



Alexei Leahu, Ion Pârțachi

PROBABILITĂȚI ȘI STATISTICĂ
(Prin exemple și probleme propuse)
Partea I. PROBABILITĂȚI



Chișinău
© 2021

CUPRINS

Prefață

1. Calculul Probabilităților

1.1. Obiectul de studiu al Teoriei Probabilităților și locul ei în Statistica Matematică, probabilitate frecvențială, probabilitate subiectivă

1.2. Noțiuni și rezultate auxiliare din Combinatorică

1.3. Spații de evenimente elementare, evenimente aleatoare și operații asupra lor, definiția axiomatică a probabilității

1.4. Proprietățile probabilității drept consecință din definiția axiomatică a probabilității

1.5. Probabilități clasice, discrete și geometrice drept cazuri particulare ale definiției axiomatică a probabilității

1.6. Probabilitate condiționată. Formula înmulțirii probabilităților

1.7. Independența evenimentelor aleatoare, formula lui Poisson

1.8. Formulele probabilității totale și a lui Bayes

2. Variabile aleatoare

2.1. Introducere

2.2. Variabilă aleatoare (unidimensională), funcția ei de distribuție (repartiție)

2.3. Variabile aleatoare de tip discret, distribuții (repartiții)

2.4. Variabile aleatoare de tip (absolut) continue, densități de distribuție (repartiție)

2.5. Variabile aleatoare mixate discrete-continue

2.6. Variabilă aleatoare multidimensională (vectorială), funcția ei de distribuție, funcții de distribuție marginale

2.7. Tipurile de variabile aleatoare multidimensionale (bidimensionale), distribuții, densități de distribuție, independența v.a.

3. Caracteristici numerice ale variabilelor aleatoare

3.1. Parametri de poziție: valoarea medie, moda, mediana, cuantile

3.2. Disperia (varianța), abaterea standard, covarianța, coeficientul de corelație, regresie liniară

3.3. Momente ale variabilei aleatoare (inițiale, centrale), asimetria, aplatizarea

4. Modele (distribuții, densități de distribuție) probabiliste uzuale, inegalități, Legea Numerelor Mari, Teorema Limită Centrală pentru variabile aleatoare independente identic distribuite

4.1. Distribuții probabiliste uzuale in caz discret (Uniformă, Bernoulli, Binomială, Geometrică, Poisson, Multinomială, Hypergeometrică, Hypergeometrică Multivariată)

4.2. Distribuții probabiliste uzuale in caz (absolut) continuu (Uniformă, Exponențială, Normală, Hi-pătrat (χ^2), T -Student)

4.3. Inegalitatea Chebyshev, Legea Numerelor Mari (în formele Chebyshev, Bernoulli, Hincin), Teorema Limită Centrală

5. Elemente de Teoria Informației

5.1. Obiectul de studiu al Teoriei Informației

5.2. Entropia ca măsură a nedeterminării/cantității de informație

5.3. Proprietățile entropiei

5.4. Transmiterea informației. Codificarea. Teoreme de codificare

BIBLIOGRAFIE

Prefață

Utilizarea modelelor și metodelor statistico-probabiliste la analiza și interpretarea datelor a devenit o practică uzuală în toate domeniile științifice, fie că este vorba de prelucrarea datelor cu caracter economic sau datelor experimentale în tehnică sau de analiză și interpretarea datelor legate de elaborarea unor noi tehnologii informaționale, etc.

Disciplina matematică **Probabilități și Statistică** cuprinde elementele de bază ale **Teoriei Probabilităților, Statisticii Descriptive și Matematice**, dar și exemple de **Aplicații** ale lor. **Teoria Probabilităților** este cea care:

1) Oferă noțiunile care stau la baza modelelor ce descriu comportamentul probabilist al fenomenelor aleatoare (eveniment aleator, probabilitate, variabilă aleatoare (v.a.) și funcția ei de distribuție (f.d.), valoare medie, dispersie și alte caracteristici numerice ale v.a., etc.);

2) Pune la dispoziție formule și rezultate teoretice fructificate în metodele de calcul ale probabilităților (formulele adunării, înmulțirii, probabilității totale, lui Bayes, dar și formulele de calcul bazate pe funcția de distribuție, distribuția sau densitatea de distribuție a v.a.);

3) Servește, grație proprietăților caracteristice f.d., drept sursă inepuizabilă de modele probabiliste, scoțând în evidență modelele probabiliste uzuale sau clasice (distribuțiile Uniformă, Bernoulli, Binomială, Geometrică, Hypergeometrică, Exponențială, Normală, Hi-patrat, Student, etc.);

4) Aflându-se la baza Teoriei Informației, permite aplicarea metodelor cantitative asupra conceptelor legate de informație, inclusiv în contextul cercetării sistemelor de prelucrare și transmitere a informației.

Statistica Matematică este cea care, prin prisma Teoriei Probabilităților, oferă un cadru riguros pentru analiza și interpretarea datelor statistice. Aceasta permite ca rezultatele prelucrării primare a datelor realizate cu ajutorul Statisticii Descriptive să fie aprofundate prin a cerceta calitatea, exactitatea estimatorilor utilizați, dar și prin lansarea, verificarea și validarea ipotezelor statistice privind legitățile probabiliste care guvernează sau generează datele statistice colectate. Cu alte cuvinte, în baza unor observații parțiale ale unui fenomen aleator putem, fie și cu o anumită aproximare și grad de încredere, anticipa comportamentul probabilist al acestuia.

1. Calculul Probabilităților

1.1. Obiectul de studiu al Teoriei Probabilităților și locul ei în Statistica Matematică, probabilitate frecvențială, probabilitate subiectivă

Noțiunea de *statistică* se naște odată cu apariția și dezvoltarea relațiilor economice, până în secolul XIX aceasta fiind tratată ca știință politică. Cuvântul în cauză provine de la latinescul *status*, care înseamnă stare.

Începând cu secolul XIX, *statistica* prinde conturul științei care, actualmente, are drept obiect de studiu metodele, procedeele de colectare, organizare, prelucrare, analiză și interpretare a datelor ce vizează rezultatele observărilor făcute asupra fenomenelor sau experimentelor aleatoare. Statistica modernă, mai ales acea parte a ei care se numește *Statistica matematică*, bazându-se esențial pe realizările științelor matematice, folosește din plin *Teoria probabilităților*.

Apariția Teoriei probabilităților ca ramură a matematicii ce studiază modele matematice ale fenomenelor aleatoare datează din secolul XVII și este legată de numele marilor matematicieni Blaise Pascal (1623-1662), Pierre Fermat (1601-1665), Christian Huygens (1629-1695) și Jacob Bernoulli (1654-1705).

Pentru a înțelege mai bine ce studiază Teoria probabilităților ca parte a Statisticii matematice, avem nevoie de câteva explicații suplimentare, ținând cont că mulțitudinea de fenomene care se întâlnesc în lumea înconjurătoare se împarte în două clase: *fenomene deterministe* și *fenomene indeterminate* sau *aleatoare*.

Astfel, spunem că fenomenul este *determinist* dacă observatorul poate anticipa cu certitudine evoluția sau derularea acestuia. În calitate de exemplu putem lua fenomenul atracției universale. Observațiile făcute asupra acestui fenomen i-au permis marelui matematician și fizician englez Isaac Newton (1642-1727) să formuleze legea atracției universale:

$$F = k \frac{m_1 m_2}{r^2} .$$

Acesta este un exemplu tipic de model matematic al unui fenomen (în cazul dat) determinist. Dealtfel, *a modela matematic (spre a fi cercetat) un fenomen, proces, experiment, eveniment sau obiect oarecare înseamnă a-l descrie cu ajutorul noțiunilor și formulelor matematice, adică a-l descrie în limbajul*

matematic. Unul și același model matematic poate descrie două fenomene diferite în esență. De exemplu, formula de mai sus poate servi în calitate de model matematic și pentru fenomenul atracției a două particule elementare (legea lui Coulomb).

Spunem despre un fenomen că este *indeterminist (aleator)*, dacă observatorul fenomenului *nu* poate anticipa cu certitudine evoluția lui. Din punct de vedere al observatorului, observațiile făcute asupra unui fenomen sau măsurătorile corespunzătoare echivalează cu o experimentare legată de fenomenul dat. Or, prin *experiment* vom înțelege observarea unui fenomen dat.

Acum putem spune că *Teoria probabilităților* și *Statistica matematică* studiază modele matematice ale experimentelor (fenomenelor) aleatoare.

Aici se impun câteva precizări. Experimentele aleatoare (indeterministe) se împart, la rândul lor, în două subclase: **(a)** *experimente aleatoare care posedă proprietatea regularității (stabilității) statistice* și **(b)** *experimente aleatoare care nu posedă proprietatea regularității statistice*.

Definiția 1. Vom spune că un experiment aleator \mathcal{E} posedă *proprietatea regularității (stabilității) statistice* dacă acesta verifică următoarele proprietăți:

- 1) *poate fi reprodus ori de câte ori dorim practic în aceleași condiții;*
- 2) *pentru orice eveniment A asociat lui \mathcal{E} frecvența lui relativă în n probe*

$$f_n(A) = \frac{\text{numărul de probe în care s-a produs } A}{\text{numărul total de probe}} = \frac{n(A)}{n}$$

oscilează în jurul unui număr notat cu $P(A)$, $P(A) \in [0, 1]$, $f_n(A)$ devenind, odată cu creșterea lui n , "tot mai aproape și mai aproape de $P(A)$ ";

- 3) *pentru două serii diferite, respectiv de n și m probe, atunci când n și m sunt foarte mari, avem că $f_n(A) \approx f_m(A)$.*

În concluzie, stabilitatea statistică a frecvențelor relative conferă verosimilitate ipotezei, conform căreia pentru orice eveniment A , posibil ca rezultat observabil al unui experiment aleator \mathcal{E} , putem defini numărul $P(A)$ cu ajutorul căruia măsurăm gradul (șansele) de realizare a lui A într-un număr foarte mare de probe. Astfel, în Teoria probabilităților devine postulat afirmația, conform căreia pentru orice eveniment A asociat unui experiment aleator \mathcal{E} există (obiectiv) un număr $P(A)$ numit probabilitate a lui A . Proprietatea

firească a acestui număr rezidă în faptul că odată cu creșterea numărului n de probe (experimente) "independente" frecvența relativă $f_n(A)$ se apropie tot mai mult de $P(A)$. Numărul $P(A)$ se numește *probabilitate statistică (sau frecvențială) a evenimentului A*. Aceasta din urmă afirmație/proprietate este conformă cu Legea Numerelor Mari, care în Teoria probabilităților reprezintă suportul matematic al proprietății regularității/stabilității statistice, ceea ce înseamnă că aceasta teorie se pliază perfect pe aplicațiile ei la rezolvarea problemelor ce apar în cadru studierii modelelor matematice ale experimentelor aleatoare ce fac parte din subclasa (a).

Exemplul 1. Considerăm în calitate de experiment aleator \mathcal{E} aruncarea unei monede simetrice (cu centrul de greutate nedeplasat) sau, cum se mai spune, "perfecte". Fie A evenimentul ce constă în apariția stemei. Observăm, astfel, că.

$$f_{1000}(A) \approx \frac{1}{2} = P(A), \quad f_{2000}(A) \approx \frac{1}{2} = P(A).$$

Prin urmare, putem afirma că probabilitatea (statistică) a apariției stemei la aruncarea monedei o singură dată este egală cu $1/2$, ceea ce înseamnă, că aruncând moneda de un număr suficient de mare de ori, stema va apărea în aproximativ 50% de cazuri .

Putem aduce și alte exemple de fenomene aleatoare: rezultatele aruncării unui zar, greutatea unui bob de grâu ales la întâmplare, numărul de bacterii într-o picătură de apă, durata vieții unui calculator produs de întreprinderea dată, numărul de apeluri telefonice înregistrate la o stație telefonică pe durata unei zile, etc., etc. Enumerarea lor poate continua la nesfârșit, însă ele toate vor avea același caracter, fiind însoțite de astfel de noțiuni imprecise (deocamdată) ca aruncare "onestă", moneda "perfectă", "probe independente", etc.

Interesant este faptul că frecvența relativă, prin urmare și probabilitatea statistică (frecvențială), posedă proprietățile puse în anul 1934 la baza Teoriei axiomatice a probabilităților, teorie lansată de către ilustrul matematician rus A. N. Kolmogorov (1903-1987). Iată aceste proprietăți care pot fi verificate cu ușurință pe cale experimentală în baza următorului exemplu.

Exemplul 2. Considerăm experimentul aleator \mathcal{E} ce constă în efectuarea un sondaj printre potențialii săi clienți de către o companie producătoare de 3 tipuri de băuturi răcoritoare A, B și C , pentru ca în funcție

de rezultatele sondajului sa se facă ajustări, privind cotele de producție a fiecărui tip de bautura. Exprimarea preferinței presupune alegerea unui singur raspuns: A , B sau C . Considerând că în eșantion au fost incluși un număr n , suficient de mare. de respondenți, atunci compania va putea identifica cotele de producție, acestea fiind direct proporționale cu frecvențele relative $f_n(A)$, $f_n(B)$ și $f_n(C)$. Se poate observa cu ușurință că:

1. $0 \leq f_n(A) \leq 1$ (rezultă din definiția frecvenței relative);
2. $f_n(A \text{ sau } B \text{ sau } C) = 1$ (deoarece raspunsurile A sau B sau C se exclud mutual, doua cate doua, iar alte raspunsuri sunt excluse, cu alte cuvinte înregistrarea ca raspuns A sau B sau C , privit ca eveniment aleator, este eveniment sigur);
3. $f_n(A \text{ sau } B) = f_n(A) + f_n(B)$, $f_n(A \text{ sau } C) = f_n(A) + f_n(C)$, $f_n(B \text{ sau } C) = f_n(B) + f_n(C)$.

Remarca1. Probabilitatea statistică nu poate fi aplicată întotdeauna, deoarece nu orice experiment poate fi repetat în condiții identice ori de câte ori dorim.

Drept exemplu putem lua rezultatul pe care-l va obține un student *concret* la examenul de Probabilități și Statistică matematică. Într-adevăr, observăm ca acest examen (experiment aleator), raportat la un student concret, nu poate fi *poate fi reprodus ori de câte ori dorim practic în aceleași condiții*. Dacă, însă, vom modifica experimentul prin a ne referi în fiecare proba (repetare de examen) la un student "ales la întâmplare", atunci acest experiment va satisface rigorile 1-3 ale regularității statistice, apelând, de exemplu, la istoricul rezultatelor de la acest examen din anii anteriori. Or, experimentele aleatoare care posedă proprietatea regularității statistice țin de fenomenele de masă. Pentru studiul experimentelor care nu posedă această proprietate, putem folosi noțiunea de probabilitate subiectivă.

Definiția 2. Prin *probabilitate subiectivă* vom înțelege acea regulă P conform căreia o persoană dată îi asociază fiecărui eveniment A un număr $P(A) \in [0, 1]$, numit *probabilitatea* evenimentului A .

Astfel, orice student concret își poate, bineînțeles, evalua probabilitatea subiectivă că va promova din prima încercare examenul de Probabilități și Statistică Matematică, această probabilitate variind în funcție de student.

La fel, putem vorbi despre probabilitatea subiectivă, evaluată, să zicem, de un expert NASA, că până în anul 2025 va avea loc prima expediție a omului pe Marte.

Pentru studiul fenomenelor aleatoare indeterminate, în afară de probabilitate *subiectivă* și probabilitate *frecvențială*, există și noțiunile de probabili-

tate *clasică*, probabilitate *geometrică*, probabilitate *discretă* și probabilitate definită în sens *axiomatic*. Toate aceste noțiuni au ca scop definirea unei modalități de măsurare a șanselor (gradelor) de realizare, gradelor de certitudine în producerea evenimentelor aleatoare date, *definiția axiomatică a probabilității*, fiind, într-un anumit sens, *acoperitoare pentru toate noțiunile de probabilitate* enumerate mai sus.

Remarca 2. În ceea ce privește fenomenele aleatoare întâlnite în domeniile științelor ingineresti acestea au, de regulă, un caracter de masă. Însă, dacă ne vom referi la domeniile economic și al afacerilor, acestea fiind vulnerabile la factorii subiectivi, trebuie să fim foarte atenți dacă fenomenele aleatoare cu caracter economic, puse de noi în discuție, satisfac rigorilor impuse asupra experimentelor aleatoare pentru ca acestea să fie tratate/clasificate ca fiind experimente ce posedă proprietatea regularității statistice sau nu.

Precizăm, ca *Teoria probabilităților și Statistica Matematică*, expusă de noi în continuare, *vizează doar modele matematice ale experimentelor aleatoare ce posedă proprietatea regularității statistice*. Cu alte cuvinte vom vorbi doar despre *probabilități obiective*.

Probleme propuse

1. La prima lecție practică, privind calculul probabilităților, studenților le-a fost propusă următoarea întrebare: "Cu ce este egală probabilitatea că, ieșind la plimbare pe bulevardul Ștefan cel Mare din Chișinău, întâmplător, vă va ieși în cale un Dinosaur?" Unul din studenți, familiarizat cu definiția clasică a probabilității încă din liceu, a răspuns că probabilitatea în cauză este egală cu $1/2$, deoarece... "este posibil unul din două cazuri cu șanse egale de realizare: ori vom întâlni un Dinosaur, ori Nu". Răspunsul acesta, fiind unul ridicol, se impune întrebarea: ce metodă de evaluare a probabilității se potrivește în acest caz? Pe cale experimentală sau subiectivă? Cu ce este egală probabilitatea respectivă? Argumentați.

2. Biblioteca Națională "Vasile Alecsandri" din Chișinău a înregistrat pe parcursul anului 2017 cititori nou veniți care, privind nivelul lor de studii, s-au repartizat conform următorului tabel:

Nivelul de studii	Frecvența
Superioare	720
Studenti/superioare incomplete	1106
Medii (liceu, școala medie)	502
Medii incomplete	52
Fără studii	0
<i>Total</i>	2380

Evaluați probabilitatea că un cititor luat la întâmplare din lista celor înregistrați la Biblioteca Națională "V. Alecsandri" va fi unul cu studii.

Evaluați probabilitățile respective pentru fiecare nivel de studii și arătați că suma lor este egală cu 1.

3. Multe instituții de învățământ preuniversitar oferă elevilor acces la internet. Astfel, conform Anuarului Statistic din SUA publicat în 1997, în anul 1996 acces la internet aveau 21 733 din 51 745 școli primare, 7286 din 14 012 școli gimnaziale și 10 682 din 17 229 licee.

Cu ce este egală probabilitatea ca:

a) Alegând la întâmplare pentru a vizita *una din școlile primare*, aceasta va fi conectată la Internet;

b) Alegând la întâmplare pentru a vizita *una din școlile gimnaziale*, aceasta va fi conectată la Internet;

c) Alegând la întâmplare pentru a vizita *unul din licee*, acesta va fi conectat la Internet;

d) Alegând la întâmplare pentru a vizita *una din instituțiile de învățământ menționate mai sus*, aceasta va fi conectată la Internet.

4. O companie producătoare de pastă de dinți a inițiat un studiu privind gradul de preferință a cumpărătorului în funcție de unul din cele 5 tipuri de ambalaje propuse de designeri. Drept rezultat, au fost înregistrate următoarele rezultate: 5 cumpărători au preferat ambalajul nr. 1, 15- ambalajul nr. 2, 30- ambalajul nr. 3, 40- ambalajul nr.4 și 10- ambalajul nr. 5.

Sunt oare aceste date suficiente pentru a depista ambalajul cu șansele cele mai mari de a fi preferat de către cumpărători? Argumentați.

5. O companie specializată în construirea blocurilor de locuit, ce constă din apartamente cu 1, 2 și 3 camere, a estimat că probabilitățile subiective, privind cererea din partea unui client, în funcție de tipul apartamentului solicitat, sunt următoarele: 0.30 pentru un apartament cu o cameră, 0.40 pentru un apartament cu 2 camere și 0.25 pentru un apartament cu 3 camere.

a) Sunt oare valide aceste estimari ale probabilităților mentionate mai sus? De ce DA sau de ce NU?

b) Ce intervenție asupra valorilor probabilităților trebuie aplicate ca acestea sa devină valide?

6. Specificați care din exemplele descrise mai jos se referă la experimentele aleatoare ce posedă proprietatea Regularității Statistice și care nu:

- a) Rezultatul unui meci concret în cadrul campionatului mondial la fotbal;
- b) Rezultatul jocului cu o singură variantă la Lotosport "6 din 49";
- c) Eșecul sau succesul unui businessman la sfârșitul primului an de activitate în cadrul unei afaceri cu caracter economic;
- d) Mărimea încălțăminte solicitate de un cumpărător ales la întâmplare dintre cei care au vizitat o secție de încălțăminte a unui Mall;
- e) Numărul de apeluri telefonice înregistrate în 24 de ore la un centru de urgență medicală.

1.2. Noțiuni și rezultate auxiliare din Combinatorică

Analiza Combinatorie sau Combinatorica este acea disciplină matematică în care sunt studiate metodele de numărare a tuturor combinațiilor ce pot fi alcătuite din elementele unei mulțimi finite. Aceasta se bazează esențial pe două principii: *Principiul adunării* și *Principiul înmulțirii* expuse mai jos.

Fie A și B sunt două mulțimi finite, atunci distingem două situații, după cum cele două mulțimi pot fi disjuncte sau nu. Se verifică cu ușurință că are loc

Principiul adunării (caz disjunct) Dacă A și B sunt mulțimi finite și disjuncte, adică $A \cap B = \emptyset$, $\text{card}(A) = n$ și $\text{card}(B) = m$, atunci

$$\text{card}(A \cup B) = n + m.$$

Corolar. Dacă A_1, A_2, \dots, A_k sunt mulțimi finite disjuncte două câte două, atunci

$$\text{card}\left(\bigcup_{i=1}^k A_i\right) = \sum_{i=1}^k \text{card } A_i$$

Principiul adunării (caz general). Dacă A și B sunt mulțimi finite, $A, B \subseteq \Omega$, $\text{card}(A) = n$, $\text{card}(B) = m$ și $\text{card}(A \cap B) = k$, atunci $\text{card}(A \cup B) = n + m - k$.

