac{S}{\sqrt{n}} t_{n-1; 1-\alpha/2}; \bar{x} + \frac{S}{\sqrt{n}} t_{n-1; \alpha/2}\right]$$

este un interval de estimare pentru m , cu coeficient de încredere $(1 - \alpha)$.

b) Interval de estimare pentru σ^2 :

Funcția $g(x_1, \dots, x_n; \sigma^2) = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \bar{x})^2$ are următoarele proprietăți: este măsurabilă în (x_1, \dots, x_n) , descrescătoare în σ^2 , iar repartiția lui $g(X_1, \dots, X_n; \sigma^2)$ nu depinde de $\theta = (m, \sigma^2)$ și anume:

$Y_i = \frac{X_i - m}{\sigma}$, $i = 1, \dots, n$ sînt identic repartizate $N(0, 1)$.

$\sum_{i=1}^n (Y_i - \bar{Y})^2$ este repartizată χ^2 cu $(n - 1)$ grade de libertate.

Repartiția χ^2 cu r grade de libertate este dată de densitatea

$$\frac{1}{2^{r/2} \Gamma\left(\frac{r}{2}\right)} x^{\frac{r}{2}-1} e^{-x/2}, \quad x \geq 0.$$

Cum $g(X_1, \dots, X_n; \sigma^2) = \sum_{i=1}^n (Y_i - \bar{Y})^2$, rezultă că $g(X_1, \dots, X_n; \sigma^2)$ este repartizată $\chi^2(n - 1)$.

Pentru $\alpha > 0$ fixat, fie $h_{n-1; \alpha}^2$ cuantila de rang α a repartiției χ^2 cu $(n-1)$ grade de libertate. Atunci

$$P(h_{n-1; \alpha/2}^2 \leq g(X_1, \dots, X_n; \sigma^2) \leq h_{n-1; 1-\alpha/2}^2) = 1 - \alpha.$$

$$\left[\frac{1}{h_{n-1; 1-\alpha/2}^2} \sum_{i=1}^n (x - \bar{x})^2; \frac{1}{h_{n-1; \alpha/2}^2} \sum_{i=1}^n (x_i - \bar{x})^2 \right]$$

este un interval de estimare pentru σ^2 , cu coeficient de încredere $(1-\alpha)$.

4.41. Fie $X_1, \dots, X_{n_1}, X_{n_1+1}, \dots, X_{n_1+n_2}$ variabile aleatoare independente în totalitate, așa încît X_1, \dots, X_{n_1} sînt repartizate $N(\mu_1, \sigma_1^2)$, iar $X_{n_1+1}, \dots, X_{n_1+n_2}$ sînt repartizate $N(\mu_2, \sigma_2^2)$ cu σ_1^2 și σ_2^2 cunoscute. Fie $\rho = \mu_1/\mu_2$ raportul celor două medii. Să se construiască o regiune de estimare pentru ρ , cu coeficient de încredere $(1-\alpha)$.

Soluție. Considerăm o nouă parametrizare: $\mu_2 = \lambda$, $\mu_1 = \lambda\rho$. Pentru o valoare $\rho = \rho_0$ fixată, variabila $\bar{X}_{(1)} = \frac{1}{n_1} \sum_{i=1}^{n_1} X_i$ este repartizată $N(\lambda\rho_0; \frac{\sigma_1^2}{n_1})$, iar $\bar{X}_{(2)} = \frac{1}{n_2} \sum_{i=n_1+1}^{n_1+n_2} X_i$ este repartizată $N(\lambda; \frac{\sigma_2^2}{n_2})$.

$$\bar{X}_{(1)} - \rho_0 \bar{X}_{(2)} \text{ este repartizată } N\left(0; \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \rho_0^2\right).$$

$$\text{Atunci } T = \frac{\bar{X}_{(1)} - \rho_0 \bar{X}_{(2)}}{\sqrt{\frac{\sigma_1^2}{n_1} + \rho_0^2 \frac{\sigma_2^2}{n_2}}} \text{ este repartizată } N(0, 1).$$

Pentru un $\alpha > 0$ fixat, fie z_α cuantila de rang α a repartiției $N(0, 1)$. Atunci mulțimea

$$\left\{ \rho \mid |\bar{x}_{(1)} - \rho \bar{x}_{(2)}| \leq z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \rho^2 \frac{\sigma_2^2}{n_2}} \right\}$$

este o regiune de estimare pentru ρ , cu coeficient de încredere $(1-\alpha)$.

4.42. Fie modelul regresiei liniare $X_i = \beta_0 + \beta_1 a_i + \varepsilon_i$, $i = 1, \dots, n$, cu a_1, \dots, a_n constante cunoscute, $\varepsilon_1, \dots, \varepsilon_n$ variabile aleatoare independente, identic repartizate $N(0, \sigma^2)$, iar $\beta_0, \beta_1, \sigma^2$

parametri necunoscuți. Să se determine intervale de estimare cu coeficient de încredere $(1-\alpha)$ pentru β_1 și pentru $\beta_0 + c\beta_1$ ($c \in R$).

Soluție. a) Estimațiile obținute prin metoda celor mai mici pătrate pentru β_0 , β_1 sînt

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (a_i - \bar{a}) x_i}{\sum_{i=1}^n (a_i - \bar{a})^2}, \quad \hat{\beta}_0 = \bar{x} - \bar{a} \hat{\beta}_1 \quad (\text{problema 4.36})$$

$\hat{\beta}_1(X_1, \dots, X_n)$ urmează o repartiție $N\left(\beta_1, \frac{\sigma^2}{\sum_{i=1}^n (a_i - \bar{a})^2}\right)$. Cum σ^2 este necunoscut, îl vom înlocui cu un estimator nedeplasat al său

$$s^2 = \frac{1}{n-2} \sum_{i=1}^n (X_i - \hat{\beta}_0 - \hat{\beta}_1 a_i)^2.$$

Variabila $\frac{(\hat{\beta}_1 - \beta_1) \sqrt{\sum_{i=1}^n (a_i - \bar{a})^2}}{s}$ urmează o repartiție Student cu $(n-2)$ grade de libertate.

Pentru $\alpha > 0$ fixat, fie $t_{n-2; \alpha}$ cuantila de rang α a repartiției $t(n-2)$. Atunci

$$\left[\hat{\beta}_1 - t_{n-2; 1-\alpha/2} \frac{s}{\sqrt{\sum_{i=1}^n (a_i - \bar{a})^2}}; \hat{\beta}_1 - t_{n-2; \alpha/2} \frac{s}{\sqrt{\sum_{i=1}^n (a_i - \bar{a})^2}} \right]$$

este un interval de estimare pentru β_1 , cu coeficient de încredere $(1-\alpha)$.

b) $\hat{\beta}_0(X_1, \dots, X_n) + c\beta_1(X_1, \dots, X_n)$ urmează o repartiție

$$N\left(\beta_0 + c\beta_1; \sigma^2 \left(\frac{1}{n} + \frac{(\bar{a} - c)^2}{\sum_{i=1}^n (a_i - \bar{a})^2} \right)\right),$$

deci variabila

$\frac{\hat{\beta}_0 + c\hat{\beta}_1 - (\beta_0 + c\beta_1)}{s \sqrt{\frac{1}{n} + \frac{(\bar{a} - c)^2}{\sum_{i=1}^n (a_i - \bar{a})^2}}}$ este repartizată Student cu $(n-2)$ grade de

libertate.

Pentru $\alpha > 0$ fixat, mulțimea

$$\left[\hat{\beta}_0 + c\hat{\beta}_1 - t_{n-2; 1-\alpha/2} S \sqrt{\frac{1}{n} + \frac{(\bar{a} - c)^2}{\sum_{i=1}^n (a_i - \bar{a})^2}} ; \right. \\ \left. \hat{\beta}_0 + c\hat{\beta}_1 - t_{n-2; \alpha/2} S \sqrt{\frac{1}{n} + \frac{(\bar{a} - c)^2}{\sum_{i=1}^n (a_i - \bar{a})^2}} \right]$$

este un interval de estimare pentru $(\beta_0 + c\beta_1)$, cu coeficient de încredere $(1 - \alpha)$.

4.43. Fie modelul de regresie liniară $X_i = \beta_0 + \beta_1 a_i + \varepsilon_i$, $i = 1, \dots, n$, cu a_1, \dots, a_n constante cunoscute, $\varepsilon_1, \dots, \varepsilon_n$ variabile independente, identic repartizate, cu $D^2(\varepsilon_i) = \sigma^2$, $i = 1, \dots, n$, iar $\beta_0, \beta_1, \sigma^2$ parametri necunoscuți.

Pentru $c_1, c_2 \in R$, considerăm parametrizarea $\theta = (\theta_1, \theta_2)$, $\theta_1 = \beta_0 + c_1\beta_1$, $\theta_2 = \beta_0 + c_2\beta_1$. Să se determine o regiune de estimare pentru parametrul bidimensional θ , cu coeficient de încredere $(1 - \alpha)$.

Soluție. Fie $\hat{\beta}_0, \hat{\beta}_1$ estimațiile lui β_0 și β_1 obținute prin metoda celor mai mici pătrate (problema 4.36).

$$M(\hat{\beta}_0) = \beta_0, \quad D^2(\hat{\beta}_0) = \sigma^2 \left(\frac{1}{n} + \frac{\bar{a}^2}{\sum_{i=1}^n (a_i - \bar{a})^2} \right),$$

$$M(\hat{\beta}_1) = \beta_1, \quad D^2(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^n (a_i - \bar{a})^2}, \quad \text{cov}(\hat{\beta}_0, \hat{\beta}_1) = -\frac{\bar{a}\sigma^2}{\sum_{i=1}^n (a_i - \bar{a})^2}.$$

Fie $\hat{\theta}_1 = \hat{\beta}_0 + c_1\hat{\beta}_1$, $\hat{\theta}_2 = \hat{\beta}_0 + c_2\hat{\beta}_1$. Atunci

$$M(\hat{\theta}_1) = \beta_0 + c_1\beta_1 = \theta_1, \quad D^2(\hat{\theta}_1) = \sigma^2 \left(\frac{1}{n} + \frac{(\bar{a} - c_1)^2}{\sum_{i=1}^n (a_i - \bar{a})^2} \right),$$

$$M(\hat{\theta}_2) = \beta_0 + c_2\beta_1 = \theta_2, \quad D^2(\hat{\theta}_2) = \sigma^2 \left(\frac{1}{n} + \frac{(\bar{a} - c_2)^2}{\sum_{i=1}^n (a_i - \bar{a})^2} \right),$$

$$\text{cov}(\hat{\theta}_1, \hat{\theta}_2) = \sigma^2 \left(\frac{1}{n} + \frac{(\bar{a} - c_1)(\bar{a} - c_2)}{\sum_{i=1}^n (a_i - \bar{a})^2} \right).$$

Utilizând o teoremă limită centrală bidimensională, stabilim că $(\hat{\beta}_0, \hat{\beta}_1)$ urmează asimptotic o repartiție normală bidimensională. Deci pentru n suficient de mare, $(\hat{\theta}_1, \hat{\theta}_2)' \sim N(2; (\theta_1, \theta_2)', \sigma^2 \Sigma)$ cu

$$\Sigma = \begin{pmatrix} \frac{1}{n} + \frac{(\bar{a} - c_1)^2}{\sum_{i=1}^n (a_i - \bar{a})^2} & \frac{1}{n} + \frac{(\bar{a} - c_1)(\bar{a} - c_2)}{\sum_{i=1}^n (a_i - \bar{a})^2} \\ \frac{1}{n} + \frac{(\bar{a} - c_1)(\bar{a} - c_2)}{\sum_{i=1}^n (a_i - \bar{a})^2} & \frac{1}{n} + \frac{(\bar{a} - c_2)^2}{\sum_{i=1}^n (a_i - \bar{a})^2} \end{pmatrix}$$

Fie $s^2 = \frac{1}{n-2} \sum_{i=1}^n (x_i - \hat{\beta}_0 - \hat{\beta}_1 a_i)^2$ o estimatie nedepasata a lui σ^2 . Deci

$$(\hat{\theta}_1, \hat{\theta}_2)' \sim N(2; (\theta_1, \theta_2)', s^2 \Sigma).$$

Atunci variabila $(\hat{\theta} - \theta)'(s^2 \Sigma)^{-1}(\hat{\theta} - \theta)$ urmeaza o repartitie χ^2 cu doua grade de libertate.

Pentru $\alpha > 0$ fixat, fie $h_{2; 1-\alpha}^2$ cuantila de rang $(1 - \alpha)$ a repartitiei $\chi^2(2)$. Multimea

$$\{0 = (\theta_1, \theta_2) | (\hat{\theta}(x_1, \dots, x_n) - \theta)'(s^2 \Sigma)^{-1}(\hat{\theta}(x_1, \dots, x_n) - \theta) \leq h_{2; 1-\alpha}^2\}$$

este o regiune de estimare pentru parametrul bidimensional θ , cu coeficient de incredere $(1 - \alpha)$.

§ 4. Verificarea ipotezelor statistice. Teste parametrice

4.44. Sa se construiasca testul cel mai puternic (bazat pe lema Neyman - Pearson) pentru verificarea unei ipoteze simple cu alternativa simpla pentru repartitiile de tip exponential. Sa se precizeze testul in cazul repartitiilor $B_m(p)$, π_λ , $N(\mu, \sigma_0^2)$ cu σ_0^2 cunoscut.

Solutie. a) Fic X_1, \dots, X_n variabile independente, identic repartizate cu densitatea de repartitie

$$p(x, \theta) = \exp \{a(\theta)b(x) + c(\theta) + d(x)\}, \theta \in R.$$

Ne propunem să verificăm ipoteza $H_0: \{\theta = \theta_0\}$ cu alternativa $H_A: \{\theta = \theta_A\}$, la pragul de semnificație α .

$$\Lambda(x_1, \dots, x_n) = \frac{p(x_1, \dots, x_n; \theta_A)}{p(x_1, \dots, x_n; \theta_0)} = \\ = \exp \{ (a(\theta_A) - a(\theta_0)) S(x_1, \dots, x_n) + n(c(\theta_A) - c(\theta_0)) \}$$

cu $S(x_1, \dots, x_n) = \sum_{i=1}^n b(x_i)$.

Regiunea critică a celui mai puternic test este de forma

$$W = \{ (x_1, \dots, x_n) \mid \ln \Lambda(x_1, \dots, x_n) \geq c_\alpha \} = \\ = \{ (x_1, \dots, x_n) \mid (a(\theta_A) - a(\theta_0)) S(x_1, \dots, x_n) \geq k_\alpha \}.$$

Dacă $a(\theta_A) - a(\theta_0) > 0$, atunci regiunea critică

$$W = \{ (x_1, \dots, x_n) \mid S(x_1, \dots, x_n) \geq k'_\alpha \}$$

este aceeași pentru orice valoare alternativă θ_A .

Pentru a determina complet regiunea critică, trebuie să cunoaștem repartiția lui $S(X_1, \dots, X_n)$ pentru H_0 adevărată. Dacă această repartiție nu poate fi determinată exact, se utilizează aproximarea normală.

b) Fie X_1, \dots, X_n variabile independente, identic repartizate $B_m(p)$. Ne propunem să verificăm ipoteza $H_0: \{p = p_0\}$ cu alternativa $H_A: \{p = p_A\}$ ($p_A < p_0$) la pragul de semnificație α .

$$P(x; p) = \exp \{ x(\ln p - \ln(1 - p)) + m \ln(1 - p) + \ln C_n^x \}$$

$$\Lambda(x_1, \dots, x_n) = \exp \left\{ \left(\ln \frac{p_A}{p_0} - \ln \frac{1-p_A}{1-p_0} \right) \sum_{i=1}^n x_i + nm \ln \frac{1-p_A}{1-p_0} \right\}$$

$$W = \left\{ (x_1, \dots, x_n) \mid \ln \frac{p_A(1-p_0)}{p_0(1-p_A)} \sum_{i=1}^n x_i \geq k_\alpha \right\} = \\ = \left\{ (x_1, \dots, x_n) \mid \sum_{i=1}^n x_i \geq k'_\alpha \right\}.$$

Statistica $S = \sum_{i=1}^n X_i$ urmează o repartiție $B_{nm}(p)$.