Demonstrație (opțional). Folosind proprietățile operațiilor asupra mulțimilor, deducem că oricare ar fi mulțimile $A, B \subseteq \Omega$ avem:

$$A = (A \cap B) \cup (A \cap \bar{B}), B = (A \cap B) \cup (\bar{A} \cap B), \\ A \cup B = (A \cap B) \cup (A \cap \bar{B}) \cup (\bar{A} \cap B).$$

Conform principiului adunării în cazul disjunct obținem:

$$\text{card}A = \text{card}(A \cap B) + \text{card}(A \cap \bar{B}),$$

$$\text{card}(B) = \text{card}(A \cap B) + \text{card}(\bar{A} \cap B),$$

$$\text{card}(A \cup B) = \text{card}(A \cap B) + \text{card}(A \cap \bar{B}) + \text{card}(\bar{A} \cap B) + \text{card}(A \cap B)$$

$$- \text{card}(A \cap B) = \text{card}(A) + \text{card}(B) - \text{card}(A \cap B) = n + m - k. \square$$

Remarcă. Folosind inducția matematică putem deduce Principiul adunării pentru un număr arbitrar k de mulțimi finite A_1, A_2, \dots, A_k :

$$\text{card} \bigcup_{i=1}^k A_i = \sum_{i=1}^k \text{card} A_i - \sum_{1 \leq i < j \leq k} \text{card}(A_i \cap A_j) + \dots \\ + (-1)^{m-1} \cdot \sum_{1 \leq i_1 < i_2 < \dots < i_m \leq k} \text{card}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_m}) + \dots \\ + (-1)^{k-1} \cdot \text{card}(A_1 \cap A_2 \cap \dots \cap A_k)$$

Exemplul 1. Considerăm o grupă de studenți despre care știm că fiecare student posedă cel puțin una din limbi străine: 20 de studenți cunosc limba engleză, 15 limba franceză, 10 limba germană, 5 limbile engleză și franceză, 5 limbile franceză și germană, 4 limbile engleză și germană și 1 student limbile engleză, franceză și germană. Câți studenți sunt în grupă ?

Notând prin E, F și G mulțimile de studenți care posedă, respectiv, limba engleză, franceză, germană și ținând cont de datele problemei, deducem: $\text{card}E = 20, \text{card}F = 15, \text{card}G = 10, \text{card}(E \cap F) = 5, \text{card}(E \cap G) = 4, \text{card}(F \cap G) = 5, \text{card}(E \cap F \cap G) = 1$ și atunci

$$\text{card}(E \cup F \cup G) = \text{card}(E) + \text{card}(F) + \text{card}(G) - \text{card}(E \cap F) - \\ \text{card}(E \cap G) - \text{card}(F \cap G) + \text{card}(E \cap F \cap G) = 32.$$

Principiul înmulțirii (în limbajul produsului cartezian). Dacă A și B sunt două mulțimi finite astfel încât $\text{card}(A) = n$ și $\text{card}(B) = m$, atunci

$$\text{card}(A \times B) = n \cdot m.$$

Demonstrație (optional). Este evident ca dacă

$$A = \{a_1, a_2, \dots, a_n\}, B = \{b_1, b_2, \dots, b_m\},$$

atunci mulțimile $A \times B$ și $\{(i, j) \mid i = \overline{1, n}, j = \overline{1, m}\}$ au același număr de elemente. Dacă în sistemul cartezian de coordonate xOy vom plasa valorile $i = \overline{1, n}$ pe axa Ox iar valorile $j = \overline{1, m}$ pe axa Oy , atunci elementului (i, j) îi corespunde punctul (i, j) din planul xOy , având, astfel, în planul xOy o rețea de $n \cdot m$ puncte. \square

Pentru orice număr k de mulțimi finite, aplicând metoda inducției matematice, putem demonstra că are loc formula:

$$\text{card}(A_1 \times A_2 \times \dots \times A_n) = \prod_{i=1}^n \text{card}A_i.$$

Demonstrația ei se bazează esențial pe faptul ca are loc egalitatea $A_1 \times A_2 \times \dots \times A_i = (A_1 \times A_2 \times \dots \times A_{i-1}) \times A_i$, $i = \overline{2, k}$.

Principiul înmulțirii în limbajul produsului cartezian poate fi reformulat în limbajul acțiunilor.

Principiul înmulțirii (în limbajul acțiunilor). Dacă o acțiune poate fi realizată în k etape succesive astfel încât etapa i poate fi realizată în n_i modalități, $i = \overline{1, k}$, atunci această acțiune poate fi realizată în $n_1 \cdot n_2 \cdot \dots \cdot n_k$ modalități.

Exemplul 2. Presupunem ca un safeu poate fi deschis cunoscând un cod de forma

$$i_1 \ i_2 \ i_3 \ i_4 \ i_5 \ i_6 ,$$

unde $i_k = \overline{0, 9}$, $k = \overline{1, 6}$. Câte coduri diferite de acest gen există?

Mulțimea Ω a tuturor codurilor de acest gen coincide cu produsul cartezian a mulțimii $\{0, 1, 2, \dots, 9\}$ de 6 ori cu ea însăși, adică

$$\Omega = \{(i_1, i_2, \dots, i_6) \mid i_k \in \{0, 1, 2, \dots, 9\}, k = \overline{1, 6}\}.$$

Acesta are, conform Principiului înmulțirii în limbajul produsului cartezian, 10^6 elemente (coduri).

Dacă avem informația ca acest cod este format din cifre diferite atunci vom observa ca Principiul înmulțirii în limbajul acțiunilor se aplică mai ușor decât Principiului înmulțirii în limbajul produsului cartezian. Într-adevăr, a forma un cod din 6 cifre diferite este echivalent cu a efectua o acțiune în 6 etape succesive, astfel încât prima etapă poate fi realizată în 10 modalități, cea de a doua în 9 modalități, etc., ultima (a șasea) în $10 - (6 - 1) = 5$ modalități. Conform *Principiului înmulțirii în limbajul acțiunilor*, numărul tuturor codurilor este egal cu $10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 = A_{10}^6$. Mai observăm că răspunsul din exemplul nostru se exprimă prin *formula aranjamentelor*.

Exemplul 3. Presupunem ca avem o mulțime Ω formată din n elemente. Atunci numărul tuturor submulțimilor pe care le putem forma din ea este egal cu

$$\text{card}\{A \mid A \subseteq \Omega\} = 2^n.$$

Într-adevăr, pentru a forma o submulțime a lui Ω e ca și cum ai realiza o acțiune în n etape succesive: la etapa cu nr. i decidem ce facem cu elementul cu nr. i , îl includem sau nu în mulțime, adică fiecare etapă, $i = \overline{1, n}$, poate fi realizată în 2 modalități. Apelând la Principiul înmulțirii obținem formula de mai sus.

Definiția 1. Fie A o mulțime formată din n elemente diferite, $A = \{a_1, a_2, \dots, a_n\}$, atunci vom numi *aranjament din n elemente luate câte k* orice mulțime ordonată de forma $(a_{i_1}, a_{i_2}, \dots, a_{i_k})$ cu proprietatea că $i_1 \neq i_2 \neq \dots \neq i_k$, $a_{i_j} \in A$, $i_j = \overline{1, n}$, $j = \overline{1, k}$. Evident, noțiunea are sens pentru $k = \overline{1, n}$. Mulțimea tuturor aranjamentelor de n elemente luate câte k se notează cu \mathcal{A}_n^k , adică

$$\mathcal{A}_n^k = \{(a_{i_1}, a_{i_2}, \dots, a_{i_k}) \mid i_1 \neq i_2 \neq \dots \neq i_k, a_{i_j} \in A, i_j = \overline{1, n}, j = \overline{1, k}\}.$$

Cardinalul acestei mulțimi se notează cu A_n^k și este numărul tuturor aranjamentelor din n elemente luate câte k .

Conform principiului înmulțirii în limbajul acțiunilor, a construi un *aranjament din n elemente luate câte k* este echivalent cu a realiza o acțiune în k etape succesive, astfel încât prima etapă poate fi realizată în n modalități, cea de a doua în $n-1$ modalități, etc., ultima (etapa nr. k) în $n - (k-1) = n - k + 1$ modalități. Or, numărul tuturor aranjamentelor din n elemente luate câte k este egal cu

$$A_n^k = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot (n-k+1) = \frac{n!}{(n-k)!}$$

Prin definiție, atunci când $k = n$, aranjamentul se numește *permutare de n elemente*. Deci mulțimea tuturor permutărilor de n elemente notată prin \mathbf{P}_n coincide cu \mathcal{A}_n^n , ceea ce înseamnă ca *numărul tuturor permutărilor de n elemente* P_n este egal cu A_n^k , adică $P_n = n!$.

Definiția 2. Orice submulțime de forma $\{a_{i_1}, a_{i_2}, \dots, a_{i_k}\}$, $i_1 \neq i_2 \neq \dots \neq i_k$, $a_{i_j} \in A$, $i_j = \overline{1, n}$, $j = \overline{1, k}$, se numește *combinare din n elemente luate câte k* . Evident noțiunea are sens pentru $k = \overline{1, n}$. Mulțimea tuturor combinațiilor de n elemente luate câte k elemente o vom nota prin

$$\mathbb{C}_n^k = \{\{a_{i_1}, a_{i_2}, \dots, a_{i_k}\} \mid i_1 \neq i_2 \neq \dots \neq i_k, a_{i_j} \in A, i_j = \overline{1, n}, j = \overline{1, k}\}.$$

Cardinalul acestei mulțimi îl vom nota cu \mathbb{C}_n^k .

Observăm ca dintr-o *combinare din n elemente luate câte k* putem forma $k!$ *aranjamente din n elemente luate câte k* . Or, a forma un aranjament din n elemente luate câte k este echivalent cu a realiza o acțiune în două etape succesive:

1. alegem o *combinare din n elemente luate câte k* , etapă pentru care avem \mathbb{C}_n^k modalități de a o efectua;
2. din această *combinare*, formăm un *aranjament din n elemente luate câte k* , etapă care se poate realiza în $k!$ modalități.

Rezultă că $A_n^k = k! \cdot \mathbb{C}_n^k$, adică

$$\mathbb{C}_n^k = \frac{A_n^k}{k!} = \frac{n!}{k!(n-k)!}.$$

Exemplul 3. Considerăm că avem o mulțime de n elemente astfel încât n_1 elemente sunt de tipul 1, n_2 elemente sunt de tipul 2, \dots , n_k elemente sunt de tipul k , $n_1 + n_2 + \dots + n_k = n$. Alegem la întâmplare, unul câte unul, toate elementele mulțimii și le aranjăm în ordinea extragerii lor. Să se calculeze cardinalul mulțimii Ω a tuturor rezultatelor posibile ce corespund acestui experiment.

Pentru a alcătui un rezultat posibil ca element a mulțimii Ω ce corespunde acestui experiment este suficient să realizăm o acțiune în k etape succesive.

Etapa 1: din n locuri disponibile pentru a aranja elementele extrase, alegem n_1 locuri pe care vom plasa elementele de tipul 1. Această acțiune o putem realiza în $\mathbb{C}_n^{n_1}$ modalități;

Etapa 2: din cele $n - n_1$ locuri, disponibile dupa etapa 1, alegem n_2 locuri pe care vom plasa elementele de tipul 2. Această acțiune o putem realiza în $\mathbb{C}_{n-n_1}^{n_2}$ modalități, etc.,

Etapa k : din cele $n - n_1 - n_2 - \dots - n_{k-1} = n_k$ locuri, disponibile dupa etapa k , alegem n_k locuri pe care vom plasa elementele de tipul k . Această acțiune o putem realiza în $\mathbb{C}_{n-n_1-n_2-\dots-n_{k-1}}^{n_k} = \mathbb{C}_{n_k}^{n_k}$ modalități.

Conform principiului înmulțirii, avem :

$$\text{card}(\Omega) = \mathbb{C}_n^{n_1} \cdot \mathbb{C}_{n-n_1}^{n_2} \cdot \dots \cdot \mathbb{C}_{n-n_1-n_2-\dots-n_{k-1}}^{n_k} = \frac{n!}{n_1!n_2! \cdot \dots \cdot n_k!}.$$

Formula obținută este, de fapt, formula de calcul pentru $P(n_1, n_2, \dots, n_k)$, numărul permutărilor a n elemente, din care n_1 elemente sunt de tipul 1, n_2 elemente sunt de tipul 2, \dots , n_k elemente sunt de tipul k , $n_1 + n_2 + \dots + n_k = n$:

$$P(n_1, n_2, \dots, n_k) = \frac{n!}{n_1!n_2! \cdot \dots \cdot n_k!}.$$

Ultima mai poarta denumirea de **formula permutarilor cu repetare**.

Exemplul 4. Presupunem că avem la dispoziție 10 cartonașe marcate cu litere astfel: $M, M, A, A, A, T, T, I, E, C$. Un copil se joacă, extăgând la întâmplare câte un cartonaș și aranjându-l în ordinea extragerii. Care este probabilitatea să obținem cuvântul MATEMATICA?

Întrucât considerăm cartonașele marcate la fel ca fiind de același tip, rezultă că avem 2 cartonașe de tip M , 3 cartonașe de tip A , 2 cartonașe de tip T , 1 cartonaș de tip I , 1 cartonaș de tip E și 1 cartonaș de tip C . Folosind formula dedusă mai sus, cardinalul spațiului de evenimente elementare asociat experimentului este

$$\text{card}(\Omega) = \frac{10!}{3!2!1!1!1!}$$

Rezultă că probabilitatea obținerii cuvântului MATEMATICA este egală cu

$$\frac{1}{\text{card}(\Omega)} = \frac{3!2!1!1!1!}{10!} = \frac{1}{2 \cdot 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9 \cdot 10} = 3.3069 \times 10^{-6}$$

Exemplul 5. Presupunem că dispunem de n cutii și r bile identice. Plasăm bilele, una câte una, la întâmplare, în una din cutii. Să se calculeze cardinalul multimii tuturor variantelor de plasare a bilelor în cutii.

Soluție. În cele ce urmează vom reprezenta n cutii prin intermediul a $n + 1$ bare verticale, iar r bile prin intermediul a r asteriscuri. De exemplu, situația când 5 bile identice, fiind plasate în 3 cutii astfel încât în prima cutie nimeresc 0 bile, în cutia a doua 2 bile și în cutia a treia 3 bile poate fi reprezentată astfel:

$$|| ** | *** |;$$

iar situația când toate bilele nimeresc în prima cutie poate fi reprezentată astfel:

$$| ***** |||.$$

Or, pentru o astfel de reprezentare schematică avem nevoie de $n + 1$ locuri pentru bare (pereții cutiilor) și r locuri pentru asteriscuri (bile). Din exemplele aduse vedem că orice repartizare concretă a r bile identice în n cutii este univoc determinată de poziția a $n - 1$ bare (pereți) interioare și a r asteriscuri (bile) pe cele $r + n - 1$ locuri interioare, cele două bare (pereți) exterioare rămânând de fiecare dată fixe. Drept consecință alegerea a $n - 1$ locuri pentru bare (sau r locuri pentru asteriscuri) din totalul de $n + r - 1$ locuri, poate fi făcută în $\mathbb{C}_{n+r-1}^{n-1} = \mathbb{C}_{n+r-1}^r$ modalități, \mathbb{C}_{n+r-1}^r , fiind cunoscut ca numărul **combinarilor din n elemente luate câte r cu repetare**.

Exemplul 6. Într-o cafenea sunt expuse 7 tipuri de înghețată. O familie formată din 4 persoane (părinții cu doi copii) a comandat la întâmplare 4 înghețate. În câte modalități poate fi realizată această comandă, care este privită ca o mulțime neordonată formată din 4 înghețate?

Soluție. Observăm că problema noastră se reduce la problema calculării numărului de *combinari din $n = 7$ elemente luate câte $r = 4$ cu repetare*, tipul înghețatei fiind considerat, convențional, cutie, iar cele 4 înghețate-bile identice. Prin urmare, comanda în cauză poate fi realizată în $\mathbb{C}_{7+4-1}^4 = 210$ modalități.

Probleme propuse.

1. Un om de afaceri are de ales unul din traseele turistice pe care le propun 2 firme turistice. Câte variante de alegere are acesta dacă:

a) Prima firmă propune 10 trasee turistice iar cea de a doua 20 de trasee, astfel încât în ofertele ambelor firme nu se regăsește niciun traseu comun;

b) Prima firmă propune 10 trasee turistice iar cea de a doua 20 de trasee, astfel încât 5 trasee se regăsesc în ofertele ambelor firme.

2. Producătorul de autoturisme Dacia Pitești produce mai multe versiuni de model noul Dacia Sandero Stepway care depind de unul din 2 tipuri de alimentare combustibil (benzină/motorină), 2 variante de volum de motor, 2 tipuri de transmisie (manuală sau automată), 3 pachete dotari interioare și 4 culori vopsea pentru exterior. Câte versiuni autoturism în total trebuie să solicite de la producător un dealer autorizat care vinde acest model de autoturism?

3. Într-un autoturism cu 5 locuri, inclusiv locul șoferului, se așează, la întâmplare, 5 persoane, din care numai 2 persoane posedă permis de conducere, fiecare persoana ocupând un singur loc. Cu ce este egal numărul total de ocupare a celor 5 locuri? Câte variante de ocupare a locurilor favorizează evenimentul că la volanul autoturismului va nimeri o persoană care posedă permis de conducere.

4. Dintr-o populație statistică, reprezentată de toți studenții unei universități, a fost efectuat un sondaj bazat pe un eșantion de volum 500, ce vizează atitudinea studenților față de fumat. Câte eșantioane diferite de volum 500 sunt posibile în total dacă numărul total al studenților este egal cu 11000, prin eșantion înțelegând orice submulțime ordonată formată din 500 de studenți aleși la întâmplare din aceasta populație statistică? Câte submulțimi diferite de acest fel sunt posibile dacă:

- a) Eșantionarea/selecția este făcută fără repetare, adică studentul, odată fiind ales, este exclus din lista tuturor studenților din care se face selectarea;
- b) Eșantionarea/selecția este făcută cu repetare, adică studentul ales poate apare în eșantion în mod repetat, fiind lăsat de fiecare dată în listă.

Care este numărul total de eșantioane diferite de volum 500, posibile în acest sondaj dacă în eșantion sunt incluși *nu* studenții aleși, ci răspunsurile lor, *DA* sau *NU*, în funcție de este sau nu fumător studentul chestionat? În acest caz, depinde oare acest număr de tipul de eșantionare (cu sau fara repetare)? Dar de numărul N de studenți care studiază la universitate, unde, bineînțeles, $N \geq 500$, în cazul selecției fără repetare?

5. Registrul Auto Român (RAR) a ales, în calitate de tip de numerotare a autoturismelor, o secvență formată din acronimul județului sau a mun. București, funcție de viza de reședință sau adresa persoanei fizice sau juridice, ce deține autoturismul, urmata de 2 cifre și 3 litere (fără diacritice) al alfabetului latin. Care este numărul maxim de autoturisme care pot fi numerotate astfel într-un județ sau mun București? Dar numărul maxim de autoturisme care pot fi numerotate astfel în România, știind că România este împărțită în 41 de județe.

6. La pizzeria Andy's au fost aduse pentru ziua curenta, n componente alimentare pentru producerea Pizzei, turtele pentru Pizza fiind produse aparte la pizerie. Câte tipuri de Pizza pot fi produse daca fiecare tip admite, cel puțin, una din aceste componente sau, cel mult, toate componentele aduse. Dar câte tipuri de Pizza pot fi produse pentru $n = 6$ daca numarul de componente folosite:

a) trebuie sa fie egal cu 3;

b) poate fi, cel puțin, egal cu 1 și, cel mult, egal cu 5.

7. Un reprezentant pentru vânzări din SUA urmează să întreprindă vizite de lucru în orașele Omaha, Dallas, Wichita și Oklahoma Sity, între care exista legătură directă cu avionul. Presupunem că ordinea vizitării lor este aleasă la întâmplare. Câte rezultate posibile corespund acestui experiment aleator?

8. În condițiile problemei anterioare reprezentantul de vânzări are sarcina sa aleagă la întâmplare si să viziteze doar trei din orașele enumerate, nu contează în ce ordine. Câte rezultate posibile corespund acestui experiment aleator?

10. Un grup de 8 participanții la o Conferință Internațională, printre care Ionescu și Petrescu, au fost cazați și repartizați de către organizatori, la întâmplare, într-un hotel particular ce are numai 4 camere: una cu 3 locuri, două cu 2 locuri si una cu un singur loc. Câte rezultate posibile sunt în acest experiment aleator? Dar câte dintre aceste rezultate favorizează faptul ca Ionescu și Petrescu vor nimeri în camera de 3 locuri?

11. Secția Literatură Tehnică a Bibliotecii Naționale "Vasile Alecsandri" din Chisinau are cărți ce țin de domeniile Matematică, Fizică, Chimie, etc., în total 16 domenii ale Științei. Experimentul rezidă în alegerea la întâmplare a 4 comenzi de carte pentru a vedea cum se repartizează acestea pe domenii. Câte rezultate posibile sunt în acest experiment aleator? Câte din aceste rezultate favorizează faptul ca toate comenzile se referă la domenii de Știința diferite? Dar câte din aceste rezultate favorizează faptul ca toate comenzile se referă la unul si același domeniu de Știința diferite?

1.3. Spații de evenimente elementare, evenimente aleatoare și operații asupra lor, definiția axiomatică a probabilității

A modela matematic un experiment aleator înseamnă, de fapt, a descrie matematic **(a)** *mulțimea de rezultate posibile* în acest experiment, **(b)** *evenimentelor aleatoare* asociate acestui experiment și **(c)** *probabilitatea*, adică

aplicația sau regula conform careia fiecărui eveniment aleator îi punem în corespondență un număr ce caracterizează șansele (gradul) lui de realizare. Vom începe cu dezideratul (a).

Spații de evenimente elementare

Chiar dacă într-un experiment aleator nu putem anticipa cu certitudine care rezultat anume se va produce, este firesc să presupunem că mulțimea tuturor rezultatelor posibile poate fi descrisă cu exactitate. Aducem, drept confirmare, câteva exemple.

Exemplul 1. Experimentul \mathcal{E} constă în aruncarea unei monede o singură dată. Mulțimea de rezultate posibile este dată de

$$\Omega = \{S, B\} = \{0, 1\} = \{\omega_1, \omega_2\},$$

unde prin $S, 0$ sau ω_1 se subînțelege "stema" iar prin $B, 1$ sau ω_2 se subînțelege "banul".

Exemplul 2. Experimentul \mathcal{E} constă în aruncarea unei monede de trei ori succesiv. Mulțimea de rezultate posibile este dată de

$$\Omega = \{SSS, SSB, SBS, SBB, BSS, BSB, BBS, BBB\}.$$

Exemplul 3. Experimentul \mathcal{E} constă în aruncarea zarului o singură dată. Mulțimea de rezultate posibile este dată de

$$\Omega = \{1, 2, 3, 4, 5, 6\} = \{i \mid i = \overline{1, 6}\} = \{\omega_i \mid i = \overline{1, 6}\}.$$

Exemplul 4. Experimentul \mathcal{E} constă în aruncarea unei monede până la prima apariție a stemei. Mulțimea de rezultate posibile este dată de

$$\Omega = \{S, BS, BBS, BBBS, \dots, \underbrace{BB \dots BS}_{n-1 \text{ ori}}, \dots\}$$

care este o *mulțime infinită numărabilă*, unde prin $\underbrace{BB \dots BS}_{n-1 \text{ ori}}$ se subînțelege că prima apariție a "stemei" este precedată de apariția "banului" de $n - 1$ ori succesiv, $n \geq 1$.

Exemplul 5. Experimentul \mathcal{E} constă în alegerea unui punct la întâmplare pe segmentul $[0, 1]$. Mulțimea de rezultate posibile este dată de *mulțimea infinită nenumărabilă*

$$\Omega = \{x \mid x \in [0, 1]\},$$

unde x este coordonata punctului ales.