Constanta k'_α se determină din condiția ca probabilitatea erorii de ordinul întâi să fie α , $P_{p_0}(S \geq k'_\alpha) = \alpha$, sau, echivalent,

$$\sum_{j=0}^{k'_\alpha-1} C_{nm}^j p_0^j (1-p_0)^{nm-j} = 1 - \alpha.$$

c) Fie X_1, \dots, X_n variabile independente, identic repartizate π_{λ_0} , și fie $H_0: \{\lambda = \lambda_0\}$, $H_A: \{\lambda = \lambda_A\} (\lambda_A < \lambda_0)$.

$$P(x; \lambda) = \exp \{x \ln \lambda - \lambda - \ln x!\}$$

$$\Lambda(x_1, \dots, x_n) = \exp \left\{ \ln \frac{\lambda_A}{\lambda_0} \sum_{i=1}^n x_i - n(\lambda_A - \lambda_0) \right\}.$$

$$W = \left\{ (x_1, \dots, x_n) \mid \ln \frac{\lambda_A}{\lambda_0} \sum_{i=1}^n x_i \geq k'_\alpha \right\} = \left\{ (x_1, \dots, x_n) \mid \sum_{i=1}^n x_i \geq k'_\alpha \right\},$$

Statistica $S = \sum_{i=1}^n X_i$ urmează o repartiție $\pi_{n\lambda}$. Constanta k'_α se determină din condiția $P_{\lambda_0}(S \geq k'_\alpha) = \alpha$ sau, echivalent,

$$\sum_{j=0}^{k'_\alpha-1} \frac{(n\lambda_0)^j}{j!} e^{-n\lambda_0} = 1 - \alpha.$$

d) Fie X_1, \dots, X_n variabile independente, identic repartizate $N(\mu, \sigma_0^2)$ cu $\mu \in R$. σ_0^2 cunoscut, și fie $H_0: \{\mu = \mu_0\}$, $H_A: \{\mu = \mu_A\} (\mu_A > \mu_0)$.

$$p(x; \mu) = \exp \left\{ \frac{\mu}{\sigma_0^2} x - \frac{1}{2\sigma_0^2} x^2 - \frac{\mu^2}{2\sigma_0^2} - \frac{1}{2} \ln \sigma_0^2 - \frac{1}{2} \ln 2\pi \right\}.$$

$$\Lambda(x_1, \dots, x_n) = \exp \left\{ \frac{\mu_A - \mu_0}{\sigma_0^2} \sum_{i=1}^n x_i - n \frac{\mu_A^2 - \mu_0^2}{2\sigma_0^2} \right\}.$$

$$W = \left\{ (x_1, \dots, x_n) \mid \frac{\mu_A - \mu_0}{\sigma_0^2} \sum_{i=1}^n x_i \geq k'_\alpha \right\} = \left\{ (x_1, \dots, x_n) \mid \frac{1}{n} \sum_{i=1}^n x_i \geq k'_\alpha \right\}.$$

Statistica $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ urmează o repartiție $N(\mu, \sigma_0^2/n)$.

Constanta k'_α se determină din condiția $P_{\mu_0}(\bar{X} \geq k'_\alpha) = \alpha$ sau, echivalent,

$$\frac{\sqrt{n}}{\sqrt{2\pi\sigma_0^2}} \int_{-\infty}^{k'_\alpha} \exp\left(-\frac{n(y - \mu_0)^2}{2\sigma_0^2}\right) dy = 1 - \alpha.$$

Dar $\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma_0}$ urmează o repartiție $N(0, 1)$. Putem impune atunci condiția $P_{\mu_0}\left(\frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma_0} \geq \frac{n(k'_\alpha - \mu_0)}{\sigma_0}\right) = \alpha$.

Notînd $z_{1-\alpha}$ cuantila de rang $(1-\alpha)$ a repartiției $N(0, 1)$, avem

$$\frac{\sqrt{n}(k'_\alpha - \mu_0)}{\sigma_0} = z_{1-\alpha}, \text{ de unde } k'_\alpha = \mu_0 + \frac{\sigma_0 z_{1-\alpha}}{\sqrt{n}}.$$

4.45. Fie X_1, \dots, X_n variabile independente, identic repartizate $N(\mu, \gamma_0^2 \mu^2)$ cu γ_0 cunoscut, $\mu \in R$.

a) Să se construiască testul cel mai puternic pentru verificarea ipotezei $H_0: \{\mu = \mu_0\}$ cu alternativa $H_A: \{\mu = \mu_A\}$, ($\mu_A > \mu_0$).

b) Pentru aceleași ipoteze să se prezinte testele bazate pe statisticile

$$S_1 = \sum_{j=1}^n (X_j - \mu_0), \quad S_2 = \sum_{j=1}^n (X_j - \mu_0)^2.$$

Soluție. a) $\Lambda(x_1, \dots, x_n) = \left(\frac{\mu_A}{\mu_0}\right)^n \exp\left\{-\frac{\sum_{i=1}^n (x_i - \mu_A)^2}{2\gamma_0^2 \mu_A^2} + \frac{\sum_{i=1}^n (x_i - \mu_0)^2}{2\gamma_0^2 \mu_0^2}\right\}$

$$\Lambda(x_1, \dots, x_n) = \left(\frac{\mu_A}{\mu_0}\right)^n \exp\left\{-\frac{\mu_0 - \mu_A}{2\gamma_0^2 \mu_0 \mu_A} \left[\left(\frac{1}{\mu_0} + \frac{1}{\mu_A}\right) \sum_{i=1}^n x_i^2 - 2 \sum_{i=1}^n x_i\right]\right\}.$$

Fie $Y_j = \left(\frac{1}{\mu_0} + \frac{1}{\mu_A}\right) X_j^2 - 2X_j$, $j = 1, \dots, n$ și $T = \sum_{j=1}^n Y_j$.

Raportul de verosimilitate este funcție de statistica T , deci construcția celui mai puternic test se va face utilizînd această statistică. Regiunea critică va fi de forma

$$W = \{(x_1, \dots, x_n) | T(x_1, \dots, x_n) \geq k_\alpha\}.$$

Pentru a determina constanta k_α trebuie să cunoaștem repartiția lui T atunci cînd H_0 este adevărată.

Pentru $\mu = \mu_0$, $Z_j = \frac{X_j - \mu_0}{\gamma_0 \mu_0}$ urmează repartiția $N(0, 1)$, $j = 1, \dots, n$

De aici, $X_j = \mu_0 + Z_j \gamma_0 \mu_0$,

$$Y_j = \frac{\mu_0 + \mu_A}{\mu_0 \mu_A} (\mu_0 + Z_j \gamma_0 \mu_0)^2 - 2(\mu_0 + Z_j \gamma_0 \mu_0), \text{ sau}$$

$$Y_j = \delta (Z_j + \beta)^2 + \gamma \text{ cu}$$

$$\delta = \frac{\mu_0 \gamma_0^2 (\mu_0 + \mu_A)}{\mu_A}, \quad \beta = \frac{\mu_0}{\gamma_0 (\mu_0 + \mu)}, \quad \gamma = -\frac{\mu_0 \mu_A}{\mu_0 + \mu_A}.$$

$$\text{Atunci } T = \delta \sum_{j=1}^n (Z_j + \beta)^2 + n\gamma.$$

$$\text{Notind } T' = \sum_{j=1}^n (Z_j + \beta)^2, \text{ avem } T = \delta T' + n\gamma.$$

Dacă H_0 este adevărată, T' urmează o repartiție χ^2 necentrată, cu n grade de libertate și parametru de necentricitate $n\beta$.

$$W = \{(x_1, \dots, x_n) \mid \delta T'(x_1, \dots, x_n) + n\gamma \geq k_\alpha\}.$$

Presupunind $\delta > 0$ obținem $W = \{(x_1, \dots, x_n) \mid T'(x_1, \dots, x_n) \geq k'_\alpha\}$, unde k'_α va fi cuantila de rang $(1-\alpha)$ a repartiției χ^2 cu n grade de libertate și parametru de necentricitate $n\beta$.

Dacă H_A este adevărată, Z_j rămân repartizate normal:

$$Z_j \sim N\left(\frac{\mu_A - \mu_0}{\gamma_0 \mu_0}, \frac{\mu_A^2}{\mu_0^2}\right) j = 1, \dots, n.$$

Atunci statistica T' va fi proporțională tot cu o repartiție χ^2 necentrată cu n grade de libertate, al cărui parametru de necentricitate poate fi determinat. Deci și funcția de putere a testului se determină utilizând o repartiție χ^2 necentrată.

Facem și observația că T' este funcție de statisticile $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$,

și $\sum_{i=1}^n (X_i - \bar{X})^2$. Printr-un calcul direct se obține

$$T' = \frac{1}{\gamma_0^2 \mu_0^2} \sum_{j=1}^n (X_j - \bar{X})^2 + n \left(\frac{\bar{X}}{\gamma_0 \mu_0} - \frac{\mu_A}{\gamma_0 (\mu_0 + \mu_A)} \right).$$

b) Dacă H_0 este adevărată, statistica $S_1 = \sum_{i=1}^n (X_i - \mu_0)$ urmează o repartiție $N(0, n\gamma_0^2\mu_0^2)$. Regiunea critică construită pe baza lui

$$S_1 \text{ este } W_1 = \left\{ (x_1, \dots, x_n) \mid \frac{S_1(x_1, \dots, x_n)}{\sqrt{n\gamma_0\mu_0}} \geq z_{1-\alpha} \right\},$$

unde $z_{1-\alpha}$ este cuantila de rang $(1-\alpha)$ a repartiției $N(0, 1)$.

Dacă H_A este adevărată, S_1 urmează repartiția $N(n(\mu_A - \mu_0), n\gamma_0^2\mu_A^0)$.

Funcția de putere a testului bazat pe S_1 este deci

$$\pi(\mu_A) = P_{\mu_A}(W_1) = 1 - \Phi_{N(0, 1)}\left(\frac{\mu_0}{\mu_A}\left(z_{1-\alpha} - \frac{\sqrt{n}(\mu_A - \mu_0)}{\gamma_0\mu_0}\right)\right).$$

Dacă H_0 este adevărată, statistica $S_2 = \sum_{j=1}^n (X_j - \mu_0)^2$ urmează o repartiție χ^2 centrată, cu n grade de libertate. Regiunea critică construită pe baza lui S_2 este

$$W_2 = \left\{ (x_1, \dots, x_n) \mid \frac{S_2(x_1, \dots, x_n)}{\gamma_0^2\mu_0^2} \geq h_{n; 1-\alpha}^2 \right\},$$

unde $h_{n; 1-\alpha}^2$ este cuantila de rang $(1-\alpha)$ a repartiției χ^2 cu n grade de libertate.

Dacă H_A este adevărată, S_2 urmează o repartiție χ^2 necentrată, cu n grade de libertate. Funcția de putere a acestui test se poate determina deci prin utilizarea repartiției χ^2 necentrate corespunzătoare.

4.46. Să se cerceteze existența testului uniform cel mai puternic pentru verificarea ipotezei $H_0: \{\mu = \mu_0\}$, cu una din alternativele $H_1: \{\mu > \mu_0\}$, $H_2: \{\mu \neq \mu_0\}$, în cazul repartiției normale $N(\mu, 1)$, $\mu \in R$.

Soluție. a) Fie X_1, \dots, X_n variabile independente, identic repartizate $N(\mu, 1)$ și fie ipoteza simplă $H_0: \{\mu = \mu_0\}$ cu alternativa compusă $H_1: \{\mu > \mu_0\}$. Fie o valoare $\mu, \mu > \mu_0$.

$$\begin{aligned} \Lambda(x_1, \dots, x_n) &= \frac{p(x_1, \dots, x_n; \mu)}{p(x_1, \dots, x_n; \mu_0)} = \\ &= \exp \left\{ -\frac{1}{2} \sum_{i=1}^n (x_i - \mu)^2 + \frac{1}{2} \sum_{i=1}^n (x_i - \mu_0)^2 \right\} = \\ &= \exp \left\{ n(\mu - \mu_0) \left(\bar{x} - \frac{\mu_0 + \mu}{2} \right) \right\}. \end{aligned}$$

Mulțimea $W = \{(x_1, \dots, x_n) | \bar{x} > k_\alpha\}$ va fi regiunea critică uniform cea mai bună la pragul de semnificație α , dacă determinăm constanta k_α din condiția

$$P_{\mu_0}(W) = \alpha.$$

Atunci

$$W = \left\{ (x_1, \dots, x_n) \mid \bar{x} > \frac{z_{1-\alpha}}{\sqrt{n}} + \mu_0 \right\},$$

unde $z_{1-\alpha}$ este cuantila de rang $(1-\alpha)$ a repartiției $N(0, 1)$.

b) Fie din nou X_1, \dots, X_n variabile independente, identic repartizate $N(\mu, 1)$ și fie ipoteza simplă $H_0: \{\mu = \mu_0\}$ cu alternativa compusă $H_2: \{\mu \neq \mu_0\}$. Fie o valoare μ , $\mu \neq \mu_0$.

$$\text{Dacă } \mu > \mu_0, \quad W_1 = \left\{ (x_1, \dots, x_n) \mid \bar{x} > \frac{z_{1-\alpha}}{\sqrt{n}} + \mu_0 \right\}$$

este o regiune critică uniform cea mai puternică pentru orice $\mu > \mu_0$,

$$\text{Dacă } \mu < \mu_0, \quad W_2 = \left\{ (x_1, \dots, x_n) \mid \bar{x} < -\frac{z_{1-\alpha}}{\sqrt{n}} + \mu_0 \right\}$$

este o regiune critică uniform cea mai puternică pentru orice $\mu < \mu_0$,

Deci, nu putem găsi un test uniform cel mai puternic pentru verificarea ipotezei H_0 cu alternativa H_2 .

4.47. Fie X_1, \dots, X_n variabile independente, identic repartizate exponențial, cu densitatea de repartiție

$$p(x; \rho) = \rho e^{-\rho x}, \quad x \geq 0, \rho > 0.$$

Să se determine testul uniform cel mai puternic pentru verificarea ipotezei $H_0: \{\rho = \rho_0\}$ cu alternativa $H: \{\rho < \rho_0\}$ la pragul de semnificație α . Să se determine funcția de putere a acestui test.

Soluție. Fie o valoare ρ , $\rho < \rho_0$.

$$\Lambda(x_1, \dots, x_n) = \frac{p(x_1, \dots, x_n; \rho)}{p(x_1, \dots, x_n; \rho_0)} = \left(\frac{\rho}{\rho_0}\right)^n \exp\left\{- (\rho - \rho_0) \sum_{i=1}^n x_i\right\}.$$

Regiunea critică va fi de formă $W = \{(x_1, \dots, x_n) | \Lambda(x_1, \dots, x_n) \geq c_\alpha\}$.

Logaritmind raportul de verosimilitate, obținem:

$$W = \left\{ (x_1, \dots, x_n) \mid \sum_{i=1}^n x_i \geq k_\alpha \right\} \text{ pentru orice } \rho < \rho_0.$$

Dacă H_0 este adevărată, $S(X_1, \dots, X_n) = \sum_{i=1}^n X_i$ are densitatea de repartiție

$$p(s, \rho_0) = \frac{\rho_0(\rho_0 s)^{n-1} e^{-\rho_0 s}}{(n-1)!}$$

Rezultă imediat că $T = 2\rho_0 S = 2\rho_0 \sum_{i=1}^n X_i$ urmează o repartiție χ^2 cu $2n$ grade de libertate, Atunci

$$W = \left\{ (x_1, \dots, x_n) \mid \sum_{i=1}^n x_i \geq \frac{1}{2\rho_0} h_{2n; 1-\alpha}^2 \right\}$$

este regiunea critică uniform cea mai bună pentru ipoteza H_0 cu alternativa H_A , la pragul de semnificație α . ($h_{2n; 1-\alpha}^2$ este cuantila de rang $(1-\alpha)$ a repartiției χ^2 cu $2n$ grade de libertate).

Funcția de putere a testului este

$$\pi(\rho) = P_\rho(W) = P\left(2\rho \sum_{i=1}^n X_i \geq \frac{\rho}{\rho_0} h_{2n; 1-\alpha}^2\right).$$

4.48. Fie X_1, \dots, X_n variabile independente, identic repartizate cu densitatea de repartiție de forma

$$p(x, \theta) = \exp\left(-\frac{1}{2}x^2 - \theta x^4\right), \quad x \in R, \quad \theta \in R.$$

Să se determine testul uniform cel mai puternic pentru verificarea ipotezei $H_0: \{\theta = 0\}$ cu alternativa $H_A: \{\theta > 0\}$, la pragul de semnificație α .

Soluție. Fie o valoare $\theta, \theta > 0$.

$$\Lambda(x_1, \dots, x_n) = \frac{p(x_1, \dots, x_n; \theta)}{p(x_1, \dots, x_n; 0)} = \exp\left\{-\theta \sum_{j=1}^n x_j^4\right\}.$$

Regiunea critică are forma $W = \left\{ (x_1, \dots, x_n) \mid \sum_{j=1}^n x_j^4 \leq k_\alpha \right\}$ pentru orice $\theta > 0$.

Nu putem stabili repartiția exactă a statisticii $\sum_{j=1}^n X_j^4$ care definește testul, dar vom utiliza repartiția sa asimptotică.

Știm că dacă X este repartizată $N(0, 1)$, atunci $M(X^{4m}) = \frac{(4m)!}{2^{2m} (2m)!}$

de unde $M(X^4) = 3$, $D^2(X^4) = 96$.

Utilizînd teorema limită centrală, variabila $\frac{\sum_{j=1}^n X_j^4 - 3n}{\sqrt{96n}}$ este repartizată asimptotic $N(0, 1)$ cînd H_0 este adevărată. Deci pentru n suficient de mare, regiunea critică uniform cea mai bună la pragul de semnificație α este

$$W = \left\{ (x_1, \dots, x_n) \mid \sum_{j=1}^n x_j^4 \leq 3n + z_\alpha \sqrt{96n} \right\},$$

unde z_α este cuantila de rang α a repartiției $N(0, 1)$.

4.49. Fie X_1, \dots, X_n variabile independente, identic repartizate, cu densitatea de repartiție $p(x; \theta)$. Să se construiască testul local cel mai puternic pentru verificarea ipotezei $H_0: \{\theta = \theta_0\}$ cu alternativa $H: \{\theta > \theta_0\}$, la pragul de semnificație α .

Aplicație. Fie X_1, \dots, X_n independente, identic repartizate $N(\theta, 1 + a\theta^2)$ cu a constantă cunoscută. Să se construiască testul local cel mai puternic pentru verificarea ipotezei $H_0: \{\theta = 0\}$ cu alternativa $H: \{\theta > 0\}$, la pragul de semnificație α .

Soluție. Considerăm ipoteza alternativă $H_A: \{\theta = \theta_0 + \delta\}$ cu $\delta > 0$.

$$\Lambda(x_1, \dots, x_n) = \frac{p(x_1, \dots, x_n; \theta_0 + \delta)}{p(x_1, \dots, x_n; \theta_0)}$$

Pentru δ suficient de mic avem:

$$\ln \Lambda(x_1, \dots, x_n) = \delta \frac{\partial \ln p(x_1, \dots, x_n; \theta_0)}{\partial \theta_0} + O(\delta^2).$$

Pentru construcția testului local cel mai puternic se utilizează statistica $U(\theta_0) = \frac{\partial \ln p(X_1, \dots, X_n; \theta_0)}{\partial \theta_0}$, care are media $M_{\theta_0}(U(\theta_0)) = 0$ și dispersia $D_{\theta_0}^2(U(\theta_0)) = I(\theta_0)$ (informația Fisher).

Regiunea critică are forma

$$W = \{(x_1, \dots, x_n) \mid U(\theta_0) \geq k_\alpha\}.$$

Constanta k_α se determină așa încît $P_{\theta_0}(W) = \alpha$.

A p l i c a Ț i e.

$$\ln p(x_j; \theta) = -\frac{1}{2} \ln 2\pi - \frac{1}{2} \ln(1 + a\theta^2) - \frac{(x_j - \theta)^2}{2(1 + a\theta^2)}.$$

$$\left. \frac{\partial \ln p(x_j; \theta)}{\partial \theta} \right|_{\theta=0} = x_j$$

$$U(0) = \sum_{j=1}^n \left. \frac{\partial \ln p(x_j; \theta)}{\partial \theta} \right|_{\theta=0} = \sum_{j=1}^n X_j.$$

Regiunea critică va avea deci forma

$$W = \left\{ (x_1, \dots, x_n) \mid \sum_{j=1}^n x_j \geq k_\alpha \right\}.$$

Dacă H_0 este adevărată, $U(0) = \sum_{j=1}^n X_j$ urmează repartiția $N(0, n)$.

Deci regiunea critică local cea mai bună la pragul de semnificație α este

$$W = \left\{ (x_1, \dots, x_n) \mid \frac{1}{\sqrt{n}} \sum_{j=1}^n x_j \geq z_{1-\alpha} \right\},$$

unde $z_{1-\alpha}$ este cuantila de rang $(1 - \alpha)$ a repartiției $N(0, 1)$.

Dacă ipoteza alternativă $H_A: \{\theta = \delta\}$ ($\delta > 0$) este adevărată, variabila $U(0) = \sum_{j=1}^n X_j$ urmează repartiția $N(n\delta, n(1 + a\delta^2))$.

Funcția de putere a testului este

$$\pi(\delta) = P_\delta(W) = 1 - \Phi_{N(0, 1)} \left(\frac{\sqrt{n} z_{1-\alpha} - n\delta}{\sqrt{n(1 + a\delta^2)}} \right).$$

4.50. Fie X_1, \dots, X_n variabile independente, repartizate respectiv $\pi_\mu, \pi_{\mu\rho}, \dots, \pi_{\mu\rho^{n-1}}$ ($\mu > 0, \rho \geq 1$).

Să se determine teste local cele mai puternice pentru următoarele ipoteze:

- a) $H: \{\mu = \mu_0\}, H_A: \{\mu > \mu_0\}$ (ρ fiind cunoscut)
- b) $H': \{\rho = \rho_0\}, H'_A: \{\rho > \rho_0\}$ (μ fiind cunoscut).

Soluție. Variabila X_j ($j = 1, \dots, n$) are densitatea de repartiție

$$p(x_j; \mu, \rho) = \frac{(\mu\rho^{j-1})^{x_j}}{x_j!} \exp\{-\mu\rho^{j-1}\}, \quad j = 1, \dots, n.$$

$$a) \frac{\partial \ln p(x_j; \mu)}{\partial \mu} = \frac{x_j}{\mu} - \rho^{j-1}.$$

$$U(\mu_0) = \frac{1}{\mu_0} \sum_{j=1}^n (X_j - \mu_0 \rho^{j-1}).$$

Dacă H este adevărată, $M_{\mu_0}(U(\mu_0)) = 0$, $D_{\mu_0}^2(U(\mu_0)) = I(\mu_0) =$
 $= \frac{1}{\mu_0} \sum_{j=1}^n \rho^{j-1}.$

Variabila $\sum_{j=1}^n X_j$ este repartizată Poisson $\pi_{\mu_0 \sum_{j=1}^n \rho^{j-1}}$.

Regiunea critică local cea mai puternică are forma:

$$W = \left\{ (x_1, \dots, x_n) \mid \frac{1}{\mu_0} \sum_{j=1}^n (x_j - \mu_0 \rho^{j-1}) \geq k_\alpha \right\}.$$

Constanta k_α se determină utilizînd cuantilele repartiției Poisson $\pi_{\mu_0 \sum_{j=1}^n \rho^{j-1}}$, sau pentru n suficient de mare, utilizînd repartiția normală asimptotică a lui $U(\mu_0)$.

$$b) \frac{\partial \ln p(x_j; \rho)}{\partial \rho} = -\mu (j-1) \rho^{j-2} + \frac{(j-1)X_j}{\rho}.$$

$$U(\rho_0) = \frac{1}{\rho_0} \sum_{j=1}^n (j-1)X_j - \mu \sum_{j=1}^n \rho_0^{j-2} (j-1).$$

Dacă H' este adevărată, $M_{\rho_0}(U(\rho_0)) = 0$, $D_{\rho_0}^2(U(\rho_0)) = I(\rho_0) =$
 $= \mu \sum_{j=1}^n (j-1)^2 \cdot \rho_0^{j-3}.$

Regiunea critică local cea mai puternică are forma:

$$W = \{(x_1, \dots, x_n) \mid U(\rho_0) \geq k_\alpha\}.$$

4.51. Fie X_1, \dots, X_n variabile independente, identic repartizate, cu densitatea de repartiție $p(x; \varphi, \lambda)$ depinzînd de parametrul θ , $\theta = (\varphi, \lambda) \in \Theta_1 \times \Theta_2$. Considerăm ipoteza compusă $H_0: \{\varphi = \varphi_0, \lambda \in \Theta_2\}$ cu alternativa compusă $H: \{\varphi \neq \varphi_0, \lambda \in \Theta_2\}$. Fie S o statistică suficientă pentru λ . Dacă H_0 este adevărată, repartiția vectorului (X_1, \dots, X_n) condiționată de S nu depinde de λ .

Pentru o valoare alternativă $\theta = (\varphi, \lambda)$ se consideră regiunea critică

$$W = \left\{ (x_1, \dots, x_n) \left| \frac{p(x_1, \dots, x_n | s; \varphi, \lambda)}{p(x_1, \dots, x_n | s; \varphi_0)} \geq c_\alpha \right. \right\},$$

unde constanta c se determină din condiția

$$P_{\varphi_0}((X_1, \dots, X_n) \in W | S = s) = \alpha.$$

Dacă W astfel definită este aceeași pentru orice valoare alternativă $\theta = (\varphi, \lambda)$, ea se numește *regiunea critică uniform cea mai bună, bazată pe statistica suficientă S* . Asemenea teste se iau în considerare atunci când raportul de verosimilitate depinde de parametrul λ .

Problemă. Fie X_1, X_2 variabile independente, repartizate Poisson π_{μ_1} , respectiv π_{μ_2} , cu $\mu_1, \mu_2 > 0$. Utilizând noțiunile introduse mai sus, să se testeze ipoteza $H_0: \{\mu_1 = \varphi_0 \mu_2\}$ (cu φ_0 cunoscută), cu alternativa $H_A: \{\mu_1 > \varphi_0 \mu_2\}$, la pragul de semnificație α .

Soluție. Considerăm o nouă parametrizare: $\varphi = \mu_1/\mu_2$, $\lambda = \mu_2$. Deci $X_1 \sim \pi_{\varphi\lambda}$ și $X_2 \sim \pi_\lambda$, iar ipotezele pe care le avem de testat sînt $H_0: \{\varphi = \varphi_0, \lambda > 0\}$ cu alternativa $H_A: \{\varphi > \varphi_0, \lambda > 0\}$.

$$P(x_1, x_2; \varphi, \lambda) = \frac{1}{x_1! x_2!} e^{-\lambda(1+\varphi)} \lambda^{x_1+x_2} \varphi^{x_1}.$$

Pentru o valoare alternativă (φ, λ) se obține raportul de verosimilitate

$$\Lambda(x_1, x_2) = \frac{P(x_1, x_2; \varphi, \lambda)}{P(x_1, x_2; \varphi_0, \lambda)} = \left(\frac{\varphi}{\varphi_0} \right)^{x_1} e^{-\lambda(\varphi - \varphi_0)}.$$

Se observă că raportul de verosimilitate depinde de parametrul λ , ori pentru a construi regiunea critică uniform cea mai bună avem nevoie de o statistică ce nu depinde de λ .

$S = X_1 + X_2$ este o statistică suficientă pentru λ , care urmează o repartitie Poisson $\pi_{\lambda(1+\varphi)}$.

$$P(x_1, s; \varphi, \lambda) = \frac{1}{x_1!(s-x_1)!} \varphi^{x_1} e^{-\lambda(1+\varphi)} \lambda^s.$$

$$P(s; \varphi, \lambda) = \frac{1}{s!} \lambda^s (1+\varphi)^s e^{-\lambda(1+\varphi)}.$$

$$P(x_1 | s; \varphi, \lambda) = C_s^{x_1} \left(\frac{\varphi}{1+\varphi} \right)^{x_1} \left(1 - \frac{\varphi}{1+\varphi} \right)^{s-x_1}.$$

Deci repartiția lui X_1 condiționată de $S = s$ este $B_s\left(\frac{\varphi_1}{1 + \varphi_1}\right)$.

Fie o valoare alternativă (φ_1, λ_1) , $\varphi_1 > \varphi_0$. Testul bazat pe statistica suficientă S este dat de regiunea critică

$$W = \left\{ x_1 \left| \frac{P(x_1 | s; \varphi_1, \lambda_1)}{P(x_1 | s; \varphi_0)} \geq c_\alpha \right. \right\}, \text{ respectiv}$$

$$W = \left\{ x_1 \left| \left[\frac{\varphi_1}{\varphi_0} \right]^{x_1} \left[\frac{1 + \varphi_0}{1 + \varphi_1} \right]^s \geq c_\alpha \right. \right\} = \{x_1 | x_1 \geq k_\alpha\}.$$

Constanta k_α se determină așa încît $P_{\varphi_0}(W | S = s) = \alpha$. Dacă H_0 este adevărată, X_1 are repartiția condiționată de $S = s$ binomială $B_s\left(\frac{\varphi_0}{1 + \varphi_0}\right)$. Atunci k_α este dată de relația

$$\sum_{x=0}^{k_\alpha} C_s^x \left(\frac{\varphi_0}{1 + \varphi_0}\right)^x \left(1 - \frac{\varphi_0}{1 + \varphi_0}\right)^{s-x} = 1 - \alpha$$

sau, echivalent $\sum_{x=0}^{k_\alpha} C_s^x \varphi_0^x (1 + \varphi_0)^{-s} = 1 - \alpha$.

4.52. Considerăm următorul model al regresiei binare: fie X_1, \dots, X_n variabile independente, identic repartizate, luînd valorile 0 sau 1, pentru care se presupune că

$$\frac{P(X_j = 1)}{P(X_j = 0)} = \exp\{\gamma + \beta y_j\}, \quad j = 1, \dots, n,$$

unde β și γ sînt parametri necunoscuți, $\beta, \gamma \in R$, iar y_1, \dots, y_n sînt scalari cunoscuți. Să se construiască testul uniform cel mai puternic bazat pe o statistică suficientă (definit în problema 4.51) pentru verificarea ipotezei $H_0: \{\beta = \beta_0, \gamma \in R\}$ cu alternativa $H_A: \{\beta > \beta_0, \gamma \in R\}$, la pragul de semnificație α .

Soluție.

$$P(X_j = x_j; \gamma, \beta) = \frac{\exp(\gamma x_j + \beta x_j y_j)}{1 + \exp(\gamma + \beta y_j)}, \quad x_j \in \{0, 1\}, \quad j = 1, \dots, n$$

$$P(x_1, \dots, x_n; \gamma, \beta) = \frac{\exp\left(\gamma \sum_{j=1}^n x_j + \beta \sum_{j=1}^n x_j y_j\right)}{\prod_{j=1}^n (1 + \exp(\gamma + \beta y_j))}$$

Fie o valoare alternativă (β, γ) , $\beta > \beta_0$. Raportul de verosimilitate are forma :

$$\Lambda(x_1, \dots, x_n) = \frac{P(x_1, \dots, x_n; \gamma, \beta)}{P(x_1, \dots, x; \gamma, \beta_0)} = \\ = \exp\left((\beta - \beta_0) \sum_{j=1}^n x_j y_j\right) \prod_{j=1}^n \frac{1 + \exp(\gamma + \beta_0 y_j)}{1 + \exp(\gamma + \beta y_j)}.$$

Se vede că $\Lambda(x_1, \dots, x_n)$ depinde de parametrul liber γ , deci nu poate fi utilizat pentru construcția testului uniform cel mai puternic.

$(R, T) = \left(\sum_{j=1}^n X_j, \sum_{j=1}^n X_j y_j\right)$ este o statistică suficientă pentru (γ, β) .

$$P(R = r, T = t; \gamma, \beta) = \frac{c(r, t) \exp(\gamma r + \beta t)}{\prod_{j=1}^n (1 + \exp(\gamma + \beta y_j))},$$

unde $c(r, t)$ este numărul de submulțimi de câte r elemente ale lui (y_1, \dots, y_n) cu proprietatea că suma elementelor respectivei submulțimi este egală cu t . Observăm că $c(r, t)$ este egal cu coeficientul lui

$\xi_1^r \xi_2^t$ din produsul $\prod_{j=1}^n (1 + \xi_1 \xi_2^{y_j})$. Atunci

$$P(R = r; \gamma, \beta) = \frac{\sum_u c(r, u) \exp(\gamma r + \beta u)}{\prod_{j=1}^n (1 + \exp(\gamma + \beta y_j))}.$$

$$P(T = t | R = r; \gamma, \beta) = \frac{c(r, t) \exp(\beta t)}{\sum_u c(r, u) \exp(\beta u)}.$$

Această repartiție condiționată nu depinde de γ și ea va fi utilizată în construcția regiunii critice căutate, pentru $H_0: \{\beta = \beta_0, \gamma \in R\}$, $H_A: \{\beta > \beta_0, \gamma \in R\}$.

$$W = \left\{ (x_1, \dots, x_n) \left| \exp(\beta - \beta_0) t \frac{\sum_u c(r, u) \exp(\beta_0 u)}{\sum_u c(r, u) \exp(\beta u)} \geq c_\alpha \right. \right\},$$

unde $r = \sum_{j=1}^n x_j$, $t = \sum_{j=1}^n x_j y_j$. Echivalent, obținem :

$$W = \{(x_1, \dots, x_n) | t \geq k_\alpha\}.$$

Constanta k_α se determină așa încît $P_{\beta_0}(W | R = r) = \alpha$, respectiv așa încît

$$\frac{\sum_{u=0}^{k_\alpha} c(r, u) \exp(\beta_0 u)}{\sum_{u=0}^{t_{max}} c(r, u) \exp(\beta_0 u)} = 1 - \alpha.$$

4.53. Să se construiască testul raportului de verosimilitate (T.R.V.) pentru următoarele ipoteze simple cu alternative compuse:

- a) $H_0: \{p = p_0\}$, $H_A: \{p \neq p_0\}$ în cazul repartiției $P(X = 1) = p$,
 $P(X = 0) = 1 - p$, $0 < p < 1$.
- b) $H_0: \{\lambda = \lambda_0\}$, $H_A: \{\lambda \neq \lambda_0\}$ în cazul repartiției π_λ , $\lambda > 0$.
- c) $H_0: \{\mu = \mu_0\}$, $H_A: \{\mu \neq \mu_0\}$ în cazul repartiției $N(\mu, \sigma_0^2)$,
 $\mu \in R$, σ_0^2 cunoscut.