Exemplul 6. Experimentul \mathcal{E} constă în înregistrarea greutății unui student ales la întâmplare de la una din universități. Mulțimea de rezultate posibile este dată de

$$\Omega = \{x \mid x > 0\}.$$

Remarca 1. Deoarece multimile finite sau infinite, cel mult, numărabile se mai numesc *mulțimi discrete*, spunem că exemplele 1-4 se referă la cazul *discret*, în caz contrar, cum sunt exemplele 5-6, spunem că se referă la cazul *continuu*. Cu toate acestea exemplele invocate ne conduc la următoarea definiție valabilă în *caz general*.

Definiția 1. Vom numi *spațiu de evenimente elementare* orice mulțime nevidă $\Omega \neq \emptyset$, ale cărei elemente reprezintă (descriu) toate rezultatele posibile într-un experiment aleator. Fiecare element al mulțimii Ω (care corespunde unui singur rezultat posibil) se numește *eveniment elementar*.

Remarca 2. Definiția nu exclude situația, când unuia și aceluiași experiment aleator îi putem asocia, în dependență de scopul urmărit, spații de evenimente elementare diferite. Astfel, în exemplul 2 dacă ne interesează numai numărul de steme apărute la aruncarea monedei de trei ori, atunci în calitate de spațiu de evenimente elementare putem lua

$$\Omega = \{0, 1, 2, 3\}.$$

Acum suntem pregătiți să răspundem dezideratului (b).

Evenimente aleatoare și operații asupra lor

Pentru a introduce noțiunea matematică cu ajutorul căreia să descriem evenimentele aleatoare vom pleca de la exemplul legat de aruncarea zarului o singură dată pentru care spațiul corespunzător de evenimente elementare este $\Omega = \{1, 2, 3, 4, 5, 6\}$. Considerăm următoarele evenimente aleatoare pe care le vom nota, aici și în continuare, cu primele majuscule ale alfabetului latin: $A = \{\text{numărul de puncte va fi par}\}$; $B = \{\text{numărul de puncte va fi impar}\}$; $C = \{\text{numărul de puncte va fi mai mic sau egal cu 4}\}$; $D = \{\text{numărul de puncte va fi mai mic decât 13}\}$; $E = \{\text{numărul de puncte va fi mai mare decât 13}\}$; $F = \{\text{numărul de puncte va fi divizibil cu 3}\}$. Observăm că putem stabili următoarea corespondență biunivocă: $A \longleftrightarrow \{2, 4, 6\}$; $B \longleftrightarrow \{1, 3, 5\}$; $C \longleftrightarrow \{1, 2, 3, 4\}$; $D \longleftrightarrow \{1, 2, 3, 4, 5, 6\} = \Omega$; $E \longleftrightarrow \emptyset$; $F \longleftrightarrow \{3, 6\}$.

În concluzie, dacă experimentului aleator îi corespunde un spațiu discret de evenimente elementare Ω , atunci evenimentele aleatoare pot fi descrise ca fiind submulțimi ale lui Ω . Așadar, având un spațiu de evenimente elementare Ω , exemplul nostru arata ca putem, cel puțin în caz discret, formula

Definiția 2. Vom numi *eveniment aleator (în caz discret)* orice submulțime $A \subseteq \Omega$. În particular, Ω se va numi *evenimentul sigur*, iar \emptyset se va numi *eveniment imposibil*. Spunem că evenimentul elementar $\omega \in \Omega$ favorizează apariția evenimentului aleator A dacă și numai dacă $\omega \in A$.

Astfel, în exemplul de mai sus, evenimentul $D = \Omega$ și se numește *eveniment sigur*, iar evenimentul E - *eveniment imposibil*.

Or, operațiile care pot fi aplicate asupra mulțimilor pot fi aplicate și asupra evenimentelor aleatoare.

Definiția 3. *Sumă (reuniune)* a două evenimente aleatoare $A, B \subseteq \Omega$ vom numi evenimentul

$$C = A \cup B = \{\omega \in \Omega \mid \omega \in A \text{ sau } \omega \in B\},$$

adică C se produce dacă *sau* se produce A , *sau* se produce B , *sau* se produce A și B concomitent. Cu alte cuvinte, *suma* evenimentelor A și B se produce atunci și numai atunci când se produce *cel puțin unul* din aceste două evenimente.

În exemplul nostru: $A \cup F = \{2, 3, 4, 6\}$, $A \cup B = \Omega$.

Definiția 4. *Produs (intersecție)* a două evenimente aleatoare $A, B \subseteq \Omega$ vom numi evenimentul

$$C = A \cap B = \{\omega \in \Omega \mid \omega \in A \text{ și } \omega \in B\},$$

adică C se produce dacă se produce *și* A și B . Cu alte cuvinte evenimentul *produs al* evenimentelor A și B are loc atunci și numai atunci când aceste două evenimente se produc concomitent. Produsul a două evenimente A și B se mai notează $A \cdot B$ sau AB .

În exemplul nostru: $A \cap F = \{6\}$, $A \cap B = \emptyset$.

Definiția 5. Pentru orice $A \subseteq \Omega$ vom numi *eveniment "non- A "* sau *eveniment opus evenimentului A* complementara acestei submulțimi în raport cu Ω , adică evenimentul notat cu \bar{A} sau A^c , unde

$$A^c = \{\omega \in \Omega \mid \omega \notin A\}$$

Or, evenimentul \bar{A} se produce atunci și numai atunci când *nu* se produce evenimentul A .

În exemplul nostru: $\bar{A} = B$, $\bar{B} = A$, $\bar{D} = E$, $\bar{E} = D$.

Definiția 6. Dacă avem două evenimente aleatoare $A, B \subseteq \Omega$ spunem că evenimentul A *implică* evenimentul B dacă și numai dacă $A \subseteq B$. Altfel spus, A *implică* B atunci și numai atunci când din faptul că s-a produs evenimentul A rezultă că s-a produs și evenimentul B . Dacă $A \subseteq B$ și $B \subseteq A$, atunci spunem că $A = B$ (A este *echivalent* cu B).

Cum operațiile asupra evenimentelor aleatoare sunt, de fapt, operații asupra mulțimilor din Ω putem formula

Propoziția 1. *Operațiile asupra evenimentelor aleatoare au următoarele proprietăți:* $A \cup A = A$; $A \cup \Omega = \Omega$; $A \cap A = A$; $A \cap \Omega = \Omega$; $A \cup \emptyset = A$; $A \cap \emptyset = \emptyset$; $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$; $\overline{A \cup B} = \bar{A} \cap \bar{B}$; $\overline{A \cap B} = \bar{A} \cup \bar{B}$.

Remarca 3. Ultimele două proprietăți se numesc formulele de *dualitate* ale lui De Morgan. Prin analogie, operațiile de *sumă* și *produs* pot fi extinse asupra evenimentelor $A_i \subseteq \Omega$, $i = 1, 2, \dots$, adică putem opera cu evenimentele $:\bigcup_{n=1}^k A_n, \bigcup_{n=1}^{\infty} A_n, \bigcap_{n=1}^k A_n, \bigcap_{n=1}^{\infty} A_n$, unde $k \geq 2$.

Definiția 7. Dacă $A, B \subseteq \Omega$ sunt astfel încât $A \cap B = \emptyset$, adică A, B nu se pot produce concomitent, atunci spunem că A și B sunt *evenimente incompatibile* (sau *disjuncte*).

Se verifică cu ușurință că în caz discret familia $\mathcal{F} = \{A \mid A \text{ este eveniment aleator legat de experimentul aleator cărui îi corespunde spațiul de evenimente elementare } \Omega\}$ coincide cu familia tuturor submulțimilor spațiului de evenimente elementare Ω , adică cu $\mathcal{F} = \{A \mid A \subseteq \Omega\}$ și aceasta verifică următoarele axiome (proprietăți):

1. Dacă $A \subseteq \mathcal{F}$, atunci $\bar{A} \in \mathcal{F}$;
2. Dacă $A_1, A_2, \dots, A_n, \dots \in \mathcal{F}$, atunci $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Definiția 8. Familia \mathcal{F} de submulțimi ale lui Ω care verifică axiomele de mai sus se numește câmp (borelian) de evenimente. Elementele lui \mathcal{F} se numesc evenimente aleatoare, iar spațiul de evenimente elementare Ω înzestrat cu un câmp de evenimente \mathcal{F} se numește *spațiu măsurabil* și se notează prin (Ω, \mathcal{F}) .

Această din urmă definiție a noțiunii de eveniment aleator este acoperitoare pentru ambele cazuri, discret și continuu, adică este valabilă în caz general. Mai mult, are loc următoarea

Propoziție 2. Dacă (Ω, \mathcal{F}) este un *spațiu măsurabil* atunci au loc următoarele proprietăți:

a) *Evenimentele sigur și imposibil fac parte și ele din câmpul de evenimente aleatoare, adică $\Omega \in \mathcal{F}$, $\emptyset \in \mathcal{F}$;*

b) *Dacă $A_1, A_2, \dots, A_n, \dots \in \mathcal{F}$, atunci și $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$.*

Cu alte cuvinte, dacă familia \mathcal{F} de submulțimi ale lui Ω este un câmp (borelian) de evenimente, atunci garantat că evenimente aleatoare vor fi mulțimile Ω și \emptyset cât și complementarele tuturor submulțimilor din Ω care fac parte din \mathcal{F} , dar și intersecția (produsul) tuturor submulțimilor $A_1, A_2, \dots, A_n, \dots$ din Ω ce fac parte din \mathcal{F} , nu numai reuniunea (suma) lor. Lucrul acesta este suficient pentru a modela matematic, toate situațiile care prezintă interes din punct de vedere practic în legătura cu evenimentele aleatoare pentru care dorim să definim noțiunea de probabilitate.

Acum, pentru a întregi modelarea matematică a experimentelor aleatoare, avem la dispoziție toate datele pentru a defini noțiunea generală de probabilitate, răspunzând, astfel, la dezideratul (c).

Pentru a da o definiție cât mai generală, așa cum arată ilustrul matematician rus A. N. Kolmogorov, părintele teoriei moderne a probabilităților, se impune o abordare axiomatică. De notat că abordarea axiomatică la construirea unei teorii nu este străină, de exemplu, Mecanicii Teoretice sau chiar și științelor economice. Astfel, Teoria Consumului se bazează pe un set de supoziții comportamentale ale consumatorilor, supoziții considerate ca fiind adevăruri axiomatiche. În cazul nostru, deoarece Probabilitatea este privită ca măsură a gradului de realizare a oricărui eveniment aleator asociat unui experiment aleator, este firesc să o privim ca pe o funcție (aplicație) definită pe multimea/familia tuturor evenimentelor aleatoare legate de acest experiment, adică pe multimea \mathcal{F} definită mai sus.

Așadar, fie (Ω, \mathcal{F}) un spațiu măsurabil, atunci putem formula

Definiția axiomatică a probabilității. Vom numi *probabilitate* orice aplicație (funcție) $P: \mathcal{F} \rightarrow \mathbb{R}$ dacă aceasta satisface

Axioma 1. Pentru orice eveniment $A \in \mathcal{F}$ probabilitatea lui $P(A) \geq 0$;

Axioma 2. Probabilitatea evenimentului sigur $P(\Omega) = 1$;

Axioma 3. Probabilitatea producerii a cel puțin unuia din evenimentele $A_1, A_2, \dots, A_n, \dots$ din \mathcal{F} , incompatibile (disjuncte) două câte două, coincide cu suma probabilităților acestor evenimente, adică $P(A_1 \cup A_2 \cup \dots \cup A_n \cup \dots) = P(A_1) + P(A_2) + \dots + P(A_n) + \dots$, pentru $\forall \{A_n\}_{n \geq 1} \in \mathcal{F}$, $A_i \cap A_j = \emptyset$, $i \neq j$, $i, j \geq 1$.

În concluzie, probabilitatea P fiind o aplicație (funcție) definită pe familia \mathcal{F} a tuturor evenimentelor aleatoare asociate experimentului dat, este firesc

sa numim probabilitate a evenimentului A valoarea $P(A)$ a acestei funcții.

Remarca 4. Dacă confruntăm proprietățile frecvențelor relative cu Definiția Axiomatică a Probabilității, deducem că aceasta din urmă este acoperitoare, de fapt, pentru proprietățile Probabilității Frecvențiale (Statistice).

Definiția 9. Vom numi *câmp de probabilitate* sau *spațiu de probabilitate* orice triplet de forma (Ω, \mathcal{F}, P) , unde (Ω, \mathcal{F}) este un spațiu măsurabil asociat unui experiment aleator iar P este o probabilitate definită pe \mathcal{F} .

Notiunea de *câmp de probabilitate* poate servi în calitate de model matematic (model probabilist) care descrie comportamentul probabilist al evenimentelor aleatoare asociate experimentului aleator în cauză. Drept confirmare putem aduce următoarele exemple.

Exemplul 7. Considerăm experimentul aleator \mathcal{E} ce constă în aruncarea unui zar "perfect" o singura dată. Atunci spațiul de evenimente elementare corespunzător $\Omega = \{1, 2, 3, 4, 5, 6\}$ descrie toate rezultatele posibile iar familia $\mathcal{F} = \{A \mid A \subseteq \Omega\}$ ale tuturor submulțimilor din Ω descrie toate evenimentele aleatoare asociate acestui experiment. Observăm că aplicația $P: \mathcal{F} \rightarrow \mathbb{R}$ conform formulei

$$P(A) = \frac{\text{Card } A}{\text{Card } \Omega} = \frac{1}{6} \text{Card } A,$$

este o probabilitate în sensul definiției de mai sus. Prin urmare tripletul (Ω, \mathcal{F}, P) este un câmp de probabilitate ce descrie comportamentul probabilist al oricarui eveniment aleator asociat experimentului \mathcal{E} . De exemplu probabilitatea evenimentului $A = \{la \text{ aruncarea unui zar "perfect" o singură dată va apare un număr par de puncte}\}$ este egală cu $P(A) = \frac{1}{6} \text{Card } \{2, 4, 6\} = \frac{3}{6} = \frac{1}{2}$.

Remarca 5. Finalmente, verificarea faptului ca modelul probabilist (Ω, \mathcal{F}, P) descrie adecvat experimentul aleator \mathcal{E} se poate face pe cale experimentală. Conform proprietății regularității statistice, dacă, odată cu creșterea lui n , pentru orice $A \in \mathcal{F}$, frecvențele relative $f_n(A) \simeq P(A)$, atunci ipoteza ca modelul în cauza este adecvat experimentului \mathcal{E} devine credibilă, dar validarea modelului presupune aplicarea unor metode avansate ce țin de Statistica Matematică.

Exemplul 8. Considerăm următorul joc de noroc: Ion și Petru aruncă pe rând o monedă "perfectă", jocul terminându-se de îndată ce unul din jucători va înregistra apariția stemei. Primul aruncă moneda Ion. Ne interesează, de exemplu, cine are șanse mai mari de a câștiga jocul. Conform Exemplului 3, spațiul de evenimente elementare $\Omega = \{S, BS, BBS,$

$BBBS, \dots, \underbrace{BB..BS}_{n-1 \text{ ori}}, \dots\} = \{\omega_1, \omega_2, \dots, \omega_n, \dots\}$. Se observă cu ușurință ca orice eveniment aleator asociat acestui joc este submultime a lui Ω . Prin urmare $\mathcal{F} = \{A \mid A \subseteq \Omega\}$. De exemplu evenimentele $\{\text{jocul va fi câștigat de Ion}\} = \{\omega_1, \omega_3, \dots, \omega_{2k+1}, \dots\}$, $\{\text{jocul va fi câștigat de Petru}\} = \{\omega_2, \omega_4, \dots, \omega_{2k}, \dots\}$. În continuare, asociem fiecărui eveniment elementar probabilitatea $P\{\omega_n\} = \frac{1}{2^n}$, $n = 1, 2, \dots$ iar fiecărui eveniment $A \in \mathcal{F}$ - probabilitatea calculată după formula $P(A) = \sum_{n:\omega_n \in A} P\{\omega_n\}$. Evident, $P(A) \geq 0$. Axioma 1 a probabilității este, astfel, îndeplinită. Este îndeplinită și Axioma 2 deoarece

$$P(\Omega) = \sum_{n \geq 1} \frac{1}{2^n} = \frac{1}{2} \left(1 + \frac{1}{2} + \frac{1}{2^2} + \dots + \frac{1}{2^n} + \dots\right) = \frac{1}{2} \cdot \frac{1}{1 - \frac{1}{2}} = 1.$$

Cum pentru $\forall \{A_n\}_{n \geq 1} \in \mathcal{F}$, $A_i \cap A_j = \emptyset$, $i \neq j$ avem ca

$$\begin{aligned} P(A_1 \cup A_2 \cup \dots \cup A_n \cup \dots) &= \sum_{n:\omega_n \in A_1 \cup A_2 \cup \dots \cup A_n \cup \dots} P(\omega_n) = \\ &= \sum_{n:\omega_n \in A_1} P(\omega_n) + \sum_{n:\omega_n \in A_2} P(\omega_n) + \dots = P(A_1) + P(A_2) + \dots \end{aligned}$$

rezulta ca si Axioma 3 este valabilă. Prin urmare funcția $P: \mathcal{F} \rightarrow \mathbb{R}$ definită mai sus este o probabilitate, iar (Ω, \mathcal{F}, P) un model de probabilitate pentru jocul de noroc descris anterior. Astfel,

$$P\{\text{jocul va fi câștigat de Ion}\} = P\{\omega_1, \omega_3, \dots, \omega_{2k+1}, \dots\} =$$

$$\frac{1}{2} + \frac{1}{2^3} + \frac{1}{2^5} + \dots = \frac{1}{2} \cdot \frac{1}{1 - \frac{1}{2^2}} = \frac{2}{3}.$$

În mod similar aflăm că $P\{\text{jocul va fi câștigat de Petru}\} = \frac{1}{3}$. Cu alte cuvinte, Ion are de 2 ori mai multe șanse de a câștiga jocul decât Petru.

Exemplul 9. Considerăm experimentul aleator ce constă în fixarea duratei x de viață, măsurată în ore, a memoriei hard a unui laptop marca HP. Evident, spatiul de evenimente elementare corespunzător este mulțimea $\Omega = \{x \mid x \geq 0\}$. Luăm în calitate de \mathcal{F} orice familie de submulțimi ale lui Ω dacă aceasta este un câmp borelian de evenimente, adică \mathcal{F} conține toate intervalele din Ω de forma $(a, b]$, $[a, b)$, (a, b) , reuniunile, intersecțiile si complementarele unor astfel de intervale. Aceasta ne permite sa afirmăm ca

aplicatia $P: \mathcal{F} \rightarrow \mathbb{R}$ dupa formula $P(A) = \int_A 10^{-6} e^{-10^{-6}x} dx$, privită ca funcție de multimii din \mathcal{F} , este o probabilitate. Într-adevăr, $P(A) \geq 0$, deoarece pentru orice $x \in A \subseteq \Omega$ funcția de sub integrală este nenegativă, iar

$$P(\Omega) = \int_{\Omega} 10^{-6} e^{-10^{-6}x} dx = \int_0^{+\infty} 10^{-6} e^{-10^{-6}x} dx = (1 - e^{-10^{-6}x}) \Big|_0^{+\infty} = 1,$$

și din proprietățile integralei rezultă că

$$\begin{aligned} P(A_1 \cup A_2 \cup \dots \cup A_n \cup \dots) &= \int_{A_1 \cup A_2 \cup \dots \cup A_n \cup \dots} 10^{-6} e^{-10^{-6}x} dx = \\ &= \int_{A_1} 10^{-6} e^{-10^{-6}x} dx + \int_{A_2} 10^{-6} e^{-10^{-6}x} dx + \dots + \int_{A_n} 10^{-6} e^{-10^{-6}x} dx + \dots = \\ &= P(A_1) + P(A_2) + \dots + P(A_n) + \dots, \end{aligned}$$

pentru $\forall \{A_n\}_{n \geq 1} \in \mathcal{F}, A_i \cap A_j = \emptyset, i \neq j, i, j \geq 1$. Aceasta înseamnă ca P verifică Axiomele 1-3 din definiția probabilității. Or, (Ω, \mathcal{F}, P) poate fi privit ca un model matematic ce descrie comportamentul probabilist al duratei vieții memoriei hard a unui laptop marca HP. Astfel, probabilitatea că acesta va avea o durată de viață mai mare decât T este egală cu

$$\int_{(T, +\infty)} 10^{-6} e^{-10^{-6}x} dx = \int_T^{+\infty} 10^{-6} e^{-10^{-6}x} dx = (1 - e^{-10^{-6}x}) \Big|_T^{+\infty} = e^{-10^{-6}T}.$$

Astfel, pentru $T = 43800$ ore, adică aproximativ 5 ani, această probabilitate este egală cu $e^{-10^{-6} \times 43800} = 0.95715$, ceea ce denotă o fiabilitate sporită pentru memoria hard a unui calculator marca HP. În ce măsură modelul matematic descris mai sus corespunde realității aceasta este o problemă ce ține de Statistica Matematică.

Probleme propuse.

1. Pentru fiecare dintre experimentele aleatoare de mai jos descrieți spațiul de evenimente elementare corespunzător.

a. La încheierea activității zilnice, contabilitatea centrului comercial UNIC duce evidența numărului de tranzacții în numerar ai primilor 100 de clienți.

b. Biroul Național de meteorologie înregistrează zilnic datele temperaturii pentru fiecare localitate. Interes prezintă, de exemplu, perechea de date ce constă din temperatura minimă și temperatura maximă exprimată în Celsius.

c. Compania Rompetrol asigură toate stațiile sale de alimentare cu combustibil inclusiv cu benzină fără plumb. Interes prezintă cantitatea totală de benzină fără plumb solicitată de toate stațiile de alimentare Rompetrol din Republica Moldova.

d. Managerul de birou al unei companii specializate în servicii de copiere contorizează numărul de copii pe care le produce o mașină de copiat înainte de a suferi un blocaj de hârtie. Interes prezintă anume acest număr de copii.

2. Pentru fiecare din exemplele aduse în problema anterioară să se specifice tipul de mulțime a rezultatelor posibile (mulțime finită, infinită numărabilă-caz discret, infinită nenumărabilă-caz continuu). Argumentați răspunsurile.

3. Pentru fiecare din exemplele **a-d**, aduse în problema **1**, să se descrie mulțimea (familia) \mathcal{F} de evenimente aleatoare (câmpul borelian de evenimente) pe care-l putem asocia (în mod firesc) experimentului corespunzător.

4. O firmă din Chișinău specializată în vânzarea autoturismelor noi japoneze are următoarea echipă de comercianți:

Prenume	Experiență de vânzare	Vârsta	Studiile	Căsătorit
Ion	4	34	liceale	<i>Da</i>
Doina	12	31	liceale	<i>Nu</i>
Mihai	21	56	superioare	<i>Da</i>
Gheorghe	9	42	liceale	<i>Da</i>
Elena	3	24	superioare	<i>Nu</i>
Mircea	7	29	liceale	<i>Nu</i>
Raluca	12	44	superioare	<i>Da</i>
Nicolae	2	25	liceale	<i>Nu</i>

Un client venit să discute procurarea unui autoturism nou alege la întâmplare un comerciant. Alegând metoda potrivită de asociere a probabilităților, calculați probabilitățile următoarelor evenimente:

a. Va fi aleasă o persoană de sex feminin; **b.** Va fi ales un barbat cu vârsta sub 40 de ani; **c.** Va fi aleasă o persoană cu o vechime de activitate în domeniul vânzărilor de cel puțin 10 ani; **d.** Va fi aleasă o persoană cu

studii superioare și căsătorită; **e.** Va fi aleasă o persoană de sex feminin , cu studii liceale și cu experiență de activitate în domeniul vânzărilor de cel puțin 5 ani; **f.** Va fi aleasă o persoană cu o vechime de activitate în domeniul vânzărilor de cel puțin 2 ani și vârsta de cel puțin 21 de ani; **g.** Va fi aleasă o persoană cu vârsta sub 60 de ani; **h.** Va fi aleasă o persoană cu vârsta su 24 de ani.