Soluție.

a) Fie X_1, \dots, X_n variabile independente, identic repartizate, cu

$$P(X_i = x_i) = p^{x_i} (1 - p)^{1 - x_i}, \quad x_i \in \{0, 1\}, \quad i = 1, \dots, n.$$

Raportul de verosimilitate este

$$\Lambda(x_1, \dots, x_n) = \frac{p_0^{\sum_{i=1}^n x_i} (1 - p_0)^{n - \sum_{i=1}^n x_i}}{\sup_p \left[p^{\sum_{i=1}^n x_i} (1 - p)^{n - \sum_{i=1}^n x_i} \right]}$$

E.V.M. pentru parametrul p este $\hat{p} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$. Atunci

$$\Lambda(x_1, \dots, x_n) = \left[\frac{p_0}{\bar{x}} \right]^{\sum_{i=1}^n x_i} \left[\frac{(1 - p_0)}{1 - \bar{x}} \right]^{n - \sum_{i=1}^n x_i}$$

Deși știm că variabila $\sum_{i=1}^n X_i$ urmează o repartiție binomială

$B_n(p)$, repartiția lui $\Lambda(X_1, \dots, X_n)$ este greu de calculat și vom utiliza rezultatul lui Neyman și Pearson privind repartiția asimptotică a variabilei $-2 \log \Lambda(X_1, \dots, X_n)$.

Pentru un $\alpha > 0$ fixat, fie $h_{1-\alpha}^2$ cuantila de rang $(1 - \alpha)$ a repartiției χ^2 cu un grad de libertate. T.R.V. pentru n suficient de mare este dat de regiunea critică

$$W = \{(x_1, \dots, x_n) \mid -2 \log \Lambda(x_1, \dots, x_n) \geq h_{1-\alpha}^2\}.$$

b) Fie X_1, \dots, X_n variabile independente, identic repartizate π_λ . Raportul de verosimilitate pentru H_0 cu alternativa H_A este

$$\Lambda(x_1, \dots, x_n) = \frac{\prod_{i=1}^n \frac{\lambda_0^{x_i}}{x_i!} \exp(-n\lambda_0)}{\sup_{\lambda} \prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!} \exp(-n\lambda)}$$

E.V.M. a parametrului λ este $\hat{\lambda} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ (problema 4.33).

Deci

$$\Lambda(x_1, \dots, x_n) = \exp(n(\bar{x} - \lambda_0)) \prod_{i=1}^n \left[\frac{\lambda_0}{\bar{x}} \right]^{x_i}.$$

T.R.V. pentru n suficient de mare, la pragul de semnificație α este dat de regiunea critică

$$W = \{(x_1, \dots, x_n) \mid -2 \log \Lambda(x_1, \dots, x_n) \geq h_{1-\alpha}^2\}.$$

c) Fie X_1, \dots, X_n variabile independente, identic repartizate $N(\mu, \sigma_0^2)$ cu σ_0^2 cunoscut. Raportul de verosimilitate pentru H_0 cu alternativa H_A este

$$\Lambda(x_1, \dots, x_n) = \frac{\exp\left\{-\frac{1}{2\sigma_0^2} \sum_{i=1}^n (x_i - \mu_0)^2\right\}}{\sup_{\mu \in R} \exp\left\{-\frac{1}{2\sigma_0^2} \sum_{i=1}^n (x_i - \mu)^2\right\}}$$

E.V.M. pentru μ este $\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$. Deci

$$\Lambda(x_1, \dots, x_n) = \exp\left\{-\frac{1}{2\sigma_0^2} \sum_{i=1}^n [(x_i - \mu_0)^2 - (x_i - \bar{x})^2]\right\}.$$

Regiunea critică a testului raportului de verosimilitate este

$$W = \{(x_1, \dots, x_n) \mid \Lambda(x_1, \dots, x_n) \leq k_\alpha\}.$$

Logaritmînd obținem forma echivalentă

$$W = \left\{ (x_1, \dots, x_n) \mid \left| \frac{\sqrt{n}(\bar{x} - \mu_0)}{\sigma_0} \right| \geq k'_\alpha \right\}$$

Dar variabila $\frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma_0}$ este repartizată $N(0, 1)$ dacă H_0 este adevărată. Deci pentru pragul de semnificație α fixat, constanta k'_α se ia egală cu $z_{1-\alpha/2}$, cuantila de rang $\left(1 - \frac{\alpha}{2}\right)$ a repartiției $N(0, 1)$.

4.54. Să se construiască T.R.V. pentru verificarea următoarelor ipoteze compuse cu alternative compuse:

- a) $H_0: \{\mu = \mu_0; \sigma^2 > 0\}$, $H_A: \{\mu \neq \mu_0; \sigma^2 > 0\}$ pentru repartiția $N(\mu, \sigma^2)$, $\mu \in R$, $\sigma^2 > 0$.
- b) $H_0: \{\mu \in R; \sigma^2 = \sigma_0^2\}$, $H_A: \{\mu \in R; \sigma^2 \neq \sigma_0^2\}$ pentru aceeași repartiție.

Soluție.

a) Spațiul parametrilor este $\Theta = \{(\mu, \sigma^2) \mid \mu \in R, \sigma^2 > 0\}$, iar $\Theta_0 = \{(\mu_0, \sigma^2) \mid \sigma^2 > 0\}$. Raportul de verosimilitate bazat pe n observații independente, identic repartizate $N(\mu, \sigma^2)$ este

$$\Lambda(x_1, \dots, x_n) = \frac{\sup_{\sigma^2 > 0} p(x_1, \dots, x_n; \mu_0, \sigma^2)}{\sup_{\mu \in R, \sigma^2 > 0} p(x_1, \dots, x_n; \mu, \sigma^2)},$$

unde $p(x_1, \dots, x_n; \mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right\}$.

Maximul de la numărător este atins pentru $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_0)^2$,

iar cel de la numitor pentru $\hat{\mu} = \bar{x}$, $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$.

Astfel, raportul de verosimilitate devine

$$\Lambda(x_1, \dots, x_n) = \frac{\left[\sum_{i=1}^n (x_i - \bar{x})^2 \right]^{n/2}}{\left[\sum_{i=1}^n (x_i - \mu_0)^2 \right]^{n/2}}.$$

T.R.V. pentru n suficient de mare, la pragul de semnificație α este dat de regiunea critică

$$W = \{(x_1, \dots, x_n) \mid -2 \log \Lambda(x_1, \dots, x_n) \geq h_{1-\alpha}^2\},$$

unde $h_{1-\alpha}^2$ este cuantila de rang $(1 - \alpha)$ a repartiției χ^2 cu un grad de libertate.

b) Spațiul parametrilor este $\Theta = \{(\mu, \sigma^2) \mid \mu \in R, \sigma^2 > 0\}$, iar $\Theta_0 = \{(\mu, \sigma_0^2) \mid \mu \in R\}$. Raportul de verosimilitate bazat pe n observații independente, identic repartizate $N(\mu, \sigma^2)$ este

$$\Lambda(x_1, \dots, x_n) = \frac{\sup_{\mu \in R} p(x_1, \dots, x_n; \mu, \sigma_0^2)}{\sup_{\mu \in R, \sigma^2 > 0} p(x_1, \dots, x_n; \mu, \sigma^2)}$$

Luind în considerare estimățiile de verosimilitate maximă corespunzătoare, obținem:

$$\Lambda(x_1, \dots, x_n) = \frac{p(x_1, \dots, x_n; \bar{x}, \hat{\sigma}_0^2)}{p(x_1, \dots, x_n; \bar{x}, \hat{\sigma}^2)} = \left[\frac{\hat{\sigma}^2}{\hat{\sigma}_0^2} \right]^{n/2} \exp\left(\frac{n}{2} \left(1 - \frac{\hat{\sigma}^2}{\hat{\sigma}_0^2} \right) \right),$$

$$\text{cu } \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2.$$

Pentru n suficient de mare, T.R.V. la pragul de semnificație α este dat de regiunea critică

$$W = \{(x_1, \dots, x_n) \mid -2 \log \Lambda(x_1, \dots, x_n) \geq h_{1-\alpha}^2\}.$$

4.55. Se consideră modelul regresiei liniare: X_1, \dots, X_n variabile independente, repartizate respectiv $\{N(y_i^t; \beta; \sigma_0^2), i = 1, \dots, n\}$, cu y_1, \dots, y_n vectori coloană s — dimensionali cunoscuți, σ_0^2 cunoscut, β vectorul coloană s — dimensional al parametrilor necunoscuți. Să se construiască T.R.V. pentru verificarea ipotezei $H_0: \{\beta = \beta_0\}$ cu alternativa $H_A: \{\beta \neq \beta_0\}$, la pragul de semnificație α .

$$\text{Soluție. } p(x_1, \dots, x_n; \beta) = \frac{1}{(2\pi\sigma_0^2)^{n/2}} \exp\left\{ -\frac{1}{2\sigma_0^2} \sum_{j=1}^n (x_j - y_j^t \beta)^2 \right\}.$$

Raportul de verosimilitate pentru H_0 cu alternativa H_A este

$$\Lambda(x_1, \dots, x_n) = \frac{p(x_1, \dots, x_n; \beta_0)}{\sup_{\beta} p(x_1, \dots, x_n; \beta)}$$

E. V.M. se obține rezolvând sistemul

$$\frac{\partial \ln p(x_1, \dots, x_n; \beta)}{\partial \beta_k} = 0, \quad k = 1, \dots, s.$$

Ea este $\hat{\beta} = \left(\sum_{j=1}^n y_j y_j^t \right)^{-1} \sum_{j=1}^n x_j y_j$.

Dintr-un calcul direct rezultă că

$$\begin{aligned} \Lambda(x_1, \dots, x_n) &= \frac{p(x_1, \dots, x_n; \beta_0)}{p(x_1, \dots, x_n; \hat{\beta})} = \\ &= \exp \left\{ -\frac{1}{2\sigma_0^2} (\hat{\beta} - \beta_0)^t \left(\sum_{j=1}^n y_j y_j^t \right) (\hat{\beta} - \beta_0) \right\}. \end{aligned}$$

Ne interesează repartiția exactă a lui $\Lambda(X_1, \dots, X_n)$, sau a lui $-2 \ln \Lambda(X_1, \dots, X_n)$.

Dacă H_0 este adevărată, $\hat{\beta}$ urmează o repartiție normală s -dimensională de medie β și matrice de covarianță $\sigma_0^2 \left(\sum_{j=1}^n y_j y_j^t \right)^{-1}$.

Atunci variabila aleatoare

$$\begin{aligned} T(X_1, \dots, X_n) &= -2 \ln \Lambda(X_1, \dots, X_n) = \\ &= \frac{1}{\sigma_0^2} (\hat{\beta} - \beta_0)^t \left(\sum_{j=1}^n y_j y_j^t \right) (\hat{\beta} - \beta_0) \end{aligned}$$

este repartizată χ^2 cu s grade de libertate.

La pragul de semnificație α , T.R.V. este dat de regiunea critică

$$W = \{(x_1, \dots, x_n) \mid T(x_1, \dots, x_n) \geq h_{s; 1-\alpha}^2\},$$

unde $h_{s; 1-\alpha}^2$ este cuantila de rang $(1 - \alpha)$ a repartiției χ^2 cu s grade de libertate.

4.56. Se consideră următorul model al regresiei poissoniene: X_1, \dots, X_n variabile independente, repartizate respectiv $\pi_{a+\theta y_j}$, $j = 1, \dots, n$, unde a este o constantă cunoscută, y_1, \dots, y_n sînt scalari cunoscuți, cu $\sum_{j=1}^n y_j = 0$, iar θ este un parametru real necunoscut, așa încît $a + \theta y_j > 0$, $j = 1, \dots, n$. Să se construiască T.R.V. pentru verificarea ipotezei $H_0: \{\theta = 0\}$, cu alternativa $H_A: \{\theta \neq 0\}$.

Soluție. $p(x_1, \dots, x_n; \theta) = \prod_{j=1}^n \left[\frac{(a + \theta y_j)^{x_j} \exp(-(a + \theta y_j))}{x_j!} \right]$.

Ecuția de verosimilitate maximă $\frac{\partial \ln p(x_1, \dots, x_n; \theta)}{\partial \theta} = 0$ revine la

$\sum_{j=1}^n \frac{x_j y_j}{a + \theta y_j} = 0$. Ea nu poate fi rezolvată prin metode exacte și de aceea vom construi T.R.V. bazat pe o statistică asimptotic echivalentă cu $\Lambda(X_1, \dots, X_n)$.

În general, $\Lambda(x_1, \dots, x_n) = \frac{p(x_1, \dots, x_n; \theta_0)}{\sup_{\theta} p(x_1, \dots, x_n; \theta)} = \frac{p(x_1, \dots, x_n; \theta_0)}{p(x_1, \dots, x_n; \hat{\theta})}$.

$T(X_1, \dots, X_n) = -2 \ln \Lambda(X_1, \dots, X_n) = 2(\ln p(X_1, \dots, X_n; \hat{\theta}) - \ln p(X_1, \dots, X_n; \theta_0))$

În ipoteza că $\theta_0 \in \text{Int } \Theta$, luând în considerare proprietățile estimației de verosimilitate maximă și utilizând dezvoltarea după formula lui Taylor a lui $\ln p(x_1, \dots, x_n; \hat{\theta})$ obținem:

$T(X_1, \dots, X_n) \simeq -n(\hat{\theta} - \theta_0)^2 \left[\frac{1}{n} \frac{\partial^2 U(\theta)}{\partial \theta^2} \Big|_{\theta=\theta_0} \right]$, unde

$U(\theta) = \frac{\partial \ln p(X_1, \dots, X_n; \theta)}{\partial \theta}$.

Dar $\frac{1}{n} \frac{\partial U(\theta)}{\partial \theta} \Big|_{\theta=\theta_0}$ converge în probabilitatea P_{θ_0} la $-\frac{I(\theta_0)}{n}$, unde $I(\theta_0)$ este informația lui Fischer.

Deci $T(X_1, \dots, X_n) \simeq (\hat{\theta} - \theta_0)^2 I(\theta_0)$, care este repartizată asimptotic χ^2 cu un grad de libertate.

Dar $\hat{\theta} - \theta_0 \simeq U(\theta_0)/I(\theta_0)$ și deci $T = U^2(\theta_0)/I(\theta_0)$ este o statistică ce-l aproximează pe $-2 \ln \Lambda(X_1, \dots, X_n)$ pentru n suficient de mare.

În cazul problemei noastre, $\theta_0 = 0$ și $U(0) = \frac{1}{a} \sum_{j=1}^n X_j y_j$, iar

$$I(0) = M_0 \left(- \frac{\partial^2 \ln p(X_1, \dots, X_n; \theta)}{\partial \theta^2} \Big|_{\theta=0} \right) = \\ = \frac{1}{a^2} \sum_{j=1}^n y_j^2 M_0(X_j) = \frac{1}{a} \sum_{j=1}^n y_j^2.$$

$$\text{Deci } T' = U^2(0)/I(0) = \frac{1}{a \sum_{j=1}^n y_j^2} \left(\sum_{j=1}^n X_j y_j \right)^2.$$

Pe baza lui T' se poate construi un test asimptotic echivalent cu testul raportului de verosimilitate.

4.57. Fie $X_1, \dots, X_n, Y_1, \dots, Y_n$ variabile aleatoare independente în totalitate, X_1, \dots, X_n identic repartizate $N(\mu_1, \sigma^2)$, iar Y_1, \dots, Y_n identic repartizate $N(\mu_2, \sigma^2)$, cu parametrii μ_1, μ_2 și σ^2 necunoscuți ($\mu_1, \mu_2 \in \mathbb{R}, \sigma^2 > 0$). Să se construiască T.R.V. pentru verificarea ipotezei egalității mediilor μ_1 și μ_2 .