5. Pentru fiecare din următoarele exemple de triplete de forma (Ω, \mathcal{F}, P) determinați dacă acesta reprezintă un model probabilist (câmp de probabilitate).

a) Spațiul de evenimente elementare $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8\}$, câmpul de evenimente aleatoare $\mathcal{F} = \{A : A \subseteq \Omega\}$, aplicația (funcția) P definită pe \mathcal{F} este dată de formula $P(A) = \sum_{k \in \Omega: k \in A} \frac{k}{36}$, $\forall A \in \mathcal{F}$; b) $\Omega = [0, +\infty)$, $\mathcal{F} = \{A : A \text{ este un subinterval a lui } \Omega \text{ sau submultime a lui } \Omega \text{ ca reuniune, intersecție sau complementara unor astfel de intervale}\}$, $P(A) = \int_A e^{-x} dx$, $\forall A \in \mathcal{F}$; c) $\Omega = \{1, 2, \dots, n, \dots\}$, $\mathcal{F} = \{A : A \subseteq \Omega\}$, $P(A) = \sum_{k \in \Omega: k \in A} \frac{k^2}{10^5}$, $\forall A \in \mathcal{F}$; d) $\Omega = (0, 1)$, $\mathcal{F} = \{A : A \text{ este un subinterval a lui } \Omega \text{ sau submultime a lui } \Omega \text{ ca reuniune, intersecție sau complementara unor astfel de intervale}\}$, $P(A) = \int_A 12x(1-x)^2 dx$, $\forall A \in \mathcal{F}$.

1.4. Proprietățile probabilității drept consecință din definiția axiomatică a probabilității

Fie (Ω, \mathcal{F}, P) un spațiu de probabilitate. Din definiția axiomatică a probabilității P , privită ca funcție de submulțimi (evenimente) ale spațiului de evenimente elementare, rezulta un set de proprietăți generale ale probabilității ce simplifică procesul de calcul ale unor probabilități ale unor evenimente aleatoare ce se exprima prin alte evenimente aleatoare ale caror probabilitati sunt cunoscute sau se identifică mai ușor. Le vom centraliza in următoarea

Propoziție (Proprietățile probabilității). Orice probabilitate P definită pe câmpul de evenimente aleatoare \mathcal{F} posedă următoarele proprietăți:

a) $0 \leq P(A) \leq 1$ pentru orice eveniment A din \mathcal{F} ;

b) $P(\bar{A}) = 1 - P(A)$ sau în formă echivalentă, $P(A) = 1 - P(\bar{A})$, pentru orice eveniment A din \mathcal{F} ;

c) Probabilitatea evenimentului imposibil este egală cu zero, cu alte cuvinte $P(\emptyset) = 0$;

d) (**Formula adunării probabilităților**). Dacă A și B sunt două evenimente aleatoare legate de unul și același câmp de evenimente \mathcal{F} , atunci

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

sau, în caz general, dacă A_1, A_2, \dots, A_n sunt n evenimente aleatoare legate de unul și același câmp de evenimente \mathcal{F} , atunci

$$\begin{aligned} P\left(\bigcup_{i=1}^n A_i\right) &= \sum_{i=1}^n P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) + \dots \\ &+ (-1)^{m-1} \cdot \sum_{1 \leq i_1 < i_2 < \dots < i_m \leq n} P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_m}) + \dots \\ &+ (-1)^{n-1} \cdot P(A_1 \cap A_2 \cap \dots \cap A_n) \end{aligned}$$

e) Dacă A și B sunt evenimente din \mathcal{F} și A implică B , adică $A \subseteq B$, atunci $P(A) \leq P(B)$ și $P(B \setminus A) = P(B) - P(A)$;

d) Dacă A și B sunt evenimente din \mathcal{F} , atunci

$$P(AB) \leq P(A) \leq P(A \cup B) \leq P(A) + P(B),$$

$$P(AB) \leq P(B) \leq P(A \cup B) \leq P(A) + P(B).$$

Următoarea teoremă vine să confirme faptul că noțiunea de Probabilitate Clasică studiată la nivel de liceu este un caz particular al noțiunii de Probabilitate Axiomatică.

Teoremă (Probabilitate Clasică). Dacă spațiul de evenimente elementare Ω corespunzător unui experiment aleator \mathcal{E} constă din n rezultate posibile ce au aceleași șanse de realizare (adică sunt echiprobabile) iar A este un eveniment aleator ce conține k elemente, atunci probabilitatea evenimentului dat se calculează după formula probabilității clasice

$$P(A) = \frac{\text{card } A}{\text{card } \Omega} = \frac{k}{n}$$

este o probabilitate și în sens axiomatic.

Demonstrație. Conform condițiilor teoremei, spațiul corespunzător de probabilitate poate fi descris ca fiind tripletul (Ω, \mathcal{F}, P) , unde $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$, câmpul de evenimente fiind $\mathcal{F} = \{A \mid A \subseteq \Omega\}$. Atunci putem arăta că aplicația $P : \mathcal{F} \rightarrow \mathbb{R}$, unde $P(A)$ se calculează după formula clasică a probabilității satisface axiomelor 1-3 ale probabilității. Într-adevăr, valabilitatea Axiomei 1 este evidentă. Deoarece evenimente elementare sunt echiprobabile, adică $P\{\omega_1\} = \dots = P\{\omega_n\} = 1/n$, rezultă că

$$P(\Omega) = 1 = P\left(\bigcup_{i=1}^n \{\omega_i\}\right) = P(\omega_1) + P(\omega_2) + \dots + P(\omega_n) = n \cdot \frac{1}{n} = 1,$$

ceea ce confirmă valabilitatea Axiomei 2.

Pentru a arăta că este valabilă Axioma 3, deoarece $\text{card } \mathcal{F} = 2^n$, ne vom referi la orice multime finită de evenimente A_1, A_2, \dots, A_k din \mathcal{F} , disjuncte două câte două. Cum

$$\bigcup_{j=1}^k A_j = \left\{ \omega \in \Omega \mid \omega \in \bigcup_{j=1}^k A_j \right\} = \bigcup_{j=1}^k \{ \omega \in \Omega \mid \omega \in A_j \},$$

$$A_i \cap A_j = \emptyset, \forall i \neq j, i, j = \overline{1, k}.$$

conform Principiului Adunării (caz disjunct), avem egalitatea $\text{Card}\left(\bigcup_{j=1}^k A_j\right) =$

$\sum_{j=1}^k \text{Card}A_j$. Prin urmare

$$P\left(\bigcup_{j=1}^k A_j\right) = \frac{\text{Card}\left(\bigcup_{j=1}^k A_j\right)}{\text{Card } \Omega} = \frac{\sum_{j=1}^k \text{Card}A_j}{\text{Card}\Omega} = \sum_{j=1}^k P(A_j). \quad \square$$

Remarcă. Teorema de mai sus arată că definiția Axiomatică a Probabilității este acoperitoare pentru Definiția Clasică a Probabilității. Echiprobabilitatea evenimentelor elementare poate fi postulată, prin urmare aceasta din urmă poate fi utilizată atunci când într-o problemă de calcul a probabilităților avem de a face cu experimente de genul aruncării unei monede "perfecte" sau unui zar "perfect", de alegere "la întâmplare" a unui element dintr-o multime finită de elemente, etc. Validarea acestor presupuneri se

poate verifica doar pe cale experimentală, comparând rezultatele obținute pe cale teoretică cu cele obținute pe cale experimentală.

Exemplul 1. Considerăm aruncarea unui zar perfect de n ori. Cu ce este egală probabilitatea că fața 6 va apare cel puțin o dată?

Soluție. Spațiul de evenimente elementare poate fi descris ca fiind mulțimea de seturi ordonate $\Omega = \{(i_1, i_2, \dots, i_n) \mid i_j = \overline{1, 6}, j = \overline{1, n}\}$. Observăm că mulțimea Ω coincide cu produsul cartezian al mulțimii $\{1, 2, 3, 4, 5, 6\}$ cu ea însăși de n ori. Prin urmare $Card \Omega = 6^n$, iar din faptul că zarul e perfect conchidem toate evenimentele elementare sunt echiprobabile, având probabilitatea $\frac{1}{6^n}$. Pentru a calcula probabilitatea evenimentului $A = \{\text{la aruncarea unui zar perfect de } n \text{ ori fața } 6 \text{ va apare cel puțin o dată}\}$ observăm că este mai ușor să aplicăm proprietatea $b)$ a probabilității. Într-adevăr, evenimentul opus \bar{A} poate fi descris ca fiind $\bar{A} = \{(i_1, i_2, \dots, i_n) \mid i_j = \overline{1, 5}, j = \overline{1, n}\}$, prin urmare $Card(\bar{A}) = 5^n$. Folosind definiția clasică a probabilității, găsim că $P(\bar{A}) = \frac{5^n}{6^n} = (\frac{5}{6})^n$. Deci $P(A) = 1 - P(\bar{A}) = 1 - (\frac{5}{6})^n$.

Exemplul 2. La examenul de "Probabilități, Statistică Matematică" sunt propuse 24 de bilete de examinare din care 20 de bilete "norocoase" și 4—"nenorocoase". La examen s-au prezentat 24 de studenți, fiecare extrăgând la întâmplare, fără repetare, câte un bilet. Care student are, în ordinea extragerii, probabilitatea mai mare de a extrage un bilet "norocos", primul, al doilea,..., sau ultimul?

Soluție. Presupunem că biletele "norocoase" sunt numerotate de la 1 până la 20 iar cele "nenorocoase" de la 21 până la 24. Atunci spațiul de evenimente elementare Ω și evenimentele $A_k = \{\text{studentul cu numărul de ordine } k \text{ va extrage un bilet "norocos"}\}$ se descriu, respectiv, ca fiind

$$\Omega = \{(i_1, i_2, \dots, i_{24}) \mid i_1 \neq i_2 \neq \dots \neq i_{24}, i_j = \overline{1, 24}, j = \overline{1, 24}\},$$

$$A_k = \{(i_1, i_2, \dots, i_k, \dots, i_{24}) \in \Omega \mid i_k = \overline{1, 20}\}, k = \overline{1, n}.$$

Observăm că Ω conține toate permutările posibile ale numerelor de la 1 la 24, prin urmare $Card \Omega = 24!$. La fel, A_k conține toate permutările din Ω , dar care au pe locul k , $k = \overline{1, 24}$ un număr "norocos" de bilet de examinare. Cu alte cuvinte ne putem imagina că evenimentele elementare din Ω care favorizează evenimentul aleator A_k pot fi obținute drept rezultat al unei acțiuni realizate în 2 etape succesive, astfel încât la prima etapă alegem și fixăm pe locul k un număr de bilet de la 1 până la 20, iar la etapa a doua punem pe locurile rămase orice rezultat al unei permutări din celelalte 23 numere de bilete rămase. Prima etapă poate fi realizată în 20 de modalități,

iar cea de a doua în $23!$ modalități. În concluzie, din Principiul Înmulțirii, deducem că numărul de evenimente elementare care favorizează extragerea unui bilet "norocos" pentru studentul nr. k , $k = \overline{1, 24}$, este egal cu $Card(A_k) = 20 \times 23!$. Or, indiferent de numărul k ,

$$P(A_k) = \frac{20 \times 23!}{24!} = \frac{20}{24} = \frac{5}{6}.$$

Morala: nu contează în ce ordine te prezinți la examen, conteaza...bagaajul de cunoștințe.

Exemplul 3. *Selecție aleatoare.* În statistica, atunci când cercetarea vizează o populație statistică, privită ca o multime de elemente omogene în raport cu o anumită caracteristică/proprietate, studiul se bazează, de fapt, pe cercetarea unei submulțimi (ordonate) de n elemente a populației statistice, numite *eșanion de volum n* . Pentru ca rezultatele cercetării făcute asupra eșanionului să redea, cu aproximarea dorită, fidel, proprietățile aceleiași caracteristici în întreaga populație, se cere ca eșanionul în cauză să fie *reprezentativ*. Aceasta înseamnă, de fapt, că *selecția este aleatoare*, adică fiecare element inclus în populația statistică este ales la întâmplare, cu alte cuvinte, are aceleași șanse de a fi selectat și fiecare eșanion, adică orice submulțime de elemente ordonate în ordinea selectării lor, are aceeași probabilitate de a fi extrasă din populația statistică. În presupunerea ca populația statistică Ω consta din N elemente diferite vom arata că fiecare eșanion de volum n are aceeași probabilitate de a fi selectat indiferent de tipul selecției *fără repetare* sau *cu repetare*. Într-adevăr:

Cazul selecției aleatoare fără repetare. Aceasta înseamnă ca elementul este extras la întâmplare și nu este returnat în populația $\Omega = \{\omega_1, \omega_2, \dots, \omega_N\}$, fiecare eșanion reprezentând o submulțime de forma $(\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_n})$: $\omega_{i_1} \neq \omega_{i_2} \neq \dots \neq \omega_{i_n}$, $\omega_{i_k} \in \Omega$, $i_k = \overline{1, N}$, $k = \overline{1, n}$, $n \leq N$. Este aplicabilă definiția clasică a probabilității pentru care spațiul de evenimente elementare coincide cu mulțimea $\mathcal{A}_N^n = \{(\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_n}) : \omega_{i_1} \neq \omega_{i_2} \neq \dots \neq \omega_{i_n}, \omega_{i_k} \in \Omega, i_k = \overline{1, N}, k = \overline{1, n}\}$. Cum

$$Card \mathcal{A}_N^n = A_N^n = N \cdot (N - 1) \cdot \dots \cdot (N - n + 1),$$

rezultă că fiecare rezultat posibil în cadrul unei atare eșanionări are aceeași probabilitate egală cu $1/A_N^n$. Dealtfel, se poate arăta, folosind metodele analizei combinatorii, ca pentru acest tip de selecție aleatoare orice element al populației poate nimeri în eșanion cu una și aceeași probabilitate egală

cu n/N (*Demonstrați!*), ceea ce corespunde criteriului de reprezentativitate a eșantionului.

Cazul selecției aleatoare cu repetare. Aceasta înseamnă ca elementul este extras la întâmplare și după ce este înregistrat, acesta este returnat în populația $\Omega = \{\omega_1, \omega_2, \dots, \omega_N\}$, fiecare eșantion reprezentând o submulțime de forma $(\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_n}) : \omega_{i_k} \in \Omega, i_k = \overline{1, N}, k = \overline{1, n}$. $n \geq 1$ Spațiul de evenimente elementare coincide cu multimea $\Omega \times \Omega \times \dots \times \Omega = \Omega^n = \{(\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_n}) : \omega_{i_k} \in \Omega, i_k = \overline{1, N}, k = \overline{1, n}\}$. Conform Principiului Înmulțirii Card $|\Omega^n| = N^n$, prin urmare fiecare rezultat posibil în cadrul unei atare eșantionări are aceeași probabilitate egală cu $1/N^n$. Cât privește probabilitatea de a nimeri în eșantion pentru orice element al populației, reeșind din caracterul selecției, aceasta este egală cu $1/n$ (*Demonstrați!*).

Exemplul 4. Presupunem că o familie posedă două autoturisme, starea lor fiind de așa natură încât probabilitatea că primul autoturism va fi funcțional, este egală cu 0.8, iar pentru al doilea, aceeași probabilitate, este egală cu 0.4. În plus, probabilitatea că ambele autoturisme vor fi, concomitent, funcționale, este egală cu 0.3.

a) Descrieți evenimentele aleatoare despre care este vorba în problemă și asociați-le probabilitățile corespunzătoare;

b) Cu ce este egală probabilitatea că va fi funcțional cel puțin un autoturism?

c) Dar probabilitatea ca niciun autoturism nu va fi funcțional?

Soluție. a) Conform condițiilor descrise în problemă putem introduce evenimentele $A = \{\text{primul autoturism va fi funcțional}\}$, $B = \{\text{al doilea autoturism va fi funcțional}\}$. Atunci evenimentul $A \cap B = \{\text{ambele autoturisme vor fi, concomitent, funcționale}\}$. Probabilitățile corespunzătoare vor fi: $P(A) = 0.8$, $P(B) = 0.4$, $P(A \cap B) = 0.3$;

b) Aici ne interesează probabilitatea evenimentului $A \cup B = \{\text{cel puțin unul din autoturisme va fi funcțional}\}$. Aplicând Formula Adunării Probabilităților găsim că

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.8 + 0.4 - 0.3 = 0.9;$$

c) Observăm că evenimentul $\overline{A \cup B} = \{\text{niciun autoturism nu va funcționa}\}$, prin urmare probabilitatea lui

$$P(\overline{A \cup B}) = 1 - P(A \cup B) = 1 - 0.9 = 0.1.$$

Exemplul 5. Un studiu privind gradul de angajare în câmpul muncii are la baza un sondaj aleator bazat pe un eșantion în care au fost incluse

1309 de persoane. Rezultatele au fost centralizate în următorul tablou:

	<i>Neangajați în câmpul muncii</i>	<i>Angajați în câmpul muncii</i>	<i>Total pe linie</i>
<i>Bărbați</i>	206	412	618
<i>Femei</i>	386	305	691
<i>Total pe coloană</i>	592	717	1309

Folosind aceste date, aflați probabilitatea că, *alegând la întâmplare* o persoană inclusă în eșantion, aceasta va fi *sau* o femeie *sau* o persoana neangajată în câmpul muncii?

Soluție. Spațiul de evenimente elementare consta din 1309 evenimente elementare. Introducem următoarele evenimente aleatoare: $A = \{ \text{persoana aleasă la întâmplare din eșantion va fi o femeie} \}$ și $B = \{ \text{persoana aleasă la întâmplare din eșantion va fi una neangajată în câmpul muncii} \}$. Atunci, deoarece alegerea este la întâmplare, conform definiției clasice, probabilitatea evenimentului căutat coincide cu $P(A \cup B)$. Dar $P(A) = \frac{691}{1309}$, $P(B) = \frac{592}{1309}$ iar $P(A \cap B) = \frac{386}{1309}$. Prin urmare $P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{691}{1309} + \frac{592}{1309} - \frac{386}{1309} = \frac{897}{1309} = 0.68526$.

Probleme propuse.

Tema: *Proprietățile probabilității, probabilitate clasică.*

1. Intr-o grupă de studenți, din care face parte și studentul Pacală, fiecare student este *sau* de sex feminin *sau* are părul blond *sau* îndrăgește disciplina Matematica. În grupă sunt 20 de studente din care 12 au părul blond și doar una din studentele cu părul blond au îndrăgit Matematica. Numarul total de studenți/studente cu părul blond este egal cu 24, din care doar 12 îndrăgesc Matematica. Numarul total de studenți/studente care îndrăgesc Matematica este egal cu 17, din care 6 sunt studente. Cu ce este egala probabilitatea ca, alegând la întâmplare un student din aceasta grupa, acesta va chiar Pacală?

2. Intr-un microbuz cu 17 locuri, inclusiv locul șoferului, au urcat 17 persoane, din care 4 persoane posedă permis de conducere al unui vehicul de acest tip. Cu ce este egală probabilitatea ca microbuzul va putea pleca, dacă știm ca fiecare persoană ocupă, la întâmplare, unul din aceste 17 locuri?

3. O grupă de studenți enumara 35 de studenti. Dintr-acestea, 20 de studenți s-au înscris în clubul sportiv al universității, 10 studenți s-au înscris în cercul de dansatori al universității, iar 10 studenți nu s-au înscris la nicio activitate. Din aceasta grupa este ales la întâmplare un student. Calculati probabilitatea ca:

- a) acesta va fi unul inscris la ambele activitati;
- b) acesta va fi unul inscris numai in clubul sportiv ;
- c) acesta va fi unul inscris numai la cercul de dans.

4. Dintr-o sută de studenți, 28 de studenți cunosc limba engleza, 30-germana, 42-franceza, 8-engleza si germana, 10-engleza si franceza, 5-germana si franceza, 2-toate trei limbi. Este ales la intamplare un student. Cu ce este egală probabilitatea ca acesta nu cunoaste nicuna din aceste trei limbi.

5. Presupunem ca un zar "perfect" este aruncat o singura data. Calculati probabilitățile următoarelor evenimente: $A = \{\text{numarul de puncte aparute va fi egal cu } 6\}$; $B = \{\text{numarul de puncte aparute va fi multiplu lui } 3\}$; $C = \{\text{numarul de puncte va fi par si ,totodata, mai mare decat } 2\}$.

6. Considerăm aruncarea unui zar "perfect" de două ori succesiv. Calculati probabilitățile următoarelor evenimente: $A = \{\text{la ambele aruncari va apare acelasi numar de puncte}\}$; $B = \{\text{numarul de puncte aparute la prima aruncare va fi mai mare decat numarul de puncte apărute la aruncarea a doua}\}$; $C = \{\text{suma punctelor aparute la ambele aruncari va fi pară}\}$; $D = \{\text{produsul punctelor apărute la ambele aruncări va fi egală cu } 6\}$.

7. Se alege la intamplare un număr natural format din 5 cifre. Calculati probabilitățile următoarelor evenimente: $A = \{\text{numărul citit de la stanga la dreapta sau invers, va ramâne neschimbat, ca de exemplu } 13531\}$ $B = \{\text{numărul va fi multiplu lui } 5\}$; $C = \{\text{numărul va fi format numai din cifre pare}\}$.

8. Considerăm o mulțime formata din primele 10 litere ale alfabetului latin. Cate alfabete formate din 3 litere putem alcătui din aceasta mulțime de litere. Cu ce este egala probabilitatea că un alfabet de acest fel ales la întâmplare va conține litera A ?

9. Dintr-un lot de 10 calculatoare, din care 3 calculatoare sunt cu defecte, sunt alese la întâmplare 3. Calculați probabilitățile următoarelor evenimente: $A = \{\text{dintre cele } 3 \text{ calculatoare alese, cel puțin unul, va avea defecte}\}$ $B = \{\text{toate calculatoarele alese vor avea defecte}\}$; $C = \{\text{dintre cele } 3 \text{ calculatoare alese exact } 2 \text{ vor avea defecte}\}$.

10. Un cub, ale cărui fețe sunt vopsite, a fost tăiat într-o mie de cubulețe de aceeași dimensiune. Cu ce este egală probabilitatea că un cubuleț extras la întâmplare va avea exact două fețe vopsite?

11. Un grup de 8 persoane ocupa fiecare, la întâmplare, unul din cele 8 scaune așezate in jurul unei mese rotunde. Cu ce este probabilitatea că 2 persoane anume vor numeri alături.

12. Un grup de 8 persoane ocupă fiecare, la intamplare, unul din cele

8 scaune așezate într-un rand de 8 locuri. Cu ce este probabilitatea că 2 persoane anume vor numeri alături.

13. Pe 5 cartonașe sunt scrise cifrele de la 1 până la 5. Se alege la întâmplare, fără întoarcere, unul după altul, 3 cartonașe, acestea fiind puse alături de la stânga la dreapta în ordinea extragerii. Calculați probabilitățile următoarelor evenimente: $A = \{\text{va apare numărul } 123\}$, $B = \{\text{nu va apare cifra } 3\}$, $C = \{\text{va apare un număr par}\}$.

14. Numerele 1, 2, ..., 9 sunt scrise în ordine aleatoare. Calculați probabilitățile următoarelor evenimente: $A = \{\text{numerele vor apare în ordinea lor crescătoare}\}$ $B = \{\text{numerele } 1 \text{ și } 2 \text{ vor nimeri alături în ordine crescătoare}\}$; $C = \{\text{pe locuri pare vor nimeri numere pare}\}$.

15. Considerăm un alfabet format din literele a, b, c, d, m . Alegem la întâmplare, succesiv, cu întoarcere, 4 litere, scriindu-le în ordinea extragerii lor. Cu ce este egală probabilitatea că vom obține cuvântul *mama*?

16. Considerăm aruncarea unui zar "perfect" de 10 ori succesiv. Calculați probabilitățile următoarelor evenimente: $A = \{\text{la nici o aruncare nu va apare fața } 6\}$ $B = \{\text{la cel puțin o aruncare va apare fața } 6\}$; $C = \{\text{exact la } 3 \text{ aruncări va apare fața } 6\}$.

17. Într-un lift al unei case cu 7 nivele, la nivelul de jos, au urcat 6 pasageri. Știind că fiecare pasager poate ieși, la întâmplare, la oricare din cele 6 nivele, calculați probabilitățile următoarelor evenimente: $A = \{\text{toți pasagerii vor ieși la nivele diferite}\}$ $B = \{\text{toți pasagerii vor ieși la același nivel}\}$; $C = \{\text{la nivelele } 4, 5 \text{ și } 6 \text{ vor ieși câte } 2 \text{ pasageri}\}$.

18. Un copil se joacă cu 11 cartonașe pe care sunt imprimată literele $I, N, F, O, R, M, A, T, I, C, A$, aranjându-le la întâmplare unul lângă altul. Cu ce este egală probabilitatea că acesta va obține, astfel, cuvântul *INFORMATICA*.

19. Considerăm aruncarea unui zar "perfect" de 6 ori succesiv. Calculați probabilitățile următoarelor evenimente: $A = \{\text{de trei ori va apare fața } 1, \text{ de două ori fața } 3 \text{ și o dată fața } 6\}$; $B = \{\text{vor apare fețe diferite}\}$; $C = \{\text{de } 3 \text{ ori va apare aceeași față}\}$.

20. Care este probabilitatea că, jucând cu o singură variantă la LOTO-SPORT "5 din 35", nu vom fi în pierdere, adică vom câștiga ceva?

21. Într-un tren cu 3 vagoane se urcă la întâmplare 7 persoane. Care este probabilitatea că în primul vagon vor urca 4 persoane?

22. *Problema cavalerului DeMere:* De câte ori trebuie să aruncăm un zar "perfect" pentru că probabilitatea apariției feței 6, cel puțin o dată, să fie mai mare decât $1/2$?

23. O grupă este formată din 23 de studenți. Calculați probabilitățile următoarelor evenimente: $A = \{\text{toți studenții vor avea zile de naștere diferite}\}$; $B = \{\text{se vor găsi, cel puțin doi studenți care au aceeași zi de naștere}\}$. **Nota.** Excludem cazul când în grupă sunt studenți gemeni.

24. Să se arate că probabilitatea de a obține în urma aruncării a 4 zaruri, cel puțin o singură dată fața 1, este mai mare decât probabilitatea de a obține după 24 de aruncări a unei perechi de zaruri cel puțin o singură dată două fețe 1. (Răspunsul explică **paradoxul cavalerului de Mere**, care considera aceste probabilități egale, fapt ce nu corespunde observărilor empirice).

25. $2n$ echipe de fotbal, printre care echipele Dacia și Zimbru, au fost împărțite, prin tragere la sorți, în 2 subgrupe a câte n echipe. Deduceți formulele de calcul pentru probabilitățile următoarelor evenimente: $A = \{\text{Dacia și Zimbru vor nimeri în grupe diferite}\}$, $B = \{\text{Dacia și Zimbru vor nimeri în aceeași grupă}\}$. Calculați aceste probabilități pentru $n = 20$.

26. O urnă conține m bile albe și n bile negre. Din această urnă făcându-se extracții cu întoarcere, să se determine formula de calcul pentru:

- Probabilitatea ca primele k bile extrase să fie negre.
- Probabilitatea ca prima bilă albă să apară la a k -a extracție.
- Probabilitatea ca printre primele k bile extrase vor fi i bile albe.

Calculați aceste probabilități pentru $m = 5$, $n = 4$.

27. Primul rand al unei săli de Cinema are $2n$ locuri. n bărbați și n femei ocupa la întâmplare, fiecare, câte un loc. Deduceți formulele de calcul pentru probabilitățile următoarelor evenimente: $A = \{\text{niciun barbat nu va nimeri alături de barbat}\}$, $B = \{\text{toti bărbații vor nimeri alături}\}$. Calculați aceste probabilități pentru $n = 10$.

28. La un turneu de tenis s-au înscris 40 de sportivi. Prin tragere la sorți aceștia au fost împărțiți în 4 subgrupe a câte 10 sportivi. Cu ce este egală probabilitatea ca 4 din cei mai puternici tenismeni vor nimeri în grupe diferite.

29. O firmă producătoare de calculatoare acceptă achiziționarea procesoarelor pentru calculatoare de la o alta firmă numai dacă în urma verificării a 5% din procesoare alese la întâmplare dintr-un lot propus spre a fi cumpărat, niciunul nu va avea defecte. Presupunem că într-un lot de 1000 de procesoare, propuse spre achiziție, se afla cinci procesoare defecte. Cu ce este egală probabilitatea că în urma controlului, lotul va fi acceptat spre a fi cumpărat?

1.5. Probabilități clasice, discrete și geometrice drept cazuri particulare ale definiției axiomatice a probabilității

După cum am văzut, din Teorema demonstrată în paragraful anterior, probabilitatea clasică devine un caz particular al Definiției Axiomatice. Folosind limbajul acesteia din urma putem formula o definiție matematică mai strictă pentru prima și anume

Definiția probabilității clasice. Vom spune ca avem de a face cu un câmp de probabilitate clasică (Ω, \mathcal{F}, P) dacă

a) Spațiul de evenimente elementare Ω conține un număr finit de evenimente elementare;

b) Familia de evenimente aleatoare \mathcal{F} este reprezentată de toate submulțimile posibile ale lui Ω ;

c) probabilitatea este o aplicație P definită pe \mathcal{F} cu valori în mulțimea numerelor reale calculate conform formulei:

$$P(A) = \frac{\text{card } A}{\text{card } \Omega}, \quad \forall A \in \mathcal{F}.$$

Remarca 1. Definiția formulată astfel atrage după sine, în mod automat, faptul ca toate evenimentele elementare sunt echiprobabile. Într-adevar, pentru orice eveniment elementar $\omega \in \Omega$ găsim că

$$P\{\omega\} = \frac{\text{card } \{\omega\}}{\text{card } \Omega} = \frac{1}{\text{card } \Omega}.$$

Însă aplicabilitatea probabilității clasice este limitată de condiția a), dar și de faptul că, chiar dacă această condiție este valabilă, nu toate evenimentele elementare sunt echiprobabile. Urmatoarele exemple confirmă această afirmație.

Exemplul 1. (Spațiul de evenimente elementare este finit, dar evenimentele elementare nu sunt echiprobabile). Considerăm aruncarea unui zar cu centrul de greutate deformat astfel încât probabilitățile apariției fețelor lui se raportează ca 1 : 2 : 3 : 4 : 5 : 6. Definiția clasică nu este aplicabilă, fapt ce se poate verifica experimental cu orice zar imperfect.

Exemplul 2. (Spațiul de evenimente elementare este infinit, dar numărabil). Considerăm jocul de noroc descris în Exemplul 8 din p.1.3, cu singura precizare ca moneda aruncată are probabilitatea apariției stemei egală cu p , $0 < p < 1$. Chiar dacă moneda era perfectă (vezi exemplul invocat), probabilitatea clasică nu este aplicabilă din cauza neîndeplinirii condiției a) din definiție.

Exemplul 3. (Evenimentele elementare sunt echiprobabile, dar spațiul de evenimente elementare este infinit nenumerabil). Considerăm experimentul

imaginar ce constă în aruncarea la întâmplare a unui punct pe segmentul $[0,1]$. Sintagma "la întâmplare" ne sugerează că toate evenimentele elementare din spațiului de evenimente elementare $\Omega = [0,1]$ sunt echiprobabile, dar probabilitatea clasică nu poate fi aplicată din același motiv ca și în exemplul anterior.

Totuși, cum arată câmpul de probabilitate în fiecare din aceste exemple? Deoarece spațiile de evenimente elementare din exemplele 1 și 2 sunt mulțimi finite sau cel mult numărabile la ele se poate aplica

Definiția probabilității discrete. Vom spune ca avem de a face cu o probabilitate discretă P dacă

a) Spațiul de evenimente elementare Ω reprezintă o mulțime finită sau infinită, cel mult, numerabilă, adică $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ sau $\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\}$;

b) Câmpul de evenimente aleatoare \mathcal{F} este reprezentat de toate submulțimile posibile ale lui Ω ;

c) P este o aplicație definită pe \mathcal{F} cu valori în mulțimea numerelor reale calculate conform formulei:

$P(A) = \text{suma probabilităților pentru fiecare eveniment elementar ce favorizează evenimentul } A = \sum_{n: \omega_n \in A} P\{\omega_n\}$, unde P verifică următoarele 2 axiome:

A1. $P\{\omega_i\} \geq 0$, pentru orice $i \geq 1$;

A2. $P(\Omega) = 1$;

$P(A)$, reprezentând un număr, se numește probabilitatea evenimentului A , iar tripletul (Ω, \mathcal{F}, P) - câmp de probabilitate discretă.

Remarca 2. Se poate arăta că orice câmp de probabilitate axiomatică aplicat la cazul discret este, de fapt un câmp de probabilitate discretă și viceversa, probabilitatea discretă fiind, întucâtva, mai ușor de manipulat în calcule.

Exemplul 1 (Continuare). Spațiul de evenimente elementare fiind $\Omega = \{1, 2, \dots, 6\}$, câmpul de evenimente \mathcal{F} este reprezentat de toate submulțimile posibile ale lui Ω . Aceasta înseamnă valabilitatea condițiilor a) și b). Pe deoparte, din faptul că $P\{1\} : P\{2\} : \dots : P\{6\} = 1 : 2 : 3 : 4 : 5 : 6$ rezultă că $P\{2\} = 2P\{1\}$, $P\{3\} = 3P\{1\}$, ..., $P\{6\} = 6P\{1\}$. Pe de alta parte $1 = P(\Omega) = P(\{1\} \cup \{2\} \cup \dots \cup \{6\}) = P\{1\} + P\{2\} + \dots + P\{6\}$. Prin urmare $P\{1\} + P\{2\} + \dots + P\{6\} = P\{1\} + 2P\{1\} + \dots + 6P\{1\} = 21P\{1\} = 1$, adică $P\{1\} = \frac{1}{21}$, $P\{2\} = \frac{2}{21}$, ..., $P\{6\} = \frac{6}{21}$. Axiomele 1,2 ale probabilității discrete fiind întrunite, rezultă, de exemplu ca $P(A) =$

$P\{\text{la o singură aruncare a zarului va apare fața pară}\} = P\{2, 4, 6\} = \frac{2+4+6}{21} = \frac{12}{21}$.
Vedem că, spre deosebire de cazul zarului simetric, $P(A) > \frac{1}{2}$.

Exemplul 2 (*Continuare*). Deoarece experimentul constă în aruncarea unei monede până la prima apariție a stemei, putem considera rezultat elementar numărul total de aruncări efectuate, adică $\Omega = \{1, 2, \dots, n, \dots\}$. Atunci câmpul de evenimente \mathcal{F} va fi reprezentat de toate submultimile posibile ale lui Ω . Nu prezintă greutate să se verifice experimental cu orice monedă ca probabilitățile $P\{k\} = P\{\text{până la prima apariție a stemei vor fi efectuate } k \text{ aruncări}\} = p(1-p)^{k-1}$, $k = 1, 2, \dots$, deîndată ce probabilitatea apariției stemei la o singură aruncarea monedei se știe ca este egală cu p , $0 < p < 1$. Cum $P\{k\} \geq 0$, iar $P(\Omega) = P\{1, 2, \dots, n, \dots\} = P\{1\} + P\{2\} + \dots + P\{n\} + \dots = p + p(1-p)^1 + \dots + p(1-p)^{n-1} + \dots = p \cdot \frac{1}{1-(1-p)} = 1$, rezultă ca putem aplica Definiția probabilității discrete. Astfel, dacă în jocul de noroc descris în exemplul 8, p.1.3, moneda nu este perfectă, adică are probabilitatea apariției stemei egală cu p , $0 < p < 1$, $p \neq \frac{1}{2}$, atunci, să zicem, probabilitatea $P(A) = P\{\text{jocul va fi câștigat de Ion}\} = P\{1\} + P\{3\} + \dots + P\{2k-1\} + \dots = p + p(1-p)^2 + \dots + p(1-p)^{2k-2} + \dots = p \cdot \frac{1}{1-(1-p)^2}$. De exemplu, pentru $p = 1/3$, avem că $P(A) = \frac{1}{3} \cdot \frac{1}{1-(1-\frac{1}{3})^2} = \frac{3}{5}$, adică probabilitatea $P(A)$ e mai mică decât atunci când moneda era simetrică, dar oricum mai mare decât probabilitatea că jocul va fi câștigat de către rivalul lui Ion, Petru.

Cât privește experimentele aleatoare similare celui din Exemplul 3, acestea pot fi modelate matematic apelând la

Definiția probabilității geometrice. Vom spune ca avem de a face cu o probabilitate geometrică P dacă

a) Spațiul de evenimente elementare Ω reprezintă o multime infinită nenumărabilă din \mathbb{R}^n pentru care $\text{mes}\Omega < +\infty$, unde mes reprezintă lungimea în \mathbb{R}^1 , aria în \mathbb{R}^2 sau volumul în \mathbb{R}^n pentru $n \geq 3$;

b) Câmpul de evenimente aleatoare \mathcal{F} este reprezentat de toate submultimile măsurabile A ale lui Ω , adică pentru care $\text{mes}A$ poate fi definită;

c) P este o aplicație definită pe \mathcal{F} cu valori în mulțimea numerelor reale calculate conform formulei:

$$P(A) = \frac{\text{mes}A}{\text{mes}\Omega}.$$

Tripletul (Ω, \mathcal{F}, P) poartă denumirea, în acest caz, de *spațiu de probabilitate geometrică*.

Astfel, în exemplul invocat, aplicând definiția probabilității geometrice aflăm că probabilitatea că un punct aruncat la intamplare pe $[0, 1]$ va nimeri

în punctul x este egală cu $P\{x\} = \text{mes}\{x\}/\text{mes}([0, 1]) = 0/1 = 0$, pentru orice x din $[0, 1]$. Dacă ne interesează, de exemplu, probabilitatea ca un punct aruncat la întâmplare pe $[0, 1]$ va nimeri în prima jumătate a acestui interval este egală cu $P([0, 0.5]) = \text{mes}([0, 0.5])/\text{mes}([0, 1]) = 0.5/1 = 0.5$. Dealtfel, observăm că $P([0, 0.5]) = P([0, 0.5]) = P([0.5, 1])$. În genere, probabilitatea ca un punct aruncat la întâmplare pe $[0, 1]$ va nimeri într-un interval (a, b) din $[0, 1]$ coincide cu lungimea acestui interval.

Remarca 3. Exemplul analizat arată ca proprietatea $c)$ a probabilității (vezi Proprietățile Probabilității din p.1.4.), conform careia, dacă evenimentul $A = \emptyset$, $P(A) = 0$, arată că reciproca acestei proprietăți nu are loc, adică există exemple de evenimente aleatoare $A \neq \emptyset$, dar care au $P(A) = 0$. Cu alte cuvinte, afirmația ca probabilitatea unui eveniment este egală cu zero nu atrage după sine afirmația că acest eveniment este imposibil. Un alt exemplu, probabilitatea ca un fir de cablu electric, situat între doi stalpi, se va rupe, în timpul unei furtuni, exact la mijloc este egală cu 0, dar evenimentul în cauză nu este imposibil. Aceste exemple fac deosebirea dintre notiunea teoretică (ideală) a probabilității și cea empirică, cum ar fi probabilitatea frecvențială.

Probleme propuse.

Tema: *Probabilități discrete.*

1. Considerăm aruncarea o singură dată a unui tetraedru regulat, ale cărui fețe sunt numerotate cu numerele de la 1 până la 4, iar centrul său de greutate este deplasat astfel încât probabilitățile apariției fiecărei fețe se raportează ca $P\{1\}:P\{2\}: \dots :P\{4\} = 1:2: \dots :4$. Calculați probabilitățile următoarelor evenimente: $A_k = \{va\ apare\ fața\ k\}$, $k = \overline{1, 4}$; $B = \{va\ apare\ o\ fata\ pară\}$; $C = \{va\ apare\ o\ față\ numerotată\ cu\ un\ număr\ prim\}$.

2. Considerăm aruncarea o singură dată a unui zar al cărui centru de greutate este deplasat astfel încât, probabilitățile apariției fiecărei dintre fețele $k = \overline{1, 5}$ coincid între ele, iar probabilitatea apariției feței 6 coincide cu suma probabilităților anterioare. Aflați probabilitățile apariției pentru fiecare față în parte, dar și probabilitățile evenimentelor $B = \{va\ apare\ un\ număr\ par\ de\ puncte\}$; $C = \{va\ apare\ un\ număr\ prim\ de\ puncte\}$.

3. Considerăm aruncarea o singură dată a unui tetraedru regulat, ale cărui fețe sunt numerotate cu numerele de la 1 până la 4, iar centrul său de greutate este deplasat astfel încât probabilitățile $P\{k\}$ ale apariției fiecărei fețe k , $k = \overline{1, 4}$, sunt legate între ele astfel: $P\{1\} : P\{2\} : P\{3\} = 1 : 2 : 3$, iar $P\{4\} = P\{1\} + P\{2\} + P\{3\}$. Calculați probabilitățile $P\{k\}$, $k = \overline{1, 4}$,

dar si probabilitatile următoarelor evenimente: $B = \{va\ apare\ o\ fa\c{t}\a\ par\acute{a}\}$; $C = \{va\ apare\ o\ fa\c{t}\a\ numerotat\acute{a}\ cu\ un\ num\acute{a}r\ prim\}$.

4. Presupunem c\c{a} alegem la \int\i\nt\amplare c\c{a}te o liter\c{a} din cuvintele *mama* si *vama*. Descrieti spa\c{t}iul de evenimente elementare si calcula\c{t}i probabilitatea c\c{a} literele extrase vor fi acelea\c{s}i.

5. Doi juc\c{a}tori, *Ion* si *Petru*, practica urm\c{a}torul joc de noroc: primul arunca moneda *Ion*; dac\c{a} apare "stema", acesta este declarat castigator; dac\c{a} nu, arunca *Petru*; dac\c{a} apare "stema", acesta este declarat castigator; dac\c{a} nu, din nou arunca moneda *Ion*; etc., etc., jocul se termina atunci cand unul din jucatori inregistreaza , primul, aparitia stemei. Pentru fiecare jucator aparte, aflati probabilitatea ca acesta va castiga jocul, stiind ca moneda este deformata astfel, incat "stema" apare cu probabilitatea $p, 0 < p < 1$? Exista oare vre-o valoare a lui $p, 0 < p < 1$ astfel incat Ion si Petru sa aib\c{a} \c{s}anse egale de castigare a jocului?

Indicatie: Sa se considere c\c{a} probabilitatea c\c{a} jocul se va termina la aruncarea $k, k = 1, 2, \dots$, este egala cu $p(1 - p)^{k-1}$.

6. Trei juc\c{a}tori, *Ion*, *Petru* si *Mihai*, practic\c{a} urm\c{a}torul joc de noroc: primul arunc\c{a} moneda *Ion*; dac\c{a} apare "stema", acesta este declarat castigator; dac\c{a} nu, arunc\c{a} *Petru*; dac\c{a} apare "stema", acesta este declarat castigator; dac\c{a} nu, arunc\c{a} moneda *Mihai*; dac\c{a} apare "stema", acesta este declarat castigator; dac\c{a} nu, atunci din nou arunc\c{a} moneda *Ion*; etc., etc., jocul se termin\c{a} atunci, cand unul din juc\c{a}tori inregistreaz\c{a} , primul, apari\c{t}ia stemei. Pentru fiecare juc\c{a}tor aparte, aflati probabilitatea ca acesta va ca\c{s}tiga jocul, , stiind ca moneda este deformata astfel, incat "stema" apare cu probabilitatea $p, 0 < p < 1$.

Indicatie: Sa se considere c\c{a} probabilitatea c\c{a} jocul se va termina la aruncarea $k, k = 1, 2, \dots$, este egala cu $p(1 - p)^{k-1}$.

Tema: *Probabilitate geometric\c{a}*.

7. *Problema \int\i\nt\al\c{n}irii.* Doua persoane si-au fixat o intalnire intre orele 12.00 \c{s}i 13.00 cu condi\c{t}ia c\c{a} primul sosit la locul \int\i\nt\al\c{n}irii \i\l a\c{s}teapt\c{a} pe al doilea cel mult 15 minute \c{s}i dac\c{a} acesta din urma nu sose\c{s}te \i\ntre timp, primul p\c{a}r\c{s}e\c{s}te locul \int\i\nt\al\c{n}irii . Consider\c{a}nd ca fiecare dintre persoane sose\c{s}te la locul \int\i\nt\al\c{n}irii in mod int\c{a}mpl\c{a}tor, sa se calculeze probabilitatea ca \int\i\nt\al\c{n}irea nu va avea loc.

8. Presupunem ca autobusele de ruta dat\c{a} circul\c{a} la intervale fixe de 30 de minute. Un poten\c{t}ial pasager sose\c{s}te la una din sta\c{t}iile acestei rute \i\nt\c{u}n moment \i\nt\c{a}mpl\c{a}tor. Cu ce este egala probabilitatea ca acesta va a\c{s}tepta, cel mult, 5 minute pana la sosirea urm\c{a}torului autobus? Dar probabilitatea

ca acesta nu va fi nevoit sa aștepte deloc?

9. Un baston de lungimea L este rupt la întâmplare în trei segmente. Cu ce este egală probabilitatea ca din aceste segmente de baston se poate construi un triunghi.