Soluție. Spațiul parametrilor este $\Theta = \{\theta = (\mu_1, \mu_2, \sigma^2) \mid \mu_1, \mu_2 \in \mathbb{R}, \sigma^2 > 0\}$ iar $\Theta_0 = \{\theta \in \Theta \mid \mu_1 = \mu_2\}$; ne propunem să verificăm ipoteza $H_0: \{\theta \in \Theta_0\}$ cu alternativa $H_A: \{\theta \in \Theta \setminus \Theta_0\}$.

$$p(x_1, \dots, x_n, y_1, \dots, y_n; \theta) = \frac{1}{(2\pi)^n \sigma^{2n}} \exp \left\{ -\frac{1}{2\sigma^2} \left(\sum_{i=1}^n (x_i - \mu_1)^2 + \sum_{i=1}^n (y_i - \mu_2)^2 \right) \right\}.$$

$$\Lambda(x_1, \dots, x_n, y_1, \dots, y_n) = \frac{\sup_{\theta \in \Theta_0} p(x_1, \dots, x_n, y_1, \dots, y_n; \theta)}{\sup_{\theta \in \Theta} p(x_1, \dots, x_n, y_1, \dots, y_n; \theta)}.$$

Maximul de la numărător este atins pentru

$$\hat{\mu} = \frac{1}{2} (\bar{x} + \bar{y}) = \frac{1}{2n} \sum_{i=1}^n (x_i + y_i)$$

$$\hat{\sigma}^2 = \frac{1}{2n} \left(\sum_{i=1}^n (x_i - \hat{\mu})^2 + \sum_{i=1}^n (y_i - \hat{\mu})^2 \right),$$

iar cel de la numitor este atins pentru $\hat{\mu}_1 = \bar{x}, \hat{\mu}_2 = \bar{y}$,

$$\hat{\sigma}^2 = \frac{1}{2n} \left(\sum_{i=1}^n (x_i - \bar{x})^2 + \sum_{i=1}^n (y_i - \bar{y})^2 \right).$$

Atunci

$$\Lambda(x_1, \dots, x_n, y_1, \dots, y_n) = \left[\frac{\hat{\sigma}^2}{\sigma^2} \right]^n.$$

$$\text{Dar } \hat{\sigma}^2 = \hat{\sigma}^2 + \frac{1}{2} (\bar{x} - \bar{y})^2.$$

Regiunea critică a testului este de forma

$$W = \left\{ (x_1, \dots, x_n, y_1, \dots, y_n) \left| \left| \frac{\bar{x} - \bar{y}}{\hat{\sigma}} \right| > k_\alpha \right. \right\}.$$

Constanta k_α se determină așa încît $P_0(W) = \alpha$ pentru orice $\theta \in \Theta_0$.

Dar variabila aleatoare $\frac{1}{\hat{\sigma}} \sqrt{\frac{n-1}{2}} (\bar{X} - \bar{Y})$ urmează o repartiție Student cu $(2n - 2)$ grade de libertate. Deci

$$W = \left\{ (x_1, \dots, x_n, y_1, \dots, y_n) \left| \frac{\bar{x} - \bar{y}}{\hat{\sigma}} < \sqrt{\frac{2}{n-1}} t_{2n-2; \alpha/2} \right. \right\} \cup \\ \cup \left\{ (x_1, \dots, x_n, y_1, \dots, y_n) \left| \frac{\bar{x} - \bar{y}}{\hat{\sigma}} > \sqrt{\frac{2}{n-1}} t_{2n-2; 1-\alpha/2} \right. \right\}.$$

unde $t_{2n-2; \alpha}$ este cuantila de rang α a repartiției $t(2n - 2)$.

4.58. Se consideră o populație supusă influenței a doi factori A și B avînd respectiv nivelele A_1, \dots, A_r și B_1, \dots, B_s . Fie θ_{ij} probabilitatea exercitării influenței simultane a celor doi factori la nivelele (A_i, B_j) , $\theta_{ij} \geq 0$ oricare ar fi i, j , $\sum_{i,j} \theta_{ij} = 1$.

Se fac n observații independente asupra populației și se notează cu n_{ij} numărul de observații ce s-au dovedit a fi influențate de (A_i, B_j) .

Notăm $\theta_i = \sum_{j=1}^s \theta_{ij}$, $\theta_j = \sum_{i=1}^r \theta_{ij}$, $i = 1, \dots, r$, $j = 1, \dots, s$.

Să se construiască T.R.V. pentru verificare a ipotezei că cei doi factori acționează în mod independent $H_0: \{\theta_{ij} = \theta_i \cdot \theta_j \text{ oricare ar fi } i, j\}$.

Soluție. Parametrul necunoscut este $\theta = (\theta_{11}, \dots, \theta_{rs})$, spațiul parametrilor este $\Theta = \left\{ \theta \mid 0 \leq \theta_{ij} \leq 1, \sum_{i,j} \theta_{ij} = 1 \right\}$, iar

$$\Theta_0 = \left\{ \theta \in \Theta \mid \theta_{ij} = \theta_i \cdot \theta_j, \sum_i \theta_i = 1, \sum_j \theta_j = 1 \right\}.$$

Vectorul $x = (n_{11}, \dots, n_{rs})$ cu $\sum_{i,j} n_{ij} = n$ urmează o repartiție multinomială de parametru θ .

$$P(x; \theta) = k \prod_{i,j} \theta_{ij}^{n_{ij}}$$

Raportul de verosimilitate este $\Lambda(x) = \frac{\sup_{\theta \in \Theta_0} p(x; \theta)}{\sup_{\theta \in \Theta} p(x; \theta)}$

$$\text{Dacă } \theta \in \Theta_0, p(x; \theta) = k \sum_{i,j} (\theta_i)^{n_i} (\theta_j)^{n_j},$$

$$\text{unde } n_i = \sum_{j=1}^s n_{ij}, n_j = \sum_{i=1}^r n_{ij}.$$

Maximul de la numărător este atins pentru $\hat{\theta}_i = \frac{n_i}{n}, \hat{\theta}_j = \frac{n_j}{n}$

iar cel de la numitor este atins pentru $\hat{\theta}_{ij} = \frac{n_{ij}}{n}$.

$$\text{Atunci } \Lambda(x) = n^n \prod_{i,j} \left[\frac{n_{ij}}{n_i n_j} \right]^{n_{ij}}$$

Determinarea repartiției exacte a statisticii Λ este foarte dificilă și de aceea utilizăm aproximarea asimptotică a acestei repartiții.

Θ are $rs - 1$ parametri liberi, iar Θ_0 are $r + s - 2$. Deci pentru definirea lui Θ_0 s-au impus $(rs - 1) - (r + s - 2) = (r - 1)(s - 1)$ restricții neredundante. $-2 \log \Lambda$ urmează o repartiție χ^2 cu $(r - 1)(s - 1)$ grade de libertate și regiunea critică a testului raportului de verosimilitate la pragul de semnificație α este

$$W = \{x \mid -2 \log \Lambda(x) \geq h_{(r-1)(s-1); 1-\alpha}^2\}.$$

4.59. Să se construiască testul secvențial al raportului probabilităților (T.S.R.P.) pentru verificarea următoarelor ipoteze simple cu alternative simple:

- a) $H_0: \{p = p_0\}, H_1: \{p = p_1\}$ pentru repartiția $P(X=1) = p, P(X=0) = 1 - p, 0 < p < 1$.
- b) $H_0: \{\lambda = \lambda_0\}, H_1: \{\lambda = \lambda_1\}$ pentru repartiția $\pi_\lambda, \lambda > 0$.
- c) $H_0: \{\mu = \mu_0\}, H_1: \{\mu = \mu_1\}$ pentru repartiția $N(\mu, \sigma_0^2)$ cu $\mu \in R, \sigma_0^2$ cunoscut.
- d) $H_0: \{\sigma^2 = \sigma_0^2\}, H_1: \{\sigma^2 = \sigma_1^2\}$ pentru repartiția $N(\mu_0, \sigma^2)$ cu $\sigma^2 > 0, \mu_0$ cunoscut.

Soluție.

a) Fie X_1, \dots, X_n variabile independente, identic repartizate cu $P(X = x) = p^x(1-p)^{1-x}$, $x \in \{0, 1\}$ și fie $k_n = \sum_{i=1}^n x_i$ numărul de „succese” în n observații (x_1, \dots, x_n) . Raportul probabilităților pentru ipotezele H_0, H_1 este:

$$R_n = \frac{p_1^n}{p_0^n} = \left[\frac{p_1}{p_0} \right]^{k_n} \left[\frac{1-p_1}{1-p_0} \right]^{n-k_n}$$

T.S.R.P. va fi dat la momentul n de regiunile:

$$W_n^a = \left\{ \log R_n \leq \log \frac{\beta}{1-\alpha} \right\}$$

$$W_n^r = \left\{ \log R_n \geq \log \frac{1-\beta}{\alpha} \right\}$$

$$W = \left\{ \log \frac{\beta}{1-\alpha} < \log R_n < \log \frac{1-\beta}{\alpha} \right\},$$

sau

$W_n^a = \left\{ (x_1, \dots, x_n) \mid \sum_{i=1}^n x_i = k_n \leq a_n \right\}$ regiunea de acceptare a lui H_0

$W_n^r = \left\{ (x_1, \dots, x_n) \mid k_n \geq r_n \right\}$ regiunea de respingere a lui H_0

$W_n = \left\{ (x_1, \dots, x_n) \mid a_n < k_n < r_n \right\}$ regiunea de continuare a selecției,

unde $a_n = -h_1 + sn$, $r_n = h_2 + sn$, cu

$$h_1 = \frac{\log \frac{1-\alpha}{\beta}}{\log \frac{p_1}{p_0} - \log \frac{1-p_1}{1-p_0}}, \quad h_2 = \frac{\log \frac{1-\beta}{\alpha}}{\log \frac{p_1}{p_0} - \log \frac{1-p_1}{1-p_0}}$$

$$s = \frac{\log \frac{1-p_0}{1-p_1}}{\log \frac{p_1}{p_0} - \log \frac{1-p_1}{1-p_0}}$$

b) Fie X_1, \dots, X_n variabile independente, identic repartizate π_λ , și fie $k_n = \sum_{i=1}^n x_i$. Raportul probabilităților pentru ipotezele H_0, H_1 este:

$$R_n = \frac{p_1^n}{p_0^n} = \left[\frac{\lambda_1}{\lambda_0} \right]^{k_n} \cdot \exp(-n(\lambda_1 - \lambda_0))$$

T.S.R.P. va fi dat la momentul n de regiunile

$W_n^a = \{(x_1, \dots, x_n) \mid k_n \leq a_n\}$ regiunea de acceptare a lui H_0

$W_n^r = \{(x_1, \dots, x_n) \mid k_n \geq r_n\}$ regiunea de respingere a lui H_0

$W_n = \{(x_1, \dots, x_n) \mid a_n < k_n < r_n\}$ regiunea de continuare a selecției, unde

$$a_n = \frac{\ln \frac{\beta}{1-\alpha}}{\ln \frac{\lambda_1}{\lambda_0}} + n \frac{\lambda_1 - \lambda_0}{\ln \frac{\lambda_1}{\lambda_0}}$$

$$r_n = \frac{\ln \frac{1-\beta}{\alpha}}{\ln \frac{\lambda_1}{\lambda_0}} + n \frac{\lambda_1 - \lambda_0}{\ln \frac{\lambda_1}{\lambda_0}}$$

c) Fie X_1, \dots, X_n variabile independente, identic repartizate $N(\mu, \sigma_0^2)$. Raportul probabilităților pentru ipotezele H_0 și H_1 este:

$$R_n(x_1, \dots, x_n) = \frac{p(x_1, \dots, x_n; \mu_1)}{p(x_1, \dots, x_n; \mu_0)}, \text{ unde}$$

$$p(x_1, \dots, x_n; \mu) = \frac{1}{(2\pi\sigma_0^2)^{n/2}} \exp \left\{ -\frac{1}{2\sigma_0^2} \sum_{i=1}^n (x_i - \mu)^2 \right\}$$

$$R_n(x_1, \dots, x_n) = \exp \left\{ (\mu_1 - \mu_0) \left(\sum_{i=1}^n x_i - \frac{n(\mu_1 + \mu_0)}{2\sigma_0^2} \right) \right\}.$$

T.S.R.P. este dat la momentul n de regiunile:

$$W_n^a = \left\{ (x_1, \dots, x_n) \mid R_n(x_1, \dots, x_n) \leq \frac{\beta}{1-\alpha} \right\}$$

$$W_n^r = \left\{ (x_1, \dots, x_n) \mid R_n(x_1, \dots, x_n) \geq \frac{1-\beta}{\alpha} \right\}$$

$$W_n = \left\{ (x_1, \dots, x_n) \mid \frac{\beta}{1-\alpha} < R_n(x_1, \dots, x_n) < \frac{1-\beta}{\alpha} \right\}.$$

Logaritmînd, obținem regiunile echivalente

$$W_n^a = \left\{ (x_1, \dots, x_n) \mid \sum_{i=1}^n x_i \leq a_n \right\} \text{ regiunea de acceptare a lui } H_0$$

$$W_n^r = \left\{ (x_1, \dots, x_n) \mid \sum_{i=1}^n x_i \geq r_n \right\} \text{ regiunea de respingere a lui } H_0$$

$$W_n = \left\{ (x_1, \dots, x_n) \mid a_n < \sum_{i=1}^n x_i < r_n \right\} \text{ regiunea de continuare a selecției,}$$

unde $a_n = -h_1 + sn$, $r_n = h_2 + sn$, cu

$$h_1 = \frac{\sigma_0^2}{\mu_1 - \mu_0} \ln \frac{1 - \alpha}{\beta}, \quad h_2 = \frac{\sigma_0^2}{\mu_1 - \mu_0} \ln \frac{1 - \beta}{\alpha}, \quad s = \frac{\mu_1 + \mu_0}{2}.$$

d) Fie X_1, \dots, X_n variabile independente, identic repartizate $N(\mu_0, \sigma^2)$. Raportul probabilităților pentru ipotezele H_0 și H_1 este

$$\begin{aligned} R_n(x_1, \dots, x_n) &= \frac{p(x_1, \dots, x_n; \sigma_1^2)}{p(x_1, \dots, x_n; \sigma_0^2)} = \\ &= \left[\frac{\sigma_0^2}{\sigma_1^2} \right] \cdot \exp \left[\left(\frac{1}{2\sigma_0^2} - \frac{1}{2\sigma_1^2} \right) \sum_{i=1}^n (x_i - \mu_0)^2 \right]. \end{aligned}$$

T.S.R.P. la momentul n este dat de regiunile

$W_n^a = \left\{ (x_1, \dots, x_n) \mid \sum_{i=1}^n (x_i - \mu_0)^2 \leq a_n \right\}$ regiunea de acceptare a lui H_0

$W_n^r = \left\{ (x_1, \dots, x_n) \mid \sum_{i=1}^n (x_i - \mu_0)^2 \geq r_n \right\}$ regiunea de respingere a lui H_0

$W_n = \left\{ (x_1, \dots, x_n) \mid a_n < \sum_{i=1}^n (x_i - \mu_0)^2 < r_n \right\}$ regiunea de continuare a selecției, unde

$a_n = -h_1 + sn$, $r_n = h_2 + sn$, cu

$$h_1 = \frac{\sigma_1^2 \sigma_0^2}{\sigma_1^2 - \sigma_0^2} \ln \frac{1 - \alpha}{\beta}, \quad h_2 = \frac{\sigma_1^2 \sigma_0^2}{\sigma_1^2 - \sigma_0^2} \ln \frac{1 - \beta}{\alpha}, \quad s = \frac{\sigma_1^2 \sigma_0^2}{\sigma_1^2 - \sigma_0^2} \ln \frac{\sigma_1}{\sigma_0}.$$

4.60. (testul t). Să se construiască testul simplu (bazat pe un interval de încredere) pentru următoarele ipoteze asupra mediei m a unei repartiții normale $N(m, \sigma^2)$, cu σ^2 necunoscut:

$H_1: \{m = m_0, \sigma^2 > 0\}$ cu alternativa

$H'_1: \{m \in (-\infty, \infty) \mid m \neq m_0, \sigma^2 > 0\}$

$H_2: \{m = m_0, \sigma^2 > 0\}$ cu alternativa $H'_2: \{m > m_0, \sigma^2 > 0\}$

$H_3: \{m = m_0, \sigma^2 > 0\}$ cu alternativa $H'_3: \{m < m_0, \sigma^2 > 0\}$

$H_4: \{m_1 \leq m \leq m_2, \sigma^2 > 0\}$ cu alternativa

$H'_4: \{m < m_1, \sigma^2 > 0\} \cup \{m > m_2, \sigma^2 > 0\}$

Aplicație. a) Se consideră o selecție bernoulliană de volum $n = 16$ dintr-o populație normală, valorile observate fiind prezentate în continuare: 16,71; 18,11; 11,19; 11,98; 24,86; 20,18; 17,49; 9,23; 15,35; 21,12; 17,23; 18,67; 23,17; 17,05; 15,43; 16,82.