1.6. Probabilitate condiționată. Formula înmulțirii probabilităților

Fie A și B doua evenimente aleatoare legate de același câmp de probabilitate (Ω, \mathcal{F}, P) , unde $P(B) > 0$. Presupunem că acest spațiu de probabilitate modelează comportamentul probabilist al unui experiment aleator \mathcal{E} . Presupunerea că $P(B) > 0$ înseamnă că într-un număr n suficient de mare de repetări a acestui experiment frecvența relativă a evenimentului B este mai mare ca zero, adică $f_n(B) > 0$. Întrebarea firească este, în ce măsură faptul (informația) că s-a produs evenimentul B influențează probabilitatea evenimentului A ? Atunci putem vorbi despre probabilitatea evenimentului A condiționată de evenimentul B notată $P(A/B)$ sau despre frecvența relativă $f_{n(B)}(A)$ a evenimentului A în $n(B)$ probe \mathcal{E} în care s-a produs evenimentul B , unde, conform Principiului Regularității Statistice, $f_{n(B)}(A) \simeq P(A/B)$ pentru $n(B)$ suficient de mare. Dar, din același Principiu avem ca,

$$P(A/B) \simeq f_{n(B)}(A) = \frac{n(A \cap B)}{n(B)} = \frac{\frac{n(A \cap B)}{n}}{\frac{n(B)}{n}} \simeq \frac{P(A \cap B)}{P(B)}.$$

Aceasta justifică următoarea

Definiție. Se numește probabilitate a evenimentului A condiționată de evenimentul B , $P(B) > 0$, mărimea notată cu $P(A/B)$ și calculată după formula

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

Exemplul 1. O companie specializată în comercializarea calculatoarelor, în urma unei cercetări statistice, că 67% din toți clienții lor cumpără calculatoare, iar 45% preferă tablete. a) Cu ce este egală probabilitatea că un client ales la întâmplare va fi unul care va cumpără și un calculator și o tabletă? b) Dar probabilitatea că acest client va cumpăra apă o tabletă, dacă știm că acesta este unul care a cumpărat un calculator?

Soluție. Deoarece, procentual vorbind, din numărul total de 100% cumpărători, 67% din ei cumpără calculatoare, iar 45% - apa tablete, rezultă că tot

ce excede 100% din suma 67%+45%, adică 12% din ei, vor fi cei care preferă să cumpere ambele tipuri de apă minerală (vezi Principiul Adunării în caz general). Introducem evenimentele $A = \{\text{un cumpărător ales la întâmplare va fi unul care va cumpăra un calculator}\}$, $B = \{\text{un cumpărător ales la întâmplare va fi unul care va cumpăra o tabletă}\}$. Atunci, folosind formula probabilității clasice, găsim :

a) $P\{\text{un client ales la întâmplare va fi unul care va cumpăra și calculator și tabletă}\} = P(A \cap B) = 12/100 = 0.12$.

b) $P(B) = 0.45$; $P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{0.12}{0.45} = 0.26667$.

Din definiție, mai exact din formula de calcul a probabilității condiționate, rezultă *formula înmulțirii probabilităților* pentru două evenimente aleatoare:

$$P(AB) = P(A)P(B/A), \text{ dacă } P(A) > 0$$

$$\text{sau } P(AB) = P(B)P(A/B), \text{ dacă } P(B) > 0.$$

Exemplul 2. Se știe că orice întreprindere producătoare, spre exemplu, de calculatoare, supune produsele sale controlului calității înainte de a fi comercializate. Presupunem că, prin metode statistico-matematice specifice controlului calității se știe că ponderea calculatoarelor cu defecte ascunse este de 2%, iar atunci când un calculator defect este supus controlului acesta poate fi admis (din greșeală) spre comercializare cu probabilitatea 0.05. Cu ce este egală probabilitatea ca un calculator ales la întâmplare va fi unul cu defecte și admis, totodată, spre comercializare.

Soluție. Introducem evenimentele $D = \{\text{un calculator ales la întâmplare va fi unul cu defecte}\}$ și $C = \{\text{un calculator ales la întâmplare va fi admis spre comercializare}\}$. Conform condițiilor din problemă $P(D) = 0.02$, $P(C/D) = 0.05$. Deoarece $P(D) > 0$, din *formula înmulțirii probabilităților* avem că probabilitatea că un calculator ales la întâmplare va fi unul cu defecte și admis, totodată, spre comercializare coincide cu

$$P(CD) = P(D)P(C/D) = 0.02 \times 0.05 = 0.001.$$

Formula înmulțirii probabilităților pentru două evenimente aleatoare este doar un caz particular a unei formule mai generale. Are loc următoarea

Teoremă (*Formula înmulțirii probabilităților în caz general*). Dacă (Ω, \mathcal{F}, P) este un câmp de probabilitate și A_1, A_2, \dots, A_n sunt evenimente aleatoare legate de acest câmp de evenimente cu proprietatea $P(A_1 A_2 \dots A_{n-1}) > 0$, atunci

$$P(A_1 A_2 \dots A_n) = P(A_1)P(A_2 / A_1)P(A_3 / A_1 A_2) \dots P(A_n / A_1 A_2 \dots A_{n-1}).$$

Exemplul 3. Cu ce este egală probabilitatea că, jucând Poker și efectuând patru extrageri succesive (fără repetare) dintr-un butuc de cărți bine amestecate, vom extrage patru ași?

Soluție. Considerăm evenimentele $A_k = \{\text{la extragerea cu numărul de ordine } k \text{ va apare un as}\}$, $k = \overline{1, 4}$. Aplicând Teorema anterioară, găsim că probabilitatea în cauză este egală cu

$$\begin{aligned} P\left(\bigcap_{k=1}^4 A_k\right) &= P(A_1 A_2 \dots A_4) = \\ &= P(A_1)P(A_2 / A_1)P(A_3 / A_1 A_2)P(A_4 / A_1 A_2 \dots A_3) = \\ &= \left(\frac{4}{52}\right) \left(\frac{3}{51}\right) \left(\frac{2}{50}\right) \left(\frac{1}{49}\right) = \frac{1}{270\,725} = 3.693\,8 \times 10^{-6}. \end{aligned}$$

Probleme propuse.

1. Studiul statistic al vânzărilor zilnice ale unui magazin specializat în vânzarea produselor electronice audio arată că fiecare din următoarele evenimente $A = \{\text{un client va cumpara cel puțin un CD-player}\}$, $B = \{\text{un client va cumpara cel puțin un set de căști audio}\}$ și $C = \{\text{un client va cumpara cel puțin un CD-player și un set de căști audio}\}$ sunt egale, respectiv, cu $P(A) = 0.6$, $P(B) = 0.75$ și $P(C) = 0.5$. Aflați valorile următoarelor probabilități $P(A \cup B)$, $P(A/B)$, $P(B/A)$, $P(\overline{A} \cap \overline{B})$, $P(\overline{A} \cup \overline{B})$, $P(\overline{A}/\overline{B})$, $P(\overline{B}/\overline{A})$.

2. În legătură cu aruncarea unui zar "perfect" de trei ori succesiv considerăm evenimentele $A = \{\text{de fiecare dată va apare o față diferită}\}$, $B = \{\text{cel puțin o dată va apare fața 6}\}$. Calculați probabilitățile $P(A/B)$, $P(B/A)$.

3. Considerăm aruncarea unei monede "perfecte" sau până când apare "stema" sau până când "banul" apare de 3 ori succesiv. Cu ce este egală probabilitatea ca "banul" să apară de 3 ori succesiv dacă se știe că la prima aruncare a aparut "banul".

4. Considerăm câmpul de probabilitate (Ω, \mathcal{F}, P) și evenimentul aleator $B \in \mathcal{F}$ cu $P(B) > 0$. Arătați că tripletul $(B, B \cap \mathcal{F}, P_B)$ reprezintă un nou câmp de probabilitate (un nou model probabilist), unde $P_B(A) = P(A/B) = \frac{P(A \cap B)}{P(B)}$ pentru orice eveniment $A \in B \cap \mathcal{F}$.

Tema: *Formula înmulțirii probabilităților.*

5. Presupunem ca un PC consta din N blocuri, calculatorul ieșind din funcțiune deîndată ce iese din funcțiune unul din blocuri, ieșirea simultană din funcțiune a două sau mai multe blocuri fiind exclusă. Depanatorul de calculatoare verifică, luând la întâmplare, unul după altul câte un bloc, până când va depista blocul defectat. Cu ce este egală probabilitatea că depanatorul va depista blocul defectat la încercarea cu numărul de ordine k , $k = 1, 2, \dots, N$.

6. Un lot de 100 de calculatoare este supus controlului calității, selectând la întâmplare 5 calculatoare. Dacă se depistează că, cel puțin, unul din aceste calculatoare este defect, atunci întreg lotul este respins. Cu ce este egala probabilitatea ca lotul de calculatoare supus controlului va fi respins, daca se stie ca 5% de calculatoare din lot sunt cu defecte?

1.7. Independența evenimentelor aleatoare, formula lui Poisson

În situația când două evenimente aleatoare A și B , legate de același câmp de probabilitate (Ω, \mathcal{F}, P) , unde $P(B) > 0$, au proprietatea că $P(A/B) = P(A)$ este firesc să spunem că evenimentul A nu depinde de B sau pe scurt A și B sunt independente. Vom arăta că definiția de mai jos este acoperitoare și pentru toate situațiile de acest tip.

Definiția independenței (*cazul a două evenimente*). Vom spune că două evenimente aleatoare A și B , legate de același câmp de probabilitate (Ω, \mathcal{F}, P) , sunt independente dacă $P(AB) = P(A)P(B)$. În caz contrar vom spune că evenimentele A și B sunt dependente.

Într-adevăr, de aici rezultă imediat că dacă A și B sunt independente, unde $P(B) > 0$, atunci $P(A / B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$, ceea ce confirmă așteptările noastre.

Remarca 1. Definiția independenței a două evenimente aleatoare, cuprinde chiar și cazurile când, cel puțin, unul din evenimentele A și B are probabilitatea egală cu zero. De exemplu, fie $P(B) = 0$. Atunci, din Axioma 1 a probabilității și din proprietatea $P(AB) \leq P(B) = 0$, rezultă că $P(AB) = 0$. Dar aceasta înseamnă că $P(AB) = P(A)P(B)$, adică A și B sunt independente.

Mai mult, are loc următoarea

Propoziția 1. Dacă evenimente aleatoare A și B , legate de același câmp de probabilitate (Ω, \mathcal{F}, P) sunt independente, atunci independente vor fi și fiecare din perechile de evenimente: (\bar{A}, B) , (A, \bar{B}) , (\bar{A}, \bar{B}) .

Exemplul 1. Angajații Corporației Excelsior sunt distribuiți în funcție de gen și tipul angajării lor în felul următor:

	Tipul de angajare			
Sexul	Vânzări	Conducere	Producție	Total
<i>Masculin</i>	825	675	750	2250
<i>Feminin</i>	1675	825	250	2750
Total	2500	1500	1000	5000

Pentru a stimula loialitatea față de companie, compania alege la întâmplare un angajat, asigurându-i lunar o scurtă vacanță plus cheltuielile aferente.

a) Este oare evenimentul $F = \{va\ fi\ aleasă\ o\ femeie\}$ independent de evenimentul $C = \{va\ fi\ ales\ o\ persoana\ din\ conducerea\ corporației\}$?

b) Dar același eveniment F , este oare independent de evenimentul $D = \{va\ fi\ ales\ o\ persoana\ din\ sectorul\ de\ Producție\}$?

Soluție. Din tabel deducem că

$$P(F) = \frac{2750}{5000} = 0.55, P(C) = \frac{1500}{5000} = 0.3, P(D) = \frac{1000}{5000} = 0.2,$$

$$P(F \cap C) = \frac{825}{5000} = 0.165, P(F \cap D) = \frac{250}{5000} = 0.05.$$

a) Cum $P(F \cap C) = 0.165 \neq P(F)P(C) = 0.55 \times 0.3$ rezultă că evenimentele F și C sunt *independente*;

b) Dar $P(F \cap D) = 0.05 \neq P(F)P(D) = 0.55 \times 0.2 = 0.11$. În concluzie, evenimentele F și D sunt *dependente*.

Exemplul 2. La ora de Statistică Matematică s-au prezentat 20 de studenți din care 8 sunt fumători, 12 poartă ochelari iar 6 din cei care poartă ochelari sunt și fumători. Este scos la tablă la întâmplare unul din studenți. Aflați dacă evenimentele $A = \{studentul\ scos\ la\ tabla\ este\ unul\ fumător\}$ și $B = \{studentul\ scos\ la\ tabla\ este\ unul\ ochelarișt\}$ sunt independente sau dependente.

Soluție. Din condițiile problemei, folosind definiția probabilității clasice, aflăm că $P(AB) = \frac{3}{10} \neq P(A)P(B) = \frac{8}{20} \cdot \frac{12}{20}$, de unde deducem că evenimentele A și B sunt dependente. Această concluzie, trebuie, însă, tratată cu prudența, aceasta fiind valabilă doar pentru acest exemplu concret, dar în care eșantionul nu este reprezentativ, deoarece numărul de 20 de studenți nu este suficient de mare pentru a lansa o cercetare statistică mai amplă.

Aceasta în pofida faptului că pare a fi plauzibilă ipoteza, conform căreia fumatul ar influența negativ acuitatea vederii. O atare cercetare, pentru un număr n de studenți (persoane) suficient de mare, dacă ar da diferențe considerabile între frecvența relativă $f_n(B)$ și frecvența relativă condiționată $f_n(B/A) = f_n(A \cap B) / f_n(A)$, atunci ipoteza în cauză ar fi confirmată statistic.

Noțiunea de independență a evenimentelor aleatoare poate fi extinsă (generalizată) asupra mai mult de două evenimente.

Definiția independenței evenimentelor (în totalitate). Vom spune ca evenimentele aleatoare A_1, A_2, \dots, A_n , legate de același câmp de probabilitate (Ω, \mathcal{F}, P) , sunt *independente (în totalitate)* dacă $P(A_{i_1} A_{i_2} \dots A_{i_k}) = P(A_{i_1}) P(A_{i_2}) P(A_{i_3}) \dots P(A_{i_k})$ pentru orice set de indici diferiți $\{i_1, i_2, \dots, i_k\}$ din mulțimea de indici $\{1, 2, \dots, n\}$, $k = 2, 3, \dots, n$.

Exercițiu. Folosind definiția de mai sus, prin negarea ei, formulați, desfășurat, **Definiția dependenței evenimentelor (în totalitate)**.

Remarca 2. Atunci când evenimentele A_1, A_2, \dots, A_n sunt independente (în totalitate) Formula Înmulțirii Probabilităților se simplifică și are forma

$$P(A_1 A_2 \dots A_n) = P(A_1) P(A_2) P(A_3) \dots P(A_n).$$

Remarca 3. Independența evenimentelor A_1, A_2, \dots, A_n două câte două, mai exact, faptul că $P(A_{i_1} A_{i_2}) = P(A_{i_1}) P(A_{i_2})$ pentru orice set de indici diferiți $\{i_1, i_2\}$ din mulțimea de indici $\{1, 2, \dots, n\}$ nu garantează independența lor *în totalitate*. Contraexemplul de mai jos confirmă această afirmație.

Contraexemplu (Trei evenimente independente două câte două, dar nu și independente în totalitate). Considerăm aruncarea o singură dată a unui tetraedru regulat, cu centru de greutate simetric, fețele cărui sunt colorate astfel: fața 1-albastru, fața 2-galben, fața 3-roșu, fața 4-albastru, galben, roșu.

Arătăm ca evenimentele $A = \{\text{va apare culoarea albastră}\}$, $G = \{\text{va apare culoarea galbenă}\}$, $R = \{\text{va apare culoarea roșie}\}$ sunt independente 2 câte 2, dar nu și în totalitate. Într-adevăr,

$$P(A) = P(G) = P(R) = \frac{1}{2}, P(AG) = \frac{1}{4} = P(A)P(G) = \frac{1}{2} \cdot \frac{1}{2},$$

$$P(AR) = \frac{1}{4} = P(A)P(R) = \frac{1}{2} \cdot \frac{1}{2}, P(GR) = \frac{1}{4} = P(G)P(R) = \frac{1}{2} \cdot \frac{1}{2},$$

dar

$$P(AGR) = \frac{1}{4} \neq P(A)P(G)P(R) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}. \quad \square$$

Propoziția 2. (Formula lui Poisson). Dacă evenimentele A_k sunt independente (în totalitate) și probabilitățile $P(A_k) = p_k$, $k=1,2,\dots,n$ sunt cunoscute, atunci probabilitatea

$$P\{\text{se va produce cel puțin unul din evenimentele } A_k\} = P\left(\bigcup_{k=1}^n A_k\right) = 1 - [(1 - P(A_1))(1 - P(A_2))\dots(1 - P(A_n))] = 1 - [(1 - p_1)(1 - p_2)\dots(1 - p_n)].$$

Exemplul 3. Un aparat constă din trei elemente care în timpul funcționării lui se pot deteriora, independent unul de altul. Notăm prin $A_i = \{\text{elementul } i \text{ se va deteriora}\}$, $i = 1, 2, 3$. Să se calculeze probabilitatea evenimentului $A = \{\text{se va deteriora un singur element}\}$, $B = \{\text{se va deteriora, cel puțin, un element}\}$, dacă se știu probabilitățile: $p_1 = P(A_1) = 0.13$, $p_2 = P(A_2) = 0.06$, $p_3 = P(A_3) = 0.12$.

Soluție. Vom exprima evenimentul aleator A prin intermediul evenimentelor A_1, A_2 și A_3 . Evenimentul A se va produce, atunci și numai atunci când, se va deteriora primul element iar al doilea – nu și al treilea – nu, **sau** se va deteriora al doilea element, iar primul – nu și al treilea – nu, **sau** se va deteriora al treilea element, iar primul – nu și al doilea – nu. Prin urmare, conform definițiilor operațiilor asupra evenimentelor aleatoare, avem:

$$P(A) = P(A_1 \cdot \overline{A_2} \cdot \overline{A_3} \cup \overline{A_1} \cdot A_2 \cdot \overline{A_3} \cup \overline{A_1} \cdot \overline{A_2} \cdot A_3).$$

Calculăm probabilitatea evenimentului A folosind succesiv: formula de adunare a probabilităților pentru evenimente incompatibile (disjuncte) două câte două, formula înmulțirii probabilităților evenimentelor independente (în totalitate) și formula de calcul al probabilității evenimentului opus.

$$P(A) = 0.13 \times (1 - 0.06) \times (1 - 0.12) + (1 - 0.13) \times 0.06 \times (1 - 0.12) + (1 - 0.13) \times (1 - 0.06) \times 0.12 = 0.25161.$$

La calcularea valorii $P(B) = P\left(\bigcup_{k=1}^3 A_k\right)$ se aplica Formula lui Poisson. Deci, $P(B) = 1 - (1 - 0.13) \times (1 - 0.06) \times (1 - 0.12) = 0.28034$. \square

Probleme propuse.

Tema: *Independența evenimentelor aleatoare.*

1. Cum ați explica pe înțelesul unei persoane inteligente, dar care nu cunoaște Teoria probabilităților, care este deosebirea dintre independența a două evenimente aleatoare și proprietatea lor de a fi incompatibile/disjuncte?

Pot fi oare independente si incompatibile concomitent două evenimente aleatoare A și B , legate de același câmp de probabilitate (Ω, \mathcal{F}, P) ?

2. Sunt oare independente două evenimente aleatoare A și B , legate de același câmp de probabilitate (Ω, \mathcal{F}, P) dacă $P(A) > 0$, $P(B) > 0$, $AB = \emptyset$?

3. Arătați că dacă un eveniment aleator A nu depinde de el însuși, atunci probabilitatea acestuia este egala cu 0 sau cu 1.

4. Juriul unui concurs consta din 3 persoane care iau decizie corecta independent unul de altul. Prima si a doua persoana iau decizie corecta cu una si aceeasi probabilitate p , $0 < p < 1$, iar cea de a treia, pentru a lua decizie, arunca o moneda "perfecta". Decizia finala se ia cu majoritate de voturi. Cu ce este egala probabilitatea ca Juriul va lua o decizie corecta.

5. Presupunem ca un PC marca DELL produs in China este de calitate superioara cu probabilitatea 0.7, iar acelasi calculator produs in Honkong este de calitate superioara cu probabilitatea 0.8. Sunt luate la întâmplare 3 PC-uri produse in China si 4 PC-uri produse in Honkong. Cu ce este egala probabilitatea ca toate calculatoarele vor fi de calitate superioara.

6. Contraexemplu (*Trei evenimente pentru care probabilitatea produsului lor coincide cu produsul probabilităților lor, dar care nu sunt independente în totalitate*). Considerăm aruncarea unui zar "perfect" de două ori succesiv. Arătați că evenimentele $A = \{\text{numărul punctelor apărute la prima aruncare nu va întrece } 3\}$, $B = \{\text{numărul punctelor apărute la prima aruncare nu va fi mai mic de } 3 \text{ și, totodată, nu mai mare de } 5\}$ și $C = \{\text{suma punctelor apărute va fi egală cu } 9\}$. Arătați ca $P(ABC) = P(A)P(B)P(C)$, dar A, B și C nu sunt independente două câte două, cu alte cuvinte, *nu sunt independente în totalitate*.

Tema: *Formula lui Poisson.*

7. Presupunem că 3% din producția de procesoare pentru telefoanele mobile **iPhone 6s**, produse de firma asociata, au defecte. Controlului sunt supuse 20 de procesoare luate la întâmplare. Cu ce este egala probabilitatea că printre ele se va depista, cel puțin, un procesor cu defecte?

8. Care este numărul minim de numere aleatoare din multimea de numere $\{1, 2, \dots, 9\}$, care te trebuie generate pe calculator, pentru a fi siguri cu probabilitatea nu mai mica decât 0.9, ca printre ele se va întâlni, cel puțin un număr par?

9. Presupunem cunoscut faptul că într-un experiment aleator \mathcal{E} probabilitatea apariției, cel puțin o dată, a evenimentului A în patru probe independente \mathcal{E} este egală cu $1/2$. Cu ce este egală probabilitatea evenimentului

A dacă aceasta este aceeași în fiecare probă \mathcal{E} .

1.8. Formulele probabilității totale și a lui Bayes

Noțiunile introduse permit calcularea probabilităților unor evenimente mai complicate, cunoscând unele probabilități, inclusiv condiționate, mai ușor de identificat. Este vorba de următoarea

Teoremă (Formulele probabilității totale și a lui Bayes). Dacă A și $H_1, H_2, \dots, H_n, \dots$ sunt evenimente aleatoare legate de același câmp de probabilitate (Ω, \mathcal{F}, P) și satisfac condițiile :

- a) evenimentul A implică producerea a cel puțin unuia din evenimentele $H_1, H_2, \dots, H_n, \dots$, adică $A \subseteq H_1 \cup H_2 \cup \dots$;
- b) evenimentele $H_1, H_2, \dots, H_n, \dots$ sunt incompatibile două câte două, adică $H_i \cap H_j = \emptyset, \forall i \neq j, i, j \geq 1$;
- c) $P(H_i) > 0$, atunci au loc

formula probabilității totale (FPT)

$$P(A) = \sum_{k \geq 1} P(A/H_k)P(H_k)$$

și **formula lui Bayes**

$$P(H_j/A) = \frac{P(H_j)P(A/H_j)}{\sum_{k \geq 1} P(A/H_k)P(H_k)}, \text{ pentru orice } j \geq 1.$$

Remarcă. Probabilitatea $P(A)$ se numește probabilitate *apriori* deoarece aceasta este calculată înainte de a se efectua experimentul aleator corespunzător; grație condiției a) evenimentele $H_1, H_2, \dots, H_n, \dots$ se mai numesc *ipoteze*, iar probabilitățile $P(H_j/A), j \geq 1$ se numesc probabilități *aposteriori* deoarece acestea sunt probabilități ale ipotezelor, probabilități recalulate în condiția că, drept rezultat, în urma experimentului s-a produs evenimentul A .

Exemplul 1 (Problema „Monty-Hall”). Această problemă a fost popularizată în Statele Unite de o emisiune de divertisment a canalului CBS din 1963.

Concurentul este pus în fața a trei uși. În spatele a două uși este câte o capră iar în spatele unei uși este un autoturism. Scopul jocului este desigur câștigarea mașinii. Jucătorul va câștiga ce se află în spatele ușii alese.

Participantul trebuie să aleagă o ușă. Apoi, prezentatorul îi deschide o altă ușă, din celelalte două rămase, în spatele căreia este o capră. Apoi îl întreabă dacă vrea sau nu să își schimbe alegerea inițială în favoarea ușii a treia. Această întrebare generează, de fapt, problema calculării a doua probabilitati corespunzătoare respectării a două strategii de joc.

Strategia 1: Participantul *nu-și schimbă* alegerea inițială.

Strategia 2: Participantul *își schimbă* alegerea inițială.

Care dintre strategii conduce la o probabilitate mai mare de a câștiga mașina?

Soluție. Deoarece participantul alege, la început, la întâmplare una din cele trei uși, rezulta că nu contează în spatele căreia uși se află autoturismul. Așa că presupunerea că mașina se află în spatele ușii numărul 1 nu afectează rezolvarea problemei. Considerăm evenimentul $A = \{\text{participantul va câștiga autoturismul}\}$. Observăm că acest eveniment implică valabilitatea a cel puțin uneia din ipoteze $H_i = \{\text{participantul va alege inițial ușa cu numărul } i\}$, $i = 1, 2, 3$.

În aceste notații, evenimentele A , H_i , verifica condițiile aplicării FPT. Deci, $P(A) = P(A/H_1)P(H_1) + P(A/H_2)P(H_2) + P(A/H_3)P(H_3)$, unde $P(H_i) = 1/3$, $i = 1, 2, 3$, indiferent de strategia folosită. În cazul folosirii Strategiei 1 avem: $P(A/H_1) = 1$, iar $P(A/H_2) = P(A/H_3) = 0$, prin urmare $P(A) = 1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3}$. În cazul folosirii Strategiei 2 avem: $P(A/H_1) = 0$, iar $P(A/H_2) = P(A/H_3) = 1$, prin urmare $P(A) = 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = \frac{2}{3}$. În concluzie, aplicarea Strategiei 2 este mai eficientă deoarece mărește șansele câștigării autoturismului de două ori față de cazul aplicării Strategiei 1.

Exemplul 2 (*Problema lui Lewis Carroll*). Într-o cutie se află o bilă, despre culoarea căreia se știe că este albă sau neagră cu una și aceeași probabilitate. Introducem în această cutie o bilă albă, după care extragem la întâmplare o bilă.

a) Cu ce este egală probabilitatea că bila extrasă va fi de culoare albă.

b) Cu ce este egală probabilitatea ca bila inițială este de culoare albă, dacă bila extrasă este albă.

Soluție. Introducem evenimentul $A = \{\text{bila extrasă va fi albă}\}$ și ipotezele $H_1 = \{\text{bila inițial aflată în cutie va fi albă}\}$, $H_2 = \{\text{bila inițial aflată în cutie va fi neagră}\}$. Din condițiile problemei deducem ca: a) $A \subseteq H_1 \cup H_2$; b) $H_1 \cap H_2 = \emptyset$; c) $P(H_1) = P(H_2) = \frac{1}{2}$, ceea ce înseamnă că sunt valabile condițiile în care putem aplica formulele probabilității totale și a lui Bayes. Prin urmare, deoarece $P(A/H_1) = 1$, $P(A/H_2) = \frac{1}{2}$,

$$a) P(A) = P(A/H_1)P(H_1) + P(A/H_2)P(H_2) = 1 \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{4};$$

$$b) P(H_1/A) = \frac{P(H_1)P(A/H_1)}{P(A)} = \frac{\frac{1}{2} \cdot 1}{\frac{3}{4}} = \frac{2}{3}.$$

În concluzie, introducerea artificială a unei bile suplimentare de culoare albă mărește probabilitatea evenimentului A de la $1/2$ la $2/3$.

Probleme propuse

1. Să se rezolve Problema „Monty-Hall” din exemplul 1 anterior în presupunerea că alegerea Strategiei 1 sau 2 este una *randomizată* cu probabilitatea p , $0 \leq p \leq 1$. Mai exact, după ce participantul a făcut alegerea inițială, acesta aruncă o monedă deformată, astfel încât Stema apare cu probabilitatea p iar Banul cu probabilitatea $1 - p$. Dacă apare Stema, atunci participantul optează pentru Strategia 1, în caz contrar optează pentru Strategia 2. Să se arate că în cazul alegerii randomizate a strategiei probabilitatea $P(A) = (4 - p)/6$, prin urmare șansele maxime de câștigare a autoturismului corespund valorii $p = 0$, ceea ce corespunde folosirii, numai și numai, a Strategiei 2.

2. Un lot de PC-uri, din care 10% sunt cu defecte, este supus controlului calitatii. Schema controlului este de așa natura, încât defectul (dacă acesta există) este depistat cu probabilitatea 0.95, iar probabilitatea ca un calculator fără defecte va fi declarat defect este egală cu 0.03. Cu ce este egală probabilitatea ca un calculator ales la întâmplare din lot va fi declarat defect? Cu ce este egală probabilitatea ca PC-ul ales la întâmplare într-adevăr este defect dacă se știe că acest PC a fost, în urma controlului, declarat a fi defect?

3. Considerăm că la un magazin de calculatoare au fost aduse un lot de PC-uri marca HP, din care 30% sunt produse în China, 20% în Singapore și 50% în Honkong. Cu ce este egală probabilitatea ca un PC cumpărat la întâmplare are defecte ascunse dacă astfel de defecte au 20% de calculatoare produse în China, 10%-cele produse în Singapore și 5%-cele produse în Honkong? Cu ce este egală probabilitatea ca PC-ul cumpărat la întâmplare este produs în China, dacă se știe că acesta s-a dovedit a avea defecte ascunse?

4. Într-o cutie sunt 20 de mingi de tenis, din care 15 sunt noi noute, iar 5 sunt folosite la joc. Pentru primul joc sunt alese la întâmplare două mingi, după care sunt puse la loc în cutie. Pentru jocul următor sunt alese, la fel, două mingi. Cu ce este egală probabilitatea ca ambele mingi alese pentru cel de al doilea joc vor fi noi noute? Cu ce este egală probabilitatea ca pentru primul joc au fost extrase 2 mingi noi noute dacă se știe că mingiile extrase

pentru cel de al doilea joc s-au dovedit a fi noi noute?

5. Avem doua cutii, astfel incat in prima cutie se afla 6 bile albe si 4 bile negre, iar intr-a doua cutie se afla 3 bile albe si 2 bile negre. Din prima cutie este extrasa la intamplare o bila si pusa intr-a doua cutie. dupa care dintr-a doua cutie este extrasa la intamplare o bila. Cu ce este egala probabilitatea ca aceasta va fi de culoare alba? Cu ce este egala probabilitatea ca din prima cutia a fost extrasa o bila alba daca se stie ca din cutia a doua a fost extrasa o bila alba?

6. Presupunem exact una din 10 000 000 de monede perfecte are imprimata Stema pe ambele parti ale ei. Cu ce este egala probabilitatea ca a fost aleasa la întâmplare moneda cu ambele fete marcate cu Stemă daca se stie ca in urma aruncarii ei de 10 ori succesiv a aparut Stema?

7. Se stie ca mesajele scurte (SMS-urile) transmise prin intermediul telefoniei mobile sunt codificate cu ajutorul cifrelor/semnalelor 0 sau 1. Presupunem ca transmiterea semnalelor este supusa bruiajelor, astfel incat sunt deformate $2/5$ semnale 0 si $1/3$ semnale 1. Presupunem ca ponderea semnalului 0 in mesajul transmis este egala cu $5/8$ iar ponderea semnalului 1 este egala cu $3/8$. Cu ce este egala probabilitatea receptionarii corecte a primului semnal din mesaj daca se stie ca a fost receptionat: a) semnalul 0; b) semnalul 1.

8. Presupunem ca avem un lot de 5 PC-uri despre care se stie, doar, ca este echiprobabila orice ipoteza H_k despre numarul k de PC-uri defecte in acest lot, $k = 0, 1, 2, \dots, 5$. Care ipoteza are probabilitatea cea mai mare daca se stie ca, alegand la intamplare un PC, acesta s-a dovedit a fi cu defecte?

9. Sa se determine probabilitatea ca intr-un lot de 1000 de calculatoare nu exista niciunul cu defecte, daca se stie ca 100 de calculatoare din acest lot, supuse controlului, s-au dovedit a fi fara defecte, presupunand ca sunt valabile, cu una si aceeasi probabilitate, oricare din ipotezele $H_k = \{ \text{numarul de calculatoare defecte, printre cele 1000 de calculatoare din lot, este egal cu } k \}$, $k = 0, 1, 2, 3, 4, 5$.

10. Intr-o cutie sunt 7 bile albe si 3 bile negre. Sunt extrase la intamplare, fara intoarcere, doua bile, din care, cel putin, una s-a dovedit a fi de culoare neagra. Cu ce este egală probabilitatea ca cealaltă bilă extrasă este albă.

11. O reprezentanță pentru vânzări de automobile cunoaște, din experiența anterioară că 10% din cei care vizitează showroom-ul și discută cu un vânzător vor cumpăra în cele din urmă o mașină. Pentru a mări șansele de succes, reprezentanța oferă o cină gratuită cu un agent de vânzări pentru

toți oamenii care sunt de acord să asculte o prezentare completă a vânzărilor. Aceasta știind că unii vor face totul pentru o cină gratuită, chiar dacă nu intenționează să cumpere o mașină, iar unii ar prefera să nu petreacă o cină cu un vânzător de mașini. Pentru a testa eficacitatea acestui stimulent de promovare a vânzărilor proiectul este derulat pentru 6 luni, constatând că 40% dintre persoanele care au cumpărat mașini au avut o cină gratuită. În plus, 10% din persoanele care nu au cumpărat mașini au avut o cină gratuită. Conducerea este interesată în a afla dacă:

a) Oamenii care acceptă cina au o probabilitate mai mare de a cumpăra o mașină nouă?

b) Care este probabilitatea ca o persoană care nu acceptă o cină gratuită să achiziționeze o mașină?

12. Pe baza unei examinări a înregistrărilor anterioare ale soldurilor conturilor unei societăți, un auditor constată că 15% au conținut erori. Din aceste solduri în eroare, 60% au fost considerate ca fiind valori neobișnuite bazate pe istoricul activității societății. Din totalul tuturor soldurilor contului, 20% prezentau valori neobișnuite. Dacă cifra pentru un anumit sold pare neobișnuită pe această bază, cu ce este egală probabilitatea că acesta conține erori?

13. Datele statistice arată ca 5% din persoanele de sex masculin suferă de daltonism pe cand doar 0.025% din persoanele de sex feminin suferă de această tulburare de vedere cromatică. Dintr-o societate formată din același număr de bărbați și de femei a fost aleasă la întâmplare o persoană. Cu ce este egală probabilitatea că aceasta persoană este una de sex masculin daca aceasta s-a dovedit a fi una afectată de daltonism.

2. Variabile aleatoare

2.1. Introducere

În legătură cu modelarea matematica a experimentelor aleatoare (ce posedă proprietatea regularității statistice) trebuie sa ținem cont că, în realitate interes prezintă, deseori, nu rezultatele posibile (evenimentele elementare) ca atare ci niște mărimi numerice ce depind de acestea. De exemplu, într-un sondaj ce vizează cercetarea veniturilor salariale ale angajaților, sa zicem, din sfera bugetară, interes prezintă venitul salarial al unui angajat inclus în eșantion, nu persoana în cauză. Cu alte cuvinte vom avea de a face

cu o marime (variabilă) ce depinde de evenimentele elementare, doar salariul variază de la angajat la angajat. Noțiunea matematică corespunzătoare este cea de *variabilă aleatoare* care în Statistică mai este cunoscută și sub denumirea de *caracteristică sau variabilă statistică*. Dar în legătură cu fiecare eveniment elementar putem asocia una, două sau mai multe caracteristici statistice. Astfel, în exemplul nostru, fiecărui angajat din sfera bugetară îi putem asocia astfel de caracteristici ca mărimea salariului, nivelul de studii, vârsta, etc. Drept consecință, se impune introducerea noțiunilor de variabile aleatoare unidimensionale, bidimensionale (multidimensionale).

2.2. Variabilă aleatoare (unidimensională), funcția ei de distribuție (repartiție)

Pentru început vom introduce noțiunea de variabilă aleatoare unidimensională.

Definiția 1. Fie (Ω, \mathcal{F}, P) un câmp de probabilitate, atunci vom numi *variabilă aleatoare (v.a.) definită pe acest câmp* orice aplicație (regulă, funcție) $X : \Omega \longrightarrow \mathbb{R}$ care verifică condiția

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F} \text{ pentru orice } x \in \mathbb{R}. \quad (1)$$

Remarca 1. Dacă suntem în cazul discret, adică, în cazul când spațiul de evenimente elementare Ω este o mulțime finită sau, cel mult, numărabilă, atunci câmpul (familia) de evenimente aleatoare \mathcal{F} coincide cu familia tuturor submulțimilor din Ω . Prin urmare, în acest caz, putem numi variabila aleatoare orice aplicație $X : \Omega \longrightarrow \mathbb{R}$, deoarece în caz discret condiția ca $\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$ este valabilă automat pentru orice $x \in \mathbb{R}$.

Evenimentul care figurează în condiția (1) se notează, pe scurt, astfel: $\{\omega : X(\omega) \leq x\}$, sau $\{X \leq x\}$. Mărimea $X(\omega)$ se numește valoare a variabilei aleatoare X . Din condiția (1) rezultă că pentru orice $x \in \mathbb{R}$ putem determina probabilitatea evenimentului aleator $\{X \leq x\}$.

În calitate de exemple de v.a. întâlnite în practică putem lua: suma de puncte apărute la aruncarea unui zar de două ori, durata funcționării unui dispozitiv electronic, numărul de particule alfa emise de o substanță radioactivă într-o unitate de timp, cantitatea anuală de precipitații atmosferice într-o anumită regiune, numărul de apeluri telefonice înregistrate pe parcursul a 24 de ore la o stație de ajutor medical, numărul de accidente auto înregistrate pe

parcursul unui anumit interval de timp, numărul despăgubirilor sau valoarea totală a despăgubirilor platite de o Companie de Asigurări, etc., etc.

Este important să facem deosebire dintre v.a. X și mulțimea ei de valori posibile \mathcal{X} . De exemplu, dacă X reprezintă mulțimea de puncte apărute la aruncarea unui zar o singură dată, atunci $X \in \mathcal{X} = \{1, 2, \dots, 6\} \subseteq \mathbb{R}$.

Definiția matematică a noțiunii de v.a. este suficient de flexibilă. Aceasta se vede din

Proprietățile variabilei aleatoare. a) Dacă X este o variabilă aleatoare, atunci pentru orice $a, b \in \mathbb{R}$, $a < b$ sunt evenimente aleatoare și prin urmare sunt definite probabilitățile lor pentru $\{X > a\}, \{X \geq a\}, \{X = a\}, \{a < X \leq b\}, \{a < X < b\}$, etc.;

b) Fie (Ω, \mathcal{F}, P) un câmp de probabilitate, $a \in \mathbb{R}$, iar $X : \Omega \rightarrow \mathbb{R}$ și $Y : \Omega \rightarrow \mathbb{R}$ sunt variabile aleatoare. Atunci sunt variabile aleatoare și funcțiile: 1) aX ; 2) X^k , $k = 1, 2, \dots$; 3) $X - Y$; 4) $X + Y$; 5) $X \cdot Y$; 6) $1/Y$ dacă $Y(\omega) \neq 0, \forall \omega \in \Omega$; 7) X / Y dacă $Y(\omega) \neq 0, \forall \omega \in \Omega$; 8) $X \pm a$.

c) În genere, dacă avem un șir finit de v.a. definite pe unul și același câmp de probabilitate, atunci v.a. va fi și orice funcție de aceste variabile, în caz că aceasta este funcție continuă. Astfel suma de v.a., diferența lor, produsul lor, minimumul sau maximumul de aceste variabile, etc., vor fi v.a.

Drept alternativă la noțiunile de v.a. X și de câmp de probabilitate (Ω, \mathcal{F}, P) pe care este definită aceasta, primate, în ansamblu, ca model matematic ce descrie comportamentul probabilist al v.a. X , se folosește pe larg, inclusiv în Statistica Matematică, noțiunea de funcție de distribuție (repartiție) a v.a. X .

Definiția 2. Fie (Ω, \mathcal{F}, P) un câmp de probabilitate și $X : \Omega \rightarrow \mathbb{R}$ o variabilă aleatoare definită pe el. Atunci funcția $F : \mathbb{R} \rightarrow \mathbb{R}$ definită prin relația

$$F(x) = P(X \leq x), \text{ pentru orice } x \in \mathbb{R}, \quad (2)$$

se numește *funcție de distribuție (f.d.) sau de repartiție a v.a. X* , numită și *funcție cumulativă de distribuție (f.c.d.)*.

Remarca 2. Funcția cumulativă de distribuție este, de fapt, imaginea teoretică a funcției empirice din Statistica Descriptivă, mai exact a frecvențelor relative cumulate crescător.

Exemplul 1. În legătură directă cu aruncarea unei monede "perfecte" o singură dată vom descrie f.d. a numărului X de "steme" apărute și vom trasa graficul ei. Observăm că valorile posibile ale lui X sunt valori din $\mathcal{X} \in \{0, 1\}$,

unde $P(X \in \mathcal{X}) = 1$. Atunci f.d. a v.a. X

$$F(X) = P(X \leq x) = \begin{cases} P(\emptyset), & \text{pentru orice } x < 0, \\ P\{\text{stema nu va apare}\}, & \text{pentru orice } 0 \leq x < 1, \\ P\{\Omega\}, & \text{pentru orice } 1 \leq x, \end{cases}$$

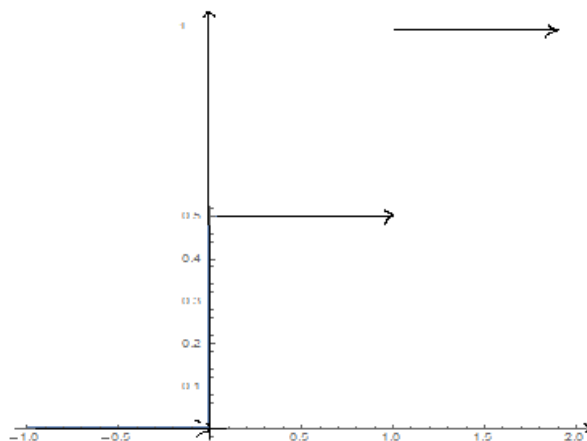
$$\begin{cases} 0, & \text{pentru orice } x < 0, \\ \frac{1}{2}, & \text{pentru orice } 0 \leq x < 1, \\ 1, & \text{pentru orice } 1 \leq x. \end{cases}$$

Folosind funcția indicator

$$I_A(x) = \begin{cases} 1, & \text{pentru } x \in A, \\ 0, & \text{pentru } x \notin A, \end{cases}$$

observăm că funcția de distribuție $F(x)$ poate fi scrisă într-o formă mai compactă: $F(x) = \frac{1}{2}I_{[0,1)}(x) + I_{[1,+\infty)}(x)$.

Graficul acestei funcții este următorul



Teoremă. (Proprietățile caracteristice ale f.d.). Dacă $F(x)$ este o f.d. a unei v.a., atunci au loc următoarele proprietăți:

1^0 . $F(x)$ este monoton nedescrescătoare, adică, $F(x_1) \leq F(x_2)$ deîndată ce $x_1 \leq x_2$;

2^0 . $F(x)$ este continuă la dreapta pentru orice $x \in \mathbb{R}$, adică, pentru orice șir monoton descrescător de valori x_n care tinde la x , atunci când n tinde

la $+\infty$, șirul corespunzător de valori $F(x_n)$ are drept limită valoarea $F(x)$, fapt ce se notează, pe scurt, $F(x+0) = F(x)$;

$$\mathbf{3}^0. F(-\infty) = 0 \text{ și } F(+\infty) = 1.$$

Remarca 3. Importanța Teoremei anterioare rezidă în faptul că proprietățile $\mathbf{1}^0$ - $\mathbf{3}^0$ sunt caracteristice *numai și numai* funcțiilor de distribuție în sensul că, are loc și reciproca acestei teoreme, conform căreia: pentru orice funcție $F : \mathbb{R} \rightarrow \mathbb{R}$, ce posedă aceste proprietăți, putem construi (neunivoc) un câmp de probabilitate (Ω, \mathcal{F}, P) și o v.a. X definită pe el, astfel încât funcția ei de distribuție va coincide cu F . În concluzie, orice funcție $F(x)$ poate fi privită ca un model matematic ce descrie o *Legitate* (distribuție, repartiție) ce guvernează comportamentul probabilist al unei v.a., deîndată ce această funcție posedă proprietățile $\mathbf{1}^0$ - $\mathbf{3}^0$. Mai mult, în postură de model probabilist, cunoscând f.d. $F(x)$, putem calcula probabilitățile cele mai importante din punct de vedere practic, probabilități legate de v.a. corespunzătoare X . Aceasta se vede din următoarea

Propoziție (Formule de calcul ale probabilităților pe baza f.d.)

Fie X o v.a. cu f.d. $F(x)$. Atunci pentru orice $a < b$, $a, b \in \mathbb{R}$, au loc următoarele formule de calcul ale probabilităților:

- a) $P(a < X \leq b) = F(b) - F(a)$;
- b) $P(X > a) = 1 - F(a)$;
- c) $P(X = a) = F(a) - F(a - 0)$, prin $F(a - 0)$ fiind notată limita la stânga a funcției F în punctul a ;
- d) $P(a \leq X \leq b) = F(b) - F(a - 0)$;
- e) $P(a < X < b) = F(b - 0) - F(a)$;
- f) $P(a \leq X < b) = F(b - 0) - F(a - 0)$.

Remarca 4. Din formula c) rezultă că, atunci când f.r. este continuă, $P(X = a) = 0$, deoarece în acest caz și $F(a - 0) = F(a)$.