Să se testeze ipoteza $H: \{m = 10, \sigma^2 > 0\}$ cu alternativa $H': \{m > 10, \sigma^2 > 0\}$, la pragul de semnificație $\alpha = 0,05$.

b) Se consideră o selecție bernoulliană de volum $n = 16$ dintr-o populație normală, valorile observate fiind prezentate în continuare: 7,63; 6,37; 6,97; 5,28; 6,74; 6,74; 12,06; 7,53; 8,81; 13,58; 12,64; 10,71; 5,41; 7,46; 4,39; 11,85.

Să se testeze ipoteza $H: \{7 \leq m \leq 9,5, \sigma^2 > 0\}$ cu alternativa $H': \{m < 7, \sigma^2 > 0\} \cup \{m > 9,5, \sigma^2 > 0\}$ la pragul de semnificație α .

Soluție. Parametrul necunoscut σ^2 va fi estimat prin $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$, care este o estimăție nedeplasată, construită pe baza selecției bernoulliene n -dimensionale (X_1, \dots, X_n) asupra variabilei X repartizată $N(m, \sigma^2)$.

Variabila aleatoare cu ajutorul căreia vom construi teste simple pentru verificarea ipotezelor considerate este $t = \frac{\sqrt{n}(\bar{X} - m)}{S}$, care urmează o repartiție Student cu $(n - 1)$ grade de libertate.

În tabelul următor prezentăm regiunile critice pentru fiecare caz considerat.

Ipoteza H_i	Alternativa H'_i	Regiunea critică pentru H_i la pragul de semnificație α
$m = m_0$ $\sigma^2 > 0$	$m \in (-\infty, \infty)$ $m \neq m_0$ $\sigma^2 > 0$	$\left\{ (x_1, \dots, x_n) \mid \frac{\sqrt{n}(\bar{x} - m_0)}{S} < t_{n-1; \frac{\alpha}{2}} \right\} \cup$ $\cup \left\{ (x_1, \dots, x_n) \mid \frac{\sqrt{n}(\bar{x} - m_0)}{S} > t_{n-1; 1 - \frac{\alpha}{2}} \right\}$
$m = m_0$ $\sigma^2 > 0$	$m > m_0$ $\sigma^2 > 0$	$\left\{ (x_1, \dots, x_n) \mid \frac{\sqrt{n}(\bar{x} - m_0)}{S} > t_{n-1; 1 - \alpha} \right\}$
$m = m_0$ $\sigma^2 > 0$	$m < m_0$ $\sigma^2 > 0$	$\left\{ (x_1, \dots, x_n) \mid \frac{\sqrt{n}(\bar{x} - m_0)}{S} < t_{n-1; \alpha} \right\}$
$m_1 < m < m_2$ $\sigma^2 > 0$	$\{m < m_1\} \cup$ $\cup \{m > m_2\}$ $\sigma^2 > 0$	$\left\{ (x_1, \dots, x_n) \mid \frac{\sqrt{n}(\bar{x} - m_1)}{S} < t_{n-1; \alpha_1} \right\} \cup$ $\cup \left\{ (x_1, \dots, x_n) \mid \frac{\sqrt{n}(\bar{x} - m_2)}{S} > t_{n-1; 1 - \alpha_2} \right\}$ cu $\alpha_1 + \alpha_2 = \alpha$.

În acest tabel $t_{n-1; \alpha}$ reprezintă cuantila de rang α a repartiției $t(n-1)$.

Aplicație

a) $\bar{x} = 17,16$; $S^2 = 17,07$; $S = 4,13$

$$t_{calc} = \frac{\sqrt{n}(\bar{x} - m_0)}{S} = 6,9$$

$$t_{15; 0,95} = 1,753.$$

Cum $t_{calc} > t_{15; 0,95}$, decidem să respingem ipoteza $H: \{m = 10\}$ și să acceptăm alternativa $H': \{m > 10\}$.

b) $\bar{x} = 8,38$; $S^2 = 8,82$; $S = 2,88$

$$\frac{\sqrt{n}(\bar{x} - m_1)}{S} = 1,9; \quad \frac{\sqrt{n}(\bar{x} - m_2)}{S} = -1,55$$

$$-t_{15; 0,025} = t_{15; 0,975} = 2,131.$$

Cum $\frac{\sqrt{n}(\bar{x} - m_1)}{S} > t_{15; 0,025}$ și $\frac{\sqrt{n}(\bar{x} - m_2)}{S} < t_{15; 0,975}$ decidem să acceptăm ipoteza $H: \{7 \leq m \leq 9,5\}$.

4.61 (testul X). Se consideră o selecție bernoulliană n -dimensională asupra unei variabile repartizate normal $N(m, \sigma^2)$ cu ambii parametri necunoscuți. Să se construiască testul simplu (bazat pe un interval de încredere) pentru verificarea următoarelor ipoteze:

$H_1: \{m \in R, \sigma^2 = \sigma_0^2\}$ cu alternativa $H'_1: \{m \in R, \sigma^2 > 0, \sigma \neq \sigma_0^2\}$

$H_2: \{m \in R, \sigma^2 = \sigma_0^2\}$ cu alternativa $H'_2: \{m \in R, \sigma^2 > \sigma_0^2\}$

$H_3: \{m \in R, \sigma^2 = \sigma_0^2\}$ cu alternativa $H'_3: \{m \in R, \sigma^2 < \sigma_0^2\}$.

Aplicație. a) Pentru selecția de volum 16 dată în problema 4.60 a) de la acest paragraf să se verifice ipoteza $H: \{m \in R, \sigma^2 = 18\}$ cu alternativa $H': \{m \in R, \sigma^2 > 18\}$ la pragul de semnificație $\alpha = 0,05$.

b) Pentru selecția de volum 16 dată în problema 4.60. b) de la acest paragraf să se verifice ipoteza $H: \{m \in R, \sigma^2 = 12\}$ cu alternativa $H': \{m \in R, \sigma^2 > 0, \sigma^2 \neq 12\}$.

Soluție. Fie $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ o estimatie nedeplasată pentru σ^2 . Variabila $\frac{(n-1)S^2}{\sigma^2}$ urmează o repartiție χ^2 cu $n-1$ grade de libertate.

Regiunile critice pentru ipotezele considerate sînt date în tabelul următor :

Ipoieza H_i	Alternativa H'_i	Regiunea critică pentru H_i la pragul de semnificație α
$\sigma^2 = \sigma_0^2$ $m \in R$	$\sigma^2 > 0$ $\sigma^2 \neq \sigma_0^2$ $m \in R$	$\left\{ (x_1, \dots, x_n) \mid \frac{(n-1)S^2}{\sigma_0^2} < h^2_{n-1; \frac{\alpha}{2}} \right\} \cup$ $\cup \left\{ (x_1, \dots, x_n) \mid \frac{(n-1)S^2}{\sigma_0^2} > h^2_{n-1; 1 - \frac{\alpha}{2}} \right\}$
$\sigma^2 = \sigma_0^2$ $m \in R$	$\sigma^2 > \sigma_0^2$ $m \in R$	$\left\{ (x_1, \dots, x_n) \mid \frac{(n-1)S^2}{\sigma_0^2} > h^2_{n-1; 1-\alpha} \right\}$
$\sigma^2 = \sigma_0^2$ $m \in R$	$\sigma^2 < \sigma_0^2$ $m \in R$	$\left\{ (x_1, \dots, x_n) \mid \frac{(n-1)S^2}{\sigma_0^2} < h^2_{n-1; \alpha} \right\}$

În acest tabel $h^2_{n-1; \alpha}$ reprezintă cuantila de rang α a repartiției χ^2 cu $(n-1)$ grade de libertate.

Aplicație. a) $S^2 = 17,07$; $\frac{(n-1)S^2}{\sigma_0^2} = 14,22$; $h^2_{15; 0,95} = 7,26$.

Decidem că respingem ipoteza H și să acceptăm alternativa $H' : \{m \in R, \sigma^2 > 18\}$.

b) $S^2 = 8,28$; $\frac{(n-1)S^2}{\sigma_0^2} = 10,35$; $h^2_{15; 0,975} = 6,26$.

Decidem să respingem ipoteza H și să acceptăm $H' : \{m \in R, \sigma^2 > 0, \sigma^2 \neq 12\}$.

4.62. Să se construiască testul χ^2 pentru compararea a r porții.

Aplicație. Apartenența la una din clasele C_1, C_2, C_3, C_4 a 500 de indivizi supuși la trei tratamente diferite T_1, T_2, T_3 este dată în tabelul de mai jos :

Tratament	Volumul eșantionului	C_1	C_2	C_3	C_4	Total
T_1	150	12	48	59	31	$n_1 = 150$
T_2	200	20	65	71	44	$n_2 = 200$
T_3	150	18	37	70	25	$n_3 = 150$
Total	$n = 500$	$n_{.1} = 50$	$n_{.2} = 150$	$n_{.3} = 200$	$n_{.4} = 100$	

Să se examineze egalitatea proporțiilor pe clase a indivizilor în cele trei tratamente.

Soluție. Considerăm r populații supuse la tratamente diferite, investigațiile făcute asupra fiecărui individ putînd oferi s răspunsuri C_1, C_2, \dots, C_s . Datele observate se ordonează ca în tabelul de mai jos :

Tratamentul	Volumul eșantionului	Nr. observațiilor				Total
		C_1	C_2	...	C_s	
T_1	n_1	n_{11}	n_{12}	...	n_{1s}	$n_{1\cdot}$
T_2	n_2	n_{21}	n_{22}	...	n_{2s}	$n_{2\cdot}$
...
T_r	n_r	n_{r1}	n_{r2}	...	n_{rs}	$n_{r\cdot}$
Total	n	$n_{\cdot 1}$	$n_{\cdot 2}$...	$n_{\cdot s}$	

Fie p_{ij} probabilitatea de apariție a răspunsului j ($j = 1, 2, \dots, s$) la un individ din populația i ($i = 1, 2, \dots, r$). $\sum_{j=1}^s p_{ij} = 1$ pentru orice $i = 1, 2, \dots, r$.

Cum p_{ij} sînt necunoscuți, vom construi în primul rînd estimațiile lor de verosimilitate maximă.

Introducem notațiile :

$$p_{i\cdot} = \sum_{j=1}^s p_{ij}, \quad p_{\cdot j} = \sum_{i=1}^r p_{ij}$$

$$n_{i\cdot} = \sum_{j=1}^s n_{ij}, \quad n_{\cdot j} = \sum_{i=1}^r n_{ij}$$

În ipoteza că observațiile făcute sînt independente avem $p_{ij} = p_{i\cdot} p_{\cdot j}$, iar funcția de verosimilitate este

$$L = \prod_{i=1}^r \prod_{j=1}^s (p_{i\cdot} p_{\cdot j})^{n_{ij}} = \prod_{i=1}^r (p_{i\cdot})^{n_{i\cdot}} \prod_{j=1}^s (p_{\cdot j})^{n_{\cdot j}}$$

Ținînd seama de faptul că $\sum_{i=1}^r p_{i\cdot} = 1$, obținem $p_{r\cdot} = 1 - \sum_{i=1}^{r-1} p_{i\cdot}$, iar funcția de verosimilitate devine

$$L = \left(1 - \sum_{i=1}^{r-1} p_{i\cdot}\right)^{n_{r\cdot}} \prod_{i=1}^{r-1} (p_{i\cdot})^{n_{i\cdot}} \prod_{j=1}^s (p_{\cdot j})^{n_{\cdot j}}$$

Logaritmînd, obținem

$$\log L = n_r \cdot \log \left(1 - \sum_{i=1}^{r-1} p_i \right) + \sum_{i=1}^{r-1} n_i \cdot \log p_i + \sum_{j=1}^s n_j \cdot \log p_j$$

Considerînd ca necunoscute $p_i, i = 1, \dots, (r-1)$, sistemul de verosimilitate maximă conduce la

$$\frac{n_i}{p_i} = \frac{n_r}{p_r}, i = 1, 2, \dots, (r-1).$$

Din proprietățile șirului de rapoarte egale obținem $\hat{p}_i = \frac{n_i}{n}, i = 1, 2, \dots, r$.

Analog se obțin $\hat{p}_j = \frac{n_j}{n}, j = 1, 2, \dots, s$.

Deci $\hat{p}_{ij} = \frac{n_i \cdot n_j}{n^2}, i = 1, 2, \dots, r; j = 1, \dots, s$.

Conform teoremei lui Pearson, variabila

$$H^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{(n_{ij} - n \hat{p}_{ij})^2}{n \hat{p}_{ij}}$$
 este repartizată χ^2 cu

$(rs - 1) - (r - 1) - (s - 1) = (r - 1)(s - 1)$ grade de libertate.

Avînd în vedere acest fapt, se construiește testul χ^2 pentru verificarea ipotezei $H: \{p_{ij} = q_j, j = 1, 2, \dots, s, i = 1, 2, \dots, r\}$.

Pentru un α fixat, fie $h_{(r-1)(s-1); 1-\alpha}^2$ cuantila de rang $(1 - \alpha)$ a repartiției χ^2 cu $(r - 1)(s - 1)$ grade de libertate. Regiunea critică pentru ipoteza H la pragul de semnificație α este

$$W = \{H^2 \geq h_{(r-1)(s-1); 1-\alpha}^2\}.$$

Aplicație. $n_i, i = 1, 2, 3$, și $n_j, j = 1, 2, 3, 4$ apar pe ultima linie și ultima coloană a tabelului de date.

Valorile estimate ale probabilităților p_{ij} sînt

$$\hat{p}_{11} = 0,03; \hat{p}_{12} = 0,09; \hat{p}_{13} = 0,12; \hat{p}_{14} = 0,06$$

$$\hat{p}_{21} = 0,04; \quad \hat{p}_{22} = 0,12; \quad \hat{p}_{23} = 0,16; \quad \hat{p}_{24} = 0,08$$

$$\hat{p}_{31} = -0,03; \quad \hat{p}_{32} = 0,09; \quad \hat{p}_{33} = 0,12; \quad \hat{p}_{34} = 0,06.$$

$$H^2_{calc} = \sum_{i=1}^3 \sum_{j=1}^4 \frac{(n_{ij} - \hat{np}_{ij})^2}{\hat{np}_{ij}} = 7,201$$

Numărul gradelor de libertate ale lui H^2 este egal cu $(3 - 1) \times (4 - 1) = 6$. Pentru $\alpha = 0,05$, cuantila de rang 0,95 a repartiției χ^2 cu 6 grade de libertate este $h^2_{0,95} = 12,6$.

Cum $H^2_{calc} < h^2_{0,95}$, decidem să acceptăm ipoteza de egalitate a proporțiilor pe clase a indivizilor în cele trei tratamente.

4.63. Se consideră două selecții bernoulliene de volume n_1, n_2 , independente, pentru populații ale căror caracteristici aleatoare X și Y sînt repartizate $N(m_1, \sigma_1^2)$, respectiv $N(m_2, \sigma_2^2)$, cu parametrii necunoscuți.

Să se construiască testul F (testul Snedecor) pentru verificarea ipotezei $H: \{\sigma_1^2 = \sigma_2^2\}$ cu una din alternativele $H'_1: \{\sigma_1^2 \neq \sigma_2^2\}$

$H'_2: \{\sigma_1^2 > \sigma_2^2\}$, $H'_3: \{\sigma_1^2 < \sigma_2^2\}$, la pragul de semnificație α

Aplicație. Se consideră două selecții independente de volume $n_1 = 5, n_2 = 6$ din populații normale, datele observate fiind prezentate în tabelul următor:

x_i	66	70	73	67	69	
y_i	72	70	64	68	75	78

Să se verifice ipoteza $H: \{\sigma_1^2 = \sigma_2^2\}$ (egalitatea dispersiilor celor două populații) la pragul de semnificație $\alpha = 0,05$.

Soluție. Pe baza primei selecții (X_1, \dots, X_{n_1}) construim o estimatie nedepasată a lui σ_1^2 , $S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_i - \bar{X})^2$, care are proprietatea că $\frac{(n_1 - 1)S_1^2}{\sigma_1^2}$ urmează o repartiție χ^2 cu $(n_1 - 1)$ grade de libertate.

Pe baza celei de-a doua selecții (Y_1, \dots, Y_{n_2}) construim o estimatie nedepasată a lui σ_2^2 , $S_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (Y_i - \bar{Y})^2$, care are proprie-

tatea că $\frac{(n_2 - 1)S_2^2}{\sigma_2^2}$ urmează o repartiție χ^2 cu $(n_2 - 1)$ grade de libertate.

În cazul în care $H: \{\sigma_1^2 = \sigma_2^2 = \sigma^2\}$ este adevărată, variabila $F = \frac{S_1^2}{S_2^2}$ este repartizată Snedecor cu $(n_1 - 1, n_2 - 1)$ grade de libertate. (Repartiția Snedecor are prin definiție densitatea

$$p(x) = \left(\frac{s_1}{s_2}\right)^{\frac{s_1}{2}} \frac{\Gamma\left(\frac{s_1 + s_2}{2}\right)}{\Gamma\left(\frac{s_1}{2}\right) \Gamma\left(\frac{s_2}{2}\right)} x^{\frac{s_2}{2} - 1} \left(1 + \frac{x}{s_2}\right)^{-\frac{s_1 + s_2}{2}}; \quad x \geq 0,$$

s_1, s_2 se numesc gradele de libertate ale repartiției)

Pentru un prag de semnificație α , decizia de acceptare sau de respingere a ipotezei $H: \{\sigma_1^2 = \sigma_2^2\}$ se ia conform următorului tabel:

Ipoteza H	Alternativă H'	Regiunea critică pentru H la pragul de semnificație α
$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 \neq \sigma_2^2$	$\{F_{calc} < f_1\} \cup \{F_{calc} \geq f_2\}$
$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 > \sigma_2^2$	$\{F_{calc} \geq f_3\}$
$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 < \sigma_2^2$	$\{F_{calc} < f_4\}$

Constantele de definire a testului se obțin din Anexa IV în modul următor:

- pentru $1 - P = \frac{\alpha}{2}$, f_2 se află la intersecția coloanei $\nu_1 = (n_1 - 1)$ cu linia $\nu_2 = (n_2 - 1)$
- pentru $1 - P = \frac{\alpha}{2}$, $1/f_1$ se află la intersecția coloanei $\nu_2 = (n_2 - 1)$ cu linia $\nu_1 = (n_1 - 1)$
- pentru $1 - P = \alpha$, f_3 se află la intersecția coloanei $\nu_1 = (n_1 - 1)$ cu linia $\nu_2 = (n_2 - 1)$
- pentru $1 - P = \alpha$, $1/f_4$ se află la intersecția coloanei $\nu_2 = (n_2 - 1)$ cu linia $\nu_1 = (n_1 - 1)$.

Aplicație. $\bar{x} = 69$; $\bar{y} = 71$; $S_1^2 = 7,5$; $S_2^2 = 28$

$$F_{calc} = 0,27$$

$$f_2 = 7,39$$
; $1/f_1 = 9,36$; $f_1 = 0,11$.

Cum $F_{calc} \in (f_1, f_2)$, decidem să acceptăm ipoteza de egalitate a dispersiilor.

4.64. Se consideră două selecții bernoulliene, independente, de volume n , respectiv r , efectuate asupra a două variabile aleatoare repartizate $N(m_1, \sigma^2)$, $N(m_2, \sigma^2)$. Să se construiască testul Student pentru verificarea ipotezei $H: \{m_1 = m_2\}$ cu una din alternativele $H'_1: \{m_1 \neq m_2\}$, $H'_2: \{m_1 > m_2\}$, $H'_3: \{m_1 < m_2\}$ la pragul de semnificație α .

Aplicație. Pentru observațiile din aplicația de la problema 4.63 (pentru care s-a verificat egalitatea dispersiilor) să se testeze ipoteza $H: \{m_1 = m_2\}$ la pragul de semnificație $\alpha = 0,05$.

Soluție. Mediile de selecție corespunzătoare celor două populații sînt repartizate respectiv: $\bar{X} \sim N\left(m_1, \frac{\sigma^2}{n}\right)$, $\bar{Y} \sim N\left(m_2, \frac{\sigma^2}{r}\right)$.

Diferența lor urmează tot o repartiție normală, și anume $\bar{X} - \bar{Y} \sim N\left(m_1 - m_2, \sigma^2\left(\frac{1}{n} + \frac{1}{r}\right)\right)$.

Pe de altă parte, variabila

$$\frac{1}{\sigma^2} (ns_x^2 + rs_y^2) = \frac{1}{\sigma^2} \left[\sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{i=1}^r (Y_i - \bar{Y})^2 \right]$$

este repartizată χ^2 cu $(n + r - 2)$ grade de libertate.

Rezultă că variabila

$$T = \frac{(\bar{X} - \bar{Y}) - (m_1 - m_2)}{\sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{i=1}^r (Y_i - \bar{Y})^2}{n + r - 2}} \cdot \sqrt{\frac{1}{n} + \frac{1}{r}}}$$

urmează o repartiție Student cu $(n + r - 2)$ grade de libertate.

Regiunile critice pentru ipoteza $H: \{m_1 = m_2\}$ cu alternativele considerate sînt date în tabelul următor:

Ipoteza H	Alternativa H'	Regiunea critică pentru H la pragul de semnificație α
$m_1 = m_2$	$m_1 \neq m_2$	$\left\{ T_{calc} < -t_{n+r-2; 1-\frac{\alpha}{2}} \right\} \cup$ $\cup \left\{ T_{calc} > t_{n+r-2; 1-\frac{\alpha}{2}} \right\}$
$m_1 = m_2$	$m_1 > m_2$	$\{ T_{calc} > t_{n+r-2; 1-\alpha} \}$
$m_1 = m_2$	$m_1 < m_2$	$\{ T_{calc} < -t_{n+r-2; 1-\alpha} \}$

Aplicație. $n = 5, r = 6, \bar{x} = 69, \bar{y} = 71$.

$$T_{calc} = -0,76$$

$$t_{9; 0,975} = 2,262; t_{9; 0,95} = 1,833$$

Cum $|T_{calc}| < 2,262$ și $-1,833 < T_{calc} < 1,833$, decidem să acceptăm ipoteza de egalitate a mediilor.

4.65. Se consideră trei selecții bernoulliene, independente, de volume egale cu $n = 5$, efectuate asupra a trei variabile aleatoare repartizate $N(m_i, \sigma^2)$, $i = 1, 2, 3$. Datele observate sînt prezentate în tabelul următor:

x_{1j}	64	58	52	38	48
x_{2j}	64	30	55	79	27
x_{3j}	42	59	47	98	24

Să se cerceteze egalitatea mediilor celor trei caracteristici, utilizînd tabelul de analiză dispersională.

Soluție. Adoptînd valoarea 50 cu o medie de lucru pentru simplificarea calculului, datele sînt prezentate în tabelul următor :

x_{1j}	14	8	2	-12	-2
x_{2j}	14	-20	5	29	-23
x_{3j}	-8	9	-3	48	-26

$$n_1 = n_2 = n_3 = 5, N = \sum_{i=1}^3 n_i = 15, r = 3.$$

Mediile de selecție sînt

$$\bar{x}_1 = 2, \bar{x}_2 = 1; \bar{x}_3 = 4; \bar{x} = 2,33.$$

Variabilele

$$V_A = \frac{1}{r-1} \sum_{i=1}^r n_i (\bar{X}_i - \bar{X})^2, V_R = \frac{1}{N-r} \sum_{i=1}^r \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$$

sînt două estimări nedepasate, independente pentru σ^2 .

Variabila $F = \frac{V_A}{V_R}$ urmează o repartiție Snedecor cu $(r-1, N-r)$ grade de libertate.

Tabelul de analiză dispersională este următorul :

Sumele de pătrate	Grade de libertate	Estimații pentru σ^2
$\sum_{i=1}^3 n_i (\bar{x}_i - \bar{x})^2 = 23,34$	2	$V_A = 11,67$
$\sum_{i=1}^3 \sum_{j=1}^5 (x_{ij} - \bar{x}_i)^2 = 5432$	12	$V_R = 452,7$
$\sum_{i=1}^3 \sum_{j=1}^5 (x_{ij} - \bar{x})^2 = 5455,34$	14	

$$F_{calc} = \frac{11,67}{452,7} = 0,026.$$

Pentru $\alpha = 0,05$, intervalul (f_1, f_2) se determină utilizând anexa V (vezi problema 4.63):

$$(f_1, f_2) = (0,002; 5,10).$$

Cum $F_{calc} \in (f_1, f_2)$, decidem să acceptăm ipoteza de egalitate a mediilor $H: \{m_1 = m_2 = m_3\}$.

Tabel 4.1. Datele din problema 4.63

i	x_i	y_i	z_i	x_i^2	y_i^2	z_i^2	$x_i y_i$	$x_i z_i$	$y_i z_i$
1	10,0	12,0	15,0	100,0	144,0	225,0	120,0	150,0	180,0
2	12,0	15,0	18,0	144,0	225,0	324,0	180,0	216,0	270,0
3	15,0	18,0	22,0	225,0	324,0	484,0	270,0	330,0	396,0
4	18,0	22,0	28,0	324,0	484,0	784,0	396,0	504,0	616,0
5	22,0	28,0	35,0	484,0	784,0	1225,0	504,0	770,0	882,0
6	28,0	35,0	45,0	784,0	1225,0	2025,0	770,0	1260,0	1575,0
7	35,0	45,0	58,0	1225,0	2025,0	3364,0	1260,0	2030,0	2706,0
8	45,0	58,0	72,0	2025,0	3364,0	5184,0	1836,0	3240,0	4608,0
9	58,0	72,0	90,0	3364,0	5184,0	8100,0	2700,0	5220,0	7200,0
10	72,0	90,0	110,0	5184,0	8100,0	12100,0	3600,0	7920,0	9900,0
11	90,0	110,0	140,0	8100,0	12100,0	19600,0	4620,0	12600,0	15400,0
12	110,0	140,0	180,0	12100,0	19600,0	32400,0	6160,0	19800,0	25200,0
13	140,0	180,0	230,0	19600,0	32400,0	52900,0	8120,0	32200,0	42420,0
14	180,0	230,0	300,0	32400,0	52900,0	90000,0	10800,0	54000,0	68400,0
15	230,0	300,0	400,0	52900,0	90000,0	160000,0	14600,0	92000,0	120000,0
16	300,0	400,0	550,0	90000,0	160000,0	302500,0	19800,0	165000,0	247500,0
17	400,0	550,0	750,0	160000,0	302500,0	562500,0	30000,0	300000,0	450000,0
18	550,0	750,0	1000,0	302500,0	562500,0	1000000,0	41250,0	550000,0	750000,0
19	750,0	1000,0	1350,0	562500,0	1000000,0	1822500,0	56250,0	1012500,0	1417500,0
20	1000,0	1350,0	1800,0	1000000,0	1822500,0	3240000,0	72000,0	1800000,0	2520000,0
21	1350,0	1800,0	2400,0	1822500,0	3240000,0	5760000,0	90000,0	3240000,0	4320000,0
22	1800,0	2400,0	3200,0	3240000,0	5760000,0	10240000,0	108000,0	5760000,0	7680000,0
23	2400,0	3200,0	4500,0	5760000,0	10240000,0	20250000,0	136000,0	10800000,0	14400000,0
24	3200,0	4500,0	6000,0	10240000,0	20250000,0	36000000,0	192000,0	19200000,0	25200000,0
25	4500,0	6000,0	8000,0	20250000,0	36000000,0	64000000,0	240000,0	36000000,0	48000000,0
26	6000,0	8000,0	11000,0	36000000,0	64000000,0	121000000,0	300000,0	66000000,0	88000000,0
27	8000,0	11000,0	15000,0	64000000,0	121000000,0	225000000,0	360000,0	120000000,0	165000000,0
28	11000,0	15000,0	20000,0	121000000,0	225000000,0	400000000,0	440000,0	220000000,0	300000000,0
29	15000,0	20000,0	28000,0	225000000,0	400000000,0	784000000,0	560000,0	420000000,0	560000000,0
30	20000,0	28000,0	38000,0	400000000,0	784000000,0	1444000000,0	760000,0	760000000,0	1008000000,0
31	28000,0	38000,0	50000,0	784000000,0	1444000000,0	2500000000,0	1000000,0	1400000000,0	1900000000,0
32	38000,0	50000,0	68000,0	1444000000,0	2500000000,0	4624000000,0	1300000,0	2600000000,0	3496000000,0
33	50000,0	68000,0	90000,0	2500000000,0	4624000000,0	8100000000,0	1700000,0	4500000000,0	6120000000,0
34	68000,0	90000,0	120000,0	4624000000,0	8100000000,0	14400000000,0	2200000,0	8160000000,0	10800000000,0
35	90000,0	120000,0	160000,0	8100000000,0	14400000000,0	25600000000,0	2800000,0	14400000000,0	19200000000,0
36	120000,0	160000,0	220000,0	14400000000,0	25600000000,0	48400000000,0	3600000,0	26400000000,0	35200000000,0
37	160000,0	220000,0	300000,0	25600000000,0	48400000000,0	90000000000,0	4800000,0	48000000000,0	63600000000,0
38	220000,0	300000,0	400000,0	48400000000,0	90000000000,0	160000000000,0	6400000,0	88000000000,0	116800000000,0
39	300000,0	400000,0	550000,0	90000000000,0	160000000000,0	302500000000,0	8400000,0	165000000000,0	217000000000,0
40	400000,0	550000,0	750000,0	160000000000,0	302500000000,0	562500000000,0	11000000,0	300000000000,0	405000000000,0
41	550000,0	750000,0	1000000,0	302500000000,0	562500000000,0	1000000000000,0	14000000,0	550000000000,0	725000000000,0
42	750000,0	1000000,0	1350000,0	562500000000,0	1000000000000,0	1822500000000,0	18000000,0	1012500000000,0	1350000000000,0
43	1000000,0	1350000,0	1800000,0	1000000000000,0	1822500000000,0	3240000000000,0	24000000,0	1800000000000,0	2400000000000,0
44	1350000,0	1800000,0	2400000,0	1822500000000,0	3240000000000,0	5760000000000,0	32000000,0	3240000000000,0	4320000000000,0
45	1800000,0	2400000,0	3200000,0	3240000000000,0	5760000000000,0	10240000000000,0	42000000,0	5760000000000,0	7680000000000,0
46	2400000,0	3200000,0	4500000,0	5760000000000,0	10240000000000,0	20250000000000,0	56000000,0	10800000000000,0	14400000000000,0
47	3200000,0	4500000,0	6000000,0	10240000000000,0	20250000000000,0	36000000000000,0	76000000,0	19200000000000,0	25200000000000,0
48	4500000,0	6000000,0	8000000,0	20250000000000,0	36000000000000,0	64000000000000,0	100000000,0	36000000000000,0	48000000000000,0
49	6000000,0	8000000,0	11000000,0	36000000000000,0	64000000000000,0	121000000000000,0	130000000,0	66000000000000,0	88000000000000,0
50	8000000,0	11000000,0	15000000,0	64000000000000,0	121000000000000,0	225000000000000,0	170000000,0	120000000000000,0	165000000000000,0
51	11000000,0	15000000,0	20000000,0	121000000000000,0	225000000000000,0	400000000000000,0	220000000,0	220000000000000,0	300000000000000,0
52	15000000,0	20000000,0	28000000,0	225000000000000,0	400000000000000,0	784000000000000,0	280000000,0	420000000000000,0	560000000000000,0
53	20000000,0	28000000,0	38000000,0	400000000000000,0	784000000000000,0	1444000000000000,0	360000000,0	760000000000000,0	1008000000000000,0
54	28000000,0	38000000,0	50000000,0	784000000000000,0	1444000000000000,0	2500000000000000,0	480000000,0	1400000000000000,0	1900000000000000,0
55	38000000,0	50000000,0	68000000,0	1444000000000000,0	2500000000000000,0	4624000000000000,0	640000000,0	2600000000000000,0	3496000000000000,0
56	50000000,0	68000000,0	90000000,0	2500000000000000,0	4624000000000000,0	8100000000000000,0	840000000,0	4500000000000000,0	6120000000000000,0
57	68000000,0	90000000,0	120000000,0	4624000000000000,0	8100000000000000,0	14400000000000000,0	1100000000,0	8160000000000000,0	10800000000000000,0
58	90000000,0	120000000,0	160000000,0	8100000000000000,0	14400000000000000,0	25600000000000000,0	1400000000,0	14400000000000000,0	19200000000000000,0
59	120000000,0	160000000,0	220000000,0	14400000000000000,0	25600000000000000,0	48400000000000000,0	1800000000,0	32400000000000000,0	43200000000000000,0
60	160000000,0	220000000,0	300000000,0	25600000000000000,0	48400000000000000,0	90000000000000000,0	2400000000,0	57600000000000000,0	76800000000000000,0
61	220000000,0	300000000,0	400000000,0	48400000000000000,0	90000000000000000,0	160000000000000000,0	3200000000,0	108000000000000000,0	144000000000000000,0
62	300000000,0	400000000,0	550000000,0	90000000000000000,0	160000000000000000,0	302500000000000000,0	4200000000,0	192000000000000000,0	252000000000000000,0
63	400000000,0	550000000,0	750000000,0	160000000000000000,0	302500000000000000,0	562500000000000000,0	5600000000,0	360000000000000000,0	480000000000000000,0
64	550000000,0	750000000,0	1000000000,0	302500000000000000,0	562500000000000000,0	1000000000000000000,0	7600000000,0	660000000000000000,0	880000000000000000,0
65	750000000,0	1000000000,0	1350000000,0	562500000000000000,0	1000000000000000000,0	1822500000000000000,0	10000000000,0	1012500000000000000,0	1350000000000000000,0
66	1000000000,0	1350000000,0	1800000000,0	1000000000000000000,0	1822500000000000000,0	3240000000000000000,0	13000000000,0	1800000000000000000,0	2400000000000000000,0
67	1350000000,0	1800000000,0	2400000000,0	1822500000000000000,0	3240000000000000000,0	5760000000000000000,0	17000000000,0	3240000000000000000,0	4320000000000000000,0
68	1800000000,0	2400000000,0	3200000000,0	3240000000000000000,0	5760000000000000000,0	10240000000000000000,0	22000000000,0	5760000000000000000,0	7680000000000000000,0
69	2400000000,0	3200000000,0	4500000000,0	5760000000000000000,0	10240000000000000000,0	20250000000000000000,0	28000000000,0	10800000000000000000,0	14400000000000000000,0
70	3200000000,0	4500000000,0	6000000000,0	10240000000000000000,0	20250000000000000000,0	36000000000000000000,0	36000000000,0	19200000000000000000,0	25200000000000000000,0
71	4500000000,0	6000000000,0	8000000000,0	20250000000000000000,0	36000000000000000000,0	64000000000000000000,0	48000000000,0	36000000000000000000,0	48000000000000000000,0
72	6000000000,0	8000000000,0	11000000000,0	36000000000000000000,0	64000000000000000000,0	121000000000000000000,0	64000000000,0	66000000000000000000,0	88000000000000000000,0
73	8000000000,0	11000000000,0	15000000000,0	64000000000000000000,0	121000000000000000000,0	225000000000000000000,0	84000000000,0	120000000000000000000,0	165000000000000000000,0
74	11000000000,0	15000000000,0	20000000000,0	121000000000000000000,0	225000000000000000000,0	400000000000000000000,0	110000000000,0	220000000000000000000,0	300000000000000000000,0
75	15000000000,0	20000000000,0	28000000000,0	225000000000000000000,0	400000000000000000000,0	784000000000000000000,0	140000000000,0	420000000000000000000,0	56000

ANEXE

Anexa I

Repartiția normală standard $N(0, 1)$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$

x	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
-0,0	0,5000	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641
-0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
-0,2	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
-0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483
-0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
-0,5	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2843	0,2810	0,2777
-0,6	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2452
-0,7	0,2420	0,2389	0,2358	0,2327	0,2297	0,2266	0,2236	0,2206	0,2177	0,2148
-0,8	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867
-0,9	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
-1,0	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379
-1,1	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1171
-1,2	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0986
-1,3	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823
-1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0721	0,0708	0,0694	0,0681
-1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0570	0,0559
-1,6	0,0548	0,0537	0,0526	0,0515	0,0505	0,0495	0,0485	0,0475	0,0465	0,0456
-1,7	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367
-1,8	0,0359	0,0351	0,0344	0,0336	0,0332	0,0329	0,0314	0,0307	0,0300	0,0293
-1,9	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0238	0,0233
-2,0	0,0227	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0184
-2,1	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143
-2,2	0,0139	0,0135	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0111
-2,3	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0086
-2,4	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0065
-2,5	0,0062	0,0060	0,0059	0,0057	0,0055	0,0054	0,0052	0,0051	0,0049	0,0048

Repartiția normală standard $N(0, 1)$

x	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
-2,6	0,0047	0,0045	0,0044	0,0043	0,0041	0,0040	0,0039	0,0038	0,0037	0,0036
-2,7	0,0035	0,0034	0,0033	0,0032	0,0030	0,0030	0,0029	0,0028	0,0027	0,0026
-2,8	0,0026	0,0025	0,0024	0,0023	0,0023	0,0022	0,0021	0,0020	0,0020	0,0019
-2,9	0,0019	0,0018	0,0017	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014
-3,0	0,0013	0,0013	0,0012	0,0012	0,0012	0,0011	0,0011	0,0011	0,0010	0,0010
0,0	0,5000	0,5040	0,5080	0,5120	0,5136	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7703	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9122	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9463	0,9474	0,9484	0,9595	0,9505	0,9515	0,9525	0,9535
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9919	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990

Repartiția χ^2 . Cuantilele h_n^{α} , $1-\alpha$

$1-\alpha$	0,995	0,990	0,975	0,950	0,900	0,800	0,700	0,500
1	—	0,0002	0,0010	0,0039	0,0158	0,0642	0,148	0,455
2	0,0100	0,0201	0,0506	0,103	0,211	0,466	0,713	1,39
3	0,0717	0,115	0,216	0,352	0,584	1,01	1,42	2,37
4	0,207	0,297	0,484	0,711	1,06	1,65	2,20	3,36
5	0,412	0,554	0,831	1,15	1,61	2,34	3,00	4,35
6	0,676	0,872	1,24	1,64	2,20	3,07	3,83	5,35
7	0,989	1,24	1,69	2,17	2,83	3,82	4,67	6,35
8	1,34	1,65	2,18	2,73	3,49	4,59	5,53	7,34
9	1,73	2,09	2,70	3,33	4,17	5,38	6,39	8,34
10	2,16	2,56	3,25	3,94	4,87	6,18	7,27	9,34
11	2,60	3,05	3,82	4,57	5,58	6,99	8,15	10,3
12	3,07	3,57	4,40	5,23	6,30	7,81	9,03	11,3
13	3,57	4,11	5,01	5,89	7,04	8,63	9,93	12,3
14	4,07	4,66	5,63	6,57	7,79	9,47	10,8	13,3
15	4,60	5,23	6,26	7,26	8,55	10,3	11,7	14,3
16	5,14	5,81	6,91	7,96	9,31	11,2	12,6	15,3
17	5,70	6,41	7,56	8,67	10,1	12,0	13,5	16,3
18	6,26	7,01	8,23	9,39	10,9	12,9	14,4	17,3
19	6,84	7,63	8,91	10,1	11,7	13,7	15,4	18,3
20	7,43	8,26	9,59	10,9	12,4	14,6	16,3	19,3
21	8,03	8,90	10,3	11,6	13,2	15,4	17,2	20,3
22	8,64	9,54	11,0	12,3	14,0	16,3	18,1	21,3
23	9,26	10,2	11,7	13,1	14,8	17,2	19,0	22,3
24	9,89	10,9	12,4	13,8	15,7	18,1	19,9	23,3
25	10,5	11,5	13,1	14,6	16,5	18,9	20,9	24,3
26	11,2	12,2	13,8	15,4	17,3	19,8	21,8	25,3
27	11,8	12,9	14,6	16,2	18,1	20,7	22,7	26,3
28	12,5	13,6	15,3	16,9	18,9	21,6	23,6	27,3
29	13,1	14,3	16,0	17,7	19,8	22,5	24,6	28,3
30	13,8	15,0	16,8	18,5	20,6	23,4	25,5	29,3

Observație:

Pentru $n > 30$ se poate admite că $\sqrt{2\chi^2} - \sqrt{2n-1}$ urmează o repartiție $N(0, 1)$.

Repartiția Student. Cuantilele $t_{1-\alpha, n}$ pentru care

$$P(t < t_{1-\alpha, n}) = 1 - \alpha$$

$1 - \alpha$	0,75	0,90	0,95	0,975	0,99	0,995
n						
1	1,000	3,078	6,314	12,706	31,821	63,657
2	0,816	1,886	2,920	4,303	6,965	9,925
3	0,765	1,638	2,353	3,182	4,541	5,841
4	0,741	1,533	2,132	2,776	3,747	4,604
5	0,727	1,476	2,015	2,571	3,365	4,032
6	0,718	1,440	1,943	2,447	3,143	3,707
7	0,711	1,415	1,895	2,365	2,998	3,499
8	0,706	1,397	1,860	2,306	2,896	3,355
9	0,703	1,383	1,833	2,262	2,821	3,250
10	0,700	1,372	1,812	2,228	2,764	3,169
11	0,697	1,363	1,796	2,201	2,718	3,106
12	0,695	1,356	1,782	2,179	2,681	3,055
13	0,694	1,350	1,771	2,160	2,650	3,012
14	0,692	1,345	1,761	2,145	2,624	2,977
15	0,691	1,341	1,753	2,131	2,602	2,947
16	0,690	1,337	1,746	2,120	2,583	2,921
17	0,689	1,333	1,740	2,110	2,567	2,898
18	0,688	1,330	1,734	2,101	2,552	2,878
19	0,688	1,328	1,729	2,093	2,539	2,861
20	0,687	1,325	1,725	2,086	2,528	2,845
21	0,686	1,323	1,721	2,080	2,518	2,831
22	0,686	1,321	1,717	2,074	2,508	2,819
23	0,685	1,319	1,714	2,069	2,500	2,807
24	0,685	1,318	1,711	2,064	2,492	2,797
25	0,684	1,316	1,708	2,060	2,485	2,787
26	0,684	1,315	1,706	2,056	2,479	2,779
27	0,684	1,314	1,703	2,052	2,473	2,771
28	0,683	1,313	1,701	2,048	2,467	2,763
29	0,683	1,311	1,699	2,045	2,462	2,756
30	0,683	1,310	1,697	2,042	2,457	2,750
35	0,681	1,306	1,690	2,030	2,438	2,724
40	0,681	1,303	1,684	2,021	2,423	2,704
80	0,679	1,291	1,671	2,000	2,390	2,660
120	0,677	1,289	1,658	1,980	2,358	2,617
n > 120	0,674	1,282	1,645	1,960	2,326	2,576

Observație. $t_{\alpha, n} = -t_{1-\alpha, n}$

Valorile $D_n; 1-\alpha$ ale testului Lilliefors

$n \backslash 1-\alpha$	0,90	0,95	0,99
4	0,352	0,381	0,417
5	0,315	0,337	0,405
6	0,294	0,319	0,364
7	0,276	0,300	0,384
8	0,261	0,285	0,331
9	0,249	0,271	0,311
10	0,239	0,258	0,294
11	0,230	0,249	0,284
12	0,223	0,242	0,275
13	0,214	0,234	0,268
14	0,207	0,227	0,261
15	0,201	0,220	0,257
16	0,195	0,213	0,250
17	0,189	0,206	0,245
18	0,184	0,200	0,239
19	0,179	0,195	0,235
20	0,174	0,190	0,231
25	0,158	0,173	0,200
30	0,144	0,161	0,187
$n > 30$	$\frac{0,805}{\sqrt{n}}$	$\frac{0,886}{\sqrt{n}}$	$\frac{1,031}{\sqrt{n}}$

Repartiția Snedecor

Calculul valorilor (f_1, f_2) pentru (v_1, v_2) grade de libertate și probabilitate ($1 - p$)

$v_1 \backslash v_2$	$1-p$	1	2	3	4	5	6	7	8	9	10	12	15	20
1	0,05	161	200	216	225	230	234	237	239	241	242	244	246	248
	0,025	648	800	864	900	922	937	948	957	963	969	977	985	993
2	0,05	18,5	19,0	19,2	19,2	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,4
	0,025	38,5	39,0	39,2	39,2	39,9	39,3	39,4	39,4	39,4	39,4	39,4	39,4	39,4
3	0,05	10,1	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,66
	0,025	17,4	16,0	15,4	15,1	14,9	14,7	14,6	14,5	14,5	14,4	14,3	14,3	14,2
4	0,05	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,80
	0,025	12,2	10,6	9,98	9,60	9,36	9,20	9,07	8,98	8,90	8,84	8,75	8,66	8,56
5	0,05	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56
	0,025	10,0	8,43	7,76	7,39	7,15	6,98	6,85	6,76	6,68	6,62	6,52	6,43	6,33
6	0,05	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,87
	0,025	8,81	7,26	6,60	6,23	5,99	5,82	5,70	5,60	5,52	5,46	5,37	5,27	5,17
7	0,05	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,44
	0,025	8,07	6,54	5,89	5,52	5,29	5,12	4,99	4,90	4,82	4,76	4,67	4,57	4,47
8	0,05	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15
	0,025	7,57	6,06	5,42	5,05	4,82	4,65	4,53	4,43	4,36	4,30	4,20	4,10	4,00
9	0,05	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94
	0,025	7,21	5,71	5,08	4,72	4,48	4,32	4,20	4,10	4,03	3,96	3,87	3,77	3,67
10	0,05	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,84	2,77
	0,025	6,94	5,46	4,83	4,47	4,24	4,07	3,95	3,85	3,78	3,72	3,62	3,52	3,42
12	0,05	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,54
	0,025	6,55	5,10	4,47	4,12	3,89	3,73	3,61	3,51	3,44	3,37	3,28	3,18	3,07

BIBLIOGRAFIE

1. Anderson, T. W. *An Introduction to Multivariate Statistical Analysis*. John Wiley, New York, 1958.
 2. Breiman, L. *Probability*. Addison—Wesley, 1968.
 3. Ciucu, G., Craiu, V., Săcuiu, I. *Probleme de teoria probabilităților*. Editura Tehnică, 1974.
 4. Ciucu, G., Tudor, C. *Probabilități și procese stocastice*. Ed. Academiei, vol. I—1978, vol II — 1979.
 5. Ciucu, G., Tudor, C. *Teoria probabilităților și aplicații*. Ed. științifică și enciclopedică, 1983.
 6. Ciucu, G., Sâmbuan, G. *Teoria probabilităților și statistică matematică*. Culegere de probleme. Ed. Tehnică, 1962.
 7. Cox, D. R., Hinkley, D. V. *Theoretical Statistics*. Chapman and Hall, London, 1974.
 8. Cox, D. R., Hinkley, D. V. *Problems and Solutions in Theoretical Statistics*. John Wiley, New York, 1978.
 9. Cuculescu, I. *Curs de teoria probabilităților*. Tipografia Universității București, 1976.
 10. Dumitrescu, M., Florea, D., Tudor, C. *Elemente de teoria probabilităților și statistică matematică. Probleme și soluții*. Tipografia Universității București, 1983.
 11. Emelianov, G. V., Schitovici, V. P., *Exerciții de teoria probabilităților și statistică matematică* (în limba rusă). Leningrad, 1967.
 12. Feller, W. *An introduction to probability Theory and its applications*. John Wiley, vol. I — 1957, vol. II — 1966.
 13. Gnedenko, B. *The theory of probability*. Mir Publishers, 1969.
 14. Vodă, ș.a. V. *Statistică și toleranțe*. Editura Tehnică, 1977.
 15. Iosifescu, M., Mihoc, Gh. ș.a. *Teoria probabilităților și statistică matematică*. Ed. didactică și pedagogică, 1970.
 16. Lamperti, J. *Probability*. Darmouth College, 1966.
 17. Lehman, E. L. *Testing Statistical Hypotheses*. John Wiley, New York.
 18. Letac, G. *Problèmes de probabilités*. Press Univ. de France, 1970.
 19. Loeve, M. *Probability theory*. Van Nostrand, 1963.
 20. Mihoc, G., Ciucu, G., Craiu, V. *Teoria probabilităților și statistică matematică*. Ed. didactică și pedagogică, 1970.
 21. Neveu, J. *Bases mathématiques du Calcul des Probabilités*. Masson, 1970.
 22. Onicescu, O., Mihoc, Gh. ș.a. *Calculul probabilităților și aplicații*. Ed. Academiei, 1956.
 23. Sâmbuan, G. ș.a. *Teoria probabilităților*. Ed. didactică și pedagogică, 1967.
 24. Silvey, S. D., *Statistical Inference*. Penguin Books, 1970.
 25. Stoyanov, J. ș.a. *Probleme de teoria probabilităților* (în limba bulgară). Sofia, 1976.
 26. Reyni, A. *Calcul des probabilités*. Dunod, 1966.
-