Remarca 5. În continuare vom spune că *graficele care au forma celui din exemplul anterior* sunt *grafice de formă scarată*. Acestea ne permit să restabilim cu ușurință repartiția v.a., mai exact, mulțimea de valori posibile ale v.a. corespunzătoare, ca fiind punctele de salt ale graficului, iar probabilitățile corespunzătoare coincid cu înălțimile treptelor în cauză.

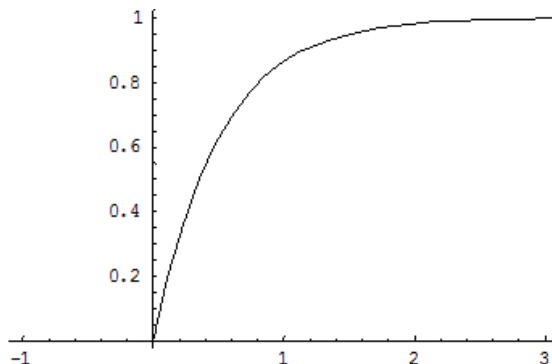
Exemplul 2. Olaf Motors, Inc. este un dealer auto într-un mic oraș sudic din SUA. Pe baza unei analize a acestuia istoriei vânzărilor sale, managerii știu că în fiecare zi numărul de mașini Toyota Prius vândute poate varia de la 0 la 5. Cum poate funcția de distribuție $F(x)$ a numărului X de mașini vândute într-o zi să fie utilizată pentru planificarea stocului zilnic dacă se

stie că

$$F(x) = 0.15I_{[0,1)}(x) + 0.45I_{[1,2)}(x) + 0.65I_{[2,3)}(x) + 0.85I_{[3,4)}(x) + \\ 0.95I_{[4,5)}(x) + I_{[5,+\infty)}(x).$$

Soluție. Variabila aleatoare $X \in \mathcal{X} = \{0, 1, 2, 3, 4, 5\}$. Deoarece $P(X \leq 4) = F(4) = 0.95$, rezultă că dacă Olaf Motors va ține zilnic în stoc 4 autoturisme Prius, atunci clienții vor fi pe deplin satisfăcuți în 95% de cazuri, pe când dacă acest dealer va ține zilnic în stoc 2 autoturisme, atunci în 35% de cazuri vor exista clienți cu solicitări de cumpărare nesatisfăcute, deoarece $P(X > 2) = 1 - F(2) = 1 - 0.65 = 0.35$.

Exemplul 3. Considerăm funcția $F(x) = (1 - e^{-2x}) \cdot I_{[0,+\infty)}(x)$. Observăm că $F(x)$ are toate proprietățile caracteristice f.d., ceea ce se vede și din graficul ei de mai jos ca este o funcție monoton crescătoare, continuă și pentru care $F(-\infty) = 0$, $F(+\infty) = 1$. Aceasta se vede și din graficul ei dat mai jos



Prin urmare putem considera ca $F(x)$ descrie comportamentul probabilist al unei v.a. X , valorile căreia sunt concentrate pe intervalul $[0, +\infty)$, deoarece:

$$P(X \leq 0) = F(0) = 1 - e^{-2 \cdot 0} = 1 - 1 = 0$$

iar

$$P(X > 0) = 1 - F(0) = 1 - (1 - e^{-2 \cdot 0}) = 1 - 0 = 1.$$

Probleme propuse.

1. Considerând funcția de distribuție $F(x)$ din exemplul anterior ca fiind funcția de distribuție a duratei vieții X (măsurată în ani) a unui corp de iluminat pe bază de neon, aflați valorile următoarelor probabilități: a) $P(X \leq 1)$; b) $P(X > 1)$; c) $P(X = 1)$; d) $P(X > 0)$. Arătați că în acest caz $P(a < X \leq b) = P(a \leq X \leq b) = P(a \leq X < b) = F(b) - F(a) = e^{-2a} - e^{-2b}$, pentru orice $a \leq b$, $a, b \in \mathbb{R}$.

2. Care din următoarele exemple reprezintă funcții de distribuție, adică sunt modele probabiliste ale unor variabile aleatoare?

a) $F(x) = (1 - p)I_{[0,1)}(x) + I_{[1,+\infty)}(x), p \in (0, 1)$;

b) $F(x) = \frac{1}{4}I_{[0,1)}(x) + \frac{1}{8}I_{[1,2)}(x) + I_{[2,+\infty)}(x)$;

c) $F(x) = \frac{1}{4}I_{[0,1)}(x) + \frac{1}{2}I_{[1,2)}(x) + I_{[2,+\infty)}(x)$;

d) $F(x) = xI_{[0,1)}(x) + I_{[1,+\infty)}(x)$; e) $F(x) = 2xI_{[0,1)}(x) + I_{[1,+\infty)}(x)$;

f) Trasați graficul fiecărei funcții.

2.3. Variabile aleatoare de tip discret, distribuții (repartiții)

În Teoria Probabilităților sunt studiate 3 tipuri de v.a.: discrete, (absolut) continue și singulare, interesante din punct de vedere al aplicațiilor fiind doar primele două. Vom începe cu variabilele aleatoare discrete (v.a.d.).

Fie X o v.a. definită pe câmpul probabilist (Ω, \mathcal{F}, P) .

Definiția 1. Variabila X se numește variabilă aleatoare de tip discret dacă mulțimea valorilor posibile ale acesteia este finită sau infinită, cel mult numerabilă, mai exact, dacă putem identifica o mulțime de forma $\mathcal{X} = \{x_1, x_2, \dots, x_n\}$ sau $\mathcal{X} = \{x_1, x_2, \dots, x_n, \dots\}$ cu proprietatea $P(X \in \mathcal{X}) = 1$, unde $x_1 < x_2 < \dots < x_n < \dots$.

Drept exemple de v.a. de tip discret putem lua numărul de steme apărute la aruncarea unei monede de n ori, numărul de puncte apărute la aruncarea unui zar o singură dată, numărul de apeluri telefonice înregistrate la Urgența Medicală pe parcursul a 24 de ore, numărul de erori descoperite în urma compilării unui soft, numărul de cumpărători, dintr-un număr total de n cumpărători, care au cumpărat un anumit produs comercial, etc., etc.

În paragraful anterior am văzut că orice v.a. X definită pe câmpul probabilist (Ω, \mathcal{F}, P) poate fi modelată matematic într-o manieră echivalentă, cu ajutorul funcției ei de distribuție $F(X) = P(X \leq x)$. După cum vom vedea în cazul v.a. de tip discret, mai există încă o alternativă, echivalentă cu f.d.

Definiția 2. Vom numi *distribuție (repartiție) probabilistă* (simplu, *dis-*

tribuție) a v.a. X setul de perechi ordonate $(x_i, p_i)_{i \geq 1}$ sau tabloul de forma

$$X : \begin{pmatrix} x_1 & x_2 & \dots & x_n & \dots \\ p_1 & p_2 & \dots & p_n & \dots \end{pmatrix}, \text{ unde } p_i = P(X = x_i) \geq 0, i \geq 1, \sum_{i \geq 1} p_i = 1.$$

Următoarea afirmație arată că, în caz discret, f.d. și distribuția v.a. X sunt două forme echivalente de modelare matematică (probabilistă) a ei.

Propoziție. *F.d. și distribuția unei v.a. X de tip discret sunt legate între ele conform următoarelor formule:*

$$a) F(x) = \sum_{i: x_i \leq x} p_i;$$

b) mulțimea de valori posibile ale v.a. X coincide cu mulțimea $\mathcal{X} = \{x_1, x_2, \dots, x_n, \dots\} = \{x \in \mathbb{R} : F(x) - F(x-0) > 0\}$ iar probabilitățile

$$p_i = P(X = x_i) = F(x_i) - F(x_i - 0), i \geq 1.$$

Remarcă. Din această propoziție rezultă că f.d. și distribuția unei v.a. de tip discret sunt două forme echivalente de modelare probabilistă ce descriu legea care guvernează comportamentul probabilist al v.a.. Mai mult, o v.a. de tip discret este definită, dacă se cunoaște legea ei de repartiție: sau sub formă de funcție de repartiție, sau sub formă de repartiție. Mai observăm că formula de legătură dintre distribuția și f.d. a v.a. X (vezi p.a) din propoziția de mai sus) mai poate fi scrisă și în forma

$$F(x) = \sum_{i \geq 1} p_i I_{[x_i, +\infty)}(x).$$

Exemplul 1. Revista *The American Almanac of Jobs and Salaries*, a comunicat, în urma unui studiu statistic pe anii 1994 – 95, că din numărul total de absolvenți licențiați la specializarea Contabilitate care și-au găsit un loc de muncă 25% s-au angajat în sectorul public. Cu ce este egală probabilitatea că alegând la întâmplare 15 astfel de absolvenți vom descoperi că cel puțin 3 dintre ei vor fi angajați din sectorul public.

Soluție. Spațiul de evenimente elementare poate fi descris ca fiind $\Omega = \{(a_1, a_2, \dots, a_{15}) : a_k \in \{0, 1\}, k = \overline{1, 15}\}$, unde

$$a_k = \begin{cases} 1, & \text{daca absolventul nr. } k \text{ va fi unul angajat din sectorul public,} \\ 0, & \text{în caz contrar.} \end{cases}$$

Câmpul de evenimente aleatoare $\mathcal{F} = \{A : A \subseteq \Omega\}$, iar pentru fiecare eveniment elementar $\{(a_1, a_2, \dots, a_{15})\}$, știind ca evenimentele aleatoare $\{a_k\}$

sunt independente si probabilitățile lor sunt egale cu

$$P\{a_k\} = (0.25)^{a_k}(1 - 0.25)^{1-a_k} = (0.25)^{a_k}(0.75)^{1-a_k}, k = \overline{1, 15},$$

putem determina probabilitatea lui

$$\begin{aligned} P\{(a_1, a_2, \dots, a_{15})\} &= P\{a_1\}P\{a_2\}\dots P\{a_{15}\} = \prod_{k=1}^{15} (0.25)^{a_k}(0.75)^{1-a_k} = \\ &= (0.25)^{\sum_{k=1}^{15} a_k} (0.75)^{15 - \sum_{k=1}^{15} a_k}. \end{aligned}$$

Deci pentru orice eveniment $A \in \mathcal{F}$ putem calcula probabilitatea lui dupa formula:

$$P(A) = \sum_{(a_1, a_2, \dots, a_{15}) \in \Omega : (a_1, a_2, \dots, a_{15}) \in A} P\{(a_1, a_2, \dots, a_{15})\}.$$

Notăm prin X numărul absolvenților angajați în sectorul public din cei 15 absolvenți angajați în câmpul muncii. Observăm ca X poate fi privită ca fiind v.a definită pe câmpul de probabilitate (Ω, \mathcal{F}, P) mai sus, unde $X \in \{0, 1, \dots, 15\}$ cu probabilitatea 1, $\{X=k\} = \{(a_1, a_2, \dots, a_{15}) \in \Omega : X(a_1, a_2, \dots, a_{15}) = k\} = \{(a_1, a_2, \dots, a_{15}) \in \Omega : a_1 + a_2 + \dots + a_{15} = k\}$, $k = \overline{1, 15}$. Prin urmare putem spune ca

$$\begin{aligned} P\{X = k\} &= \sum_{(a_1, a_2, \dots, a_{15}) \in \Omega : (a_1, a_2, \dots, a_{15}) \in \{X=k\}} P\{(a_1, a_2, \dots, a_{15})\} = \\ &= \sum_{(a_1, a_2, \dots, a_{15}) \in \Omega : a_1 + a_2 + \dots + a_{15} = k} P\{(a_1, a_2, \dots, a_{15})\} = \\ &= \sum_{(a_1, a_2, \dots, a_{15}) \in \Omega : a_1 + a_2 + \dots + a_{15} = k} (0.25)^{\sum_{k=1}^{15} a_k} (0.75)^{15 - \sum_{k=1}^{15} a_k} = \\ &= (0.25)^k (0.75)^{15-k} \cdot \sum_{(a_1, a_2, \dots, a_{15}) \in \Omega : a_1 + a_2 + \dots + a_{15} = k} 1 = \end{aligned}$$

$$\begin{aligned} &card(\{(a_1, a_2, \dots, a_{15}) \in \Omega : a_1 + a_2 + \dots + a_{15} = k\}) \cdot (0.25)^k (0.75)^{15-k} = \\ &= \mathbb{C}_n^k (0.25)^k (0.75)^{15-k}, k = \overline{1, 15}. \end{aligned}$$

V.a. X are distribuția dată de formula $P\{X = k\} = \mathbb{C}_n^k (0.25)^k (0.75)^{15-k}$, $k = \overline{1, 15}$. Atunci evenimentul $A = \{\text{cel puțin 3 din 15 absolvenți vor fi angajați din sectorul public}\} = \{X \geq 3\}$, iar $P(A) = 1 - P(\bar{A}) = 1 - P(X < 3)$. Dar,

$$P(X < 3) = P(\{X = 0\} \cup \{X = 1\} \cup \{X = 2\}) = \\ P(\{X = 0\} \cup \{X = 1\} \cup \{X = 2\}) =$$

$$P\{X = 0\} + P\{X = 1\} + P\{X = 2\} = \sum_{k=0}^3 \mathbb{C}_n^k (0.25)^k (0.75)^{15-k} = 0.236088.$$

Or, $P(X \leq 3) = 1 - 0.236088 = 0.76391$.

Modelul (distribuția) probabilistă dată de formula

$$P\{X = k\} = \mathbb{C}_n^k (0.25)^k (0.75)^{15-k}, k = \overline{1, 15}$$

reprezintă un caz particular al distribuției numită *binomială*.

Probleme propuse.

1. O unitate medicală specializată tratează 10 pacienți afectați de o infecție bacterială letală. Aceștia li se aplică un antibiotic eficiența cărui este confirmată experimental, ca fiind egală cu 95% din cazuri favorabile raportat la numărul total de cazuri. Dacă tratamentul nu este eficient, atunci pacientul decedează.

a) Identificați v.a. ce descrie numărul de pacienți care au supraviețuit infecției în urma aplicării antibioticului. Descrieți mulțimea de valori posibile a acestei variabile. Ce reprezintă câmpul de evenimente aleator corespunzător acestei v.a.

b) Prin analogie cu Exemplul 1 din p.2.3., asociați repartiția (modelul) ce descrie comportamentul probabilist al v.a. introduse.

c) Calculați probabilitatea ca vor supraviețui toți cei 10 pacienți supuși tratamentului.

d) Dar cu ce este egală probabilitatea ca vor deceda, cel mult, 2 pacienți.

e) În cazul decedării, să zicem, a 5 pacienți, Ministerul Ocrotirii Sănătății declanșează o investigație privind corectitudinea practicilor unității medicale. Ce argumente de ordin probabilist pot fi aduse în favoarea acestei investigații guvernamentale.

f) Aflați f.d. al v.a. cercetate și trasați graficul ei.

2. Magazinul "Torrent Computers" specializat în vânzarea calculatoarelor a stabilit, în baza activității sale de lungă durată, că numărul X de calculatoare vândute zilnic este guvernată de distribuția

$$X : \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 0.05 & 0.1 & 0.2 & 0.2 & 0.2 & 0.15 & 0.1 \end{pmatrix}.$$

Calculați următoarele probabilități: a) $P(3 \leq X \leq 6)$; b) $P(X > 3)$; c) $P(X \leq 4)$; d) $P(2 < X \leq 5)$.

3. V.a. X reprezintă suma punctelor aparute la aruncarea unui zar "perfect" de doua ori succesiv. Aflați distribuția probabilistă a v.a. X și calculați probabilitatea ca X va lua valori din intervalul $[-1, 4)$.

4. V.a. X reprezintă produsul punctelor aparute la aruncarea unui zar "perfect" de doua ori succesiv. Aflați distribuția probabilistă a v.a. X și calculați probabilitatea ca X va lua valori din intervalul $[-1, 0)$.

5. V.a. X reprezintă numărul minim din cele doua numere de puncte aparute la aruncarea unui zar "perfect" de doua ori succesiv. Aflați distribuția probabilistă a v.a. X și calculați probabilitatea ca X va lua valori din intervalul $[-1, 5)$.

6. V.a. X reprezintă numărul maxim din cele doua numere de puncte aparute la aruncarea unui zar "perfect" de doua ori succesiv. Aflați distribuția probabilistă a v.a. X și calculați probabilitatea ca X va lua valori din intervalul $[-1, 5)$.

7. Monetăria statului utilizează o mașină de ștanțat monede. La fiecare ștanțare se produc 10 monede. Numărul de ordine a ștanțării de la care mașina se defectează și începe sa produca monede rebutate poate fi privit ca o v.a. X a carei distribuție probabilistă are forma $P(X = k) = q(1 - p)^{k-1} I_{\{1,2,3,\dots\}}(k)$, unde $p \in (0, 1)$.

a) Exista vre-o restricție la alegerea constantei q ? Daca da, precizați care anume? **b)** Este oare v.a. X de tip discret sau de tip (absolut) continuu? De ce? **c)** Dacă se știe ca probabilitatea defectării mașinii chiar de la prima ștanțare este egală cu 0.5, concretizați cum arată distribuția probabilistă. Calculați probabilitatea că mașina se va defecta la ștanțarea a 10-a. **d)** Deduceți formula de calcul pentru valorile f.d. $F(x) = P(X \leq x)$ și calculați cu ajutorul ei probabilitatea că mașina va efectua fără defecte cel puțin 10 ștanțări. **e)** Calculați probabilitatea că mașina va efectua fără defecte cel puțin 20 ștanțări dacă se știe că aceasta a efectuat fără defecte cel puțin 10 ștanțări.

2.4. Variabile aleatoare de tip (absolut) continue, densități de distribuție (repartiție)

Pentru început vom relua Exemplitul 3 din p.1.5 în care experimentul aleator considerat rezidă în aruncarea /alegerea la întâmplare a unui punct din intervalul $[0, 1]$. Din punct de vedere matematic acest experiment se descrie, cum am văzut, de câmpul de probabilitate geometrică (Ω, \mathcal{F}, P) , unde

$\Omega = [0, 1]$, $\mathcal{F} = \{A \subseteq \Omega : \text{există } \text{mes}A\}$ iar $P(A) = \frac{\text{mes}A}{\text{mes}\Omega}$. Dacă vom nota prin X coordonata unui punct aruncat la întâmplare pe intervalul $[0, 1]$, atunci X devine v.a. definită pe câmpul de probabilitate (Ω, \mathcal{F}, P) , unde $X(\omega) = \omega$, pentru orice $\omega \in \Omega$. Mai mult ca atât, dat fiind specificul câmpului de probabilitate geometrică funcția ei de distribuție

$$F(x) = P(X \leq x) = \frac{\text{mes}(\Omega \cap (-\infty, x))}{\text{mes}\Omega} = \begin{cases} 0, & \text{dacă } x < 0, \\ x, & \text{dacă } x \in [0, 1], \\ 1, & \text{dacă } x > 1. \end{cases}$$

Observăm că f.d. $F(x)$ este nu numai continuă, dar și derivabilă aproape pentru orice $x \in \mathbb{R}$, mai exact, cu excepția punctelor $x \in \{0, 1\}$ și derivata ei

$$f(x) = \frac{dF(x)}{dx} = \begin{cases} 0, & \text{dacă } x \notin [0, 1], \\ 1, & \text{dacă } x \in [0, 1], \end{cases} \geq 0, \text{ pentru orice } x \in \mathbb{R}.$$

În plus,

$$F(x) = \int_{-\infty}^x f(u) du.$$

Faptul că v.a. X din exemplul nostru ia valori dintr-o multime *infinită nenumărabilă* (de putere continuum) cu f.d. continuă și pentru care există o funcție nenegativă $f : \mathbb{R} \rightarrow [0, +\infty)$ cu proprietatea de mai sus face ca toate v.a. cu proprietăți similare să scoată în evidență un tip aparte de v.a. marcat de următoarea

Definiție. Vom spune că v.a. X este o v.a. de tip (absolut) continuă dacă pentru funcția ei de distribuție $F(x) = P(X \leq x)$ există o funcție $f(x) \geq 0, \forall x \in \mathbb{R}$, astfel încât

$$F(x) = \int_{-\infty}^x f(u) du.$$

Funcția $f(x)$ cu aceste două proprietăți se numește *densitate de distribuție (repartiție) a probabilităților*, pe scurt, (*d.d.*).

Remarca 1. F.d., dar și d.d. analizate în exemplul de mai sus poartă denumirea de *distribuție uniformă pe* $[0, 1]$, iar experimentul (imaginar) descris mai sus poate fi simulat (imitat) pe calculator, folosind orice limbaj de programare evoluat (C++, Java, etc.) în baza funcției program *random float*, accesarea căreia conduce la generarea unei valori a unei v.a. X uniform distribuite pe acest interval. Dealtfel aceasta funcție program, împreună cu varianta ei discretă stau la baza algoritmilor de simulare (generare) a valorilor v.a. X cu distribuția probabilistă arbitrară. Cu alte cuvinte, cu ajutorul calculatorului putem simula orice fenomen aleator, deîndată ce este cunoscută legitatea probabilistă sub forma ei de f.d., distribuție sau d.d..

Drept alte exemple de v.a. de tip (absolut) continue putem lua durata vieții unui dispozitiv electronic, durata dintre două apeluri telefonice succesive înregistrate la un post telefonic, durata dintre două despăgubiri succesive acordate de o companie de asigurări, eroarea de măsurare cu ajutorul unui instrument de măsurare, înălțimea unui barbat (femeie) ales (aleasă) la întâmplare dintr-o populație a unei țări anume, etc., toate acestea fiind exemple de v.a. ale caror valori posibile formează o mulțime infinită nenumărabilă.

Remarca 2. Cuvântul "*absolut*" din Definiție nu este luat întâmplător în paranteze, deoarece în Teoria probabilităților se regăsesc exemple de v.a. ale căror f.d. este continuă, dar care nu satisfac condițiilor suplimentare impuse asupra f.d. pentru ca v.a. corespunzătoare să fie *v.a. de tip (absolut) continuă*. Este vorba de *v.a. de tip singular* care în acest curs nu vor fi abordate, datorită faptului ca aceste nu au o aplicație practică evidentă. Folosirea cuvântului "*densitate*", la fel, nu apare întâmplător în denumirea de *densitate de repartiție a probabilităților*, deoarece acesta ne amintește de sensul fizic al densității. Într-adevăr, pentru orice v.a. X de tip absolut continuă cu f.d. $F(x)$, are d.d. $f(x)$ ce posedă proprietatea că pentru o valoare $\Delta x > 0$, oricât de mică ar fi ea,

$$P(x < X \leq x + \Delta x) = F(x + \Delta x) - F(x) = \int_{-\infty}^{x+\Delta x} f(u)du - \int_{-\infty}^x f(u)du =$$

$$\int_x^{x+\Delta x} f(u)du \simeq f(x)\Delta x.$$

În plus, toate v.a. X de tip (absolut) continuu, dat fiind faptul că acestea au funcții de distribuție continue, au proprietatea că pentru orice număr real a

$$P(X = a) = F(a) - F(a - 0) = F(a) - F(a) = 0.$$

Din același motiv au loc următoarele formule de calcul ale probabilităților în baza f.d. sau a d.d.

$$\begin{aligned} P(a < X \leq b) &= P(a \leq X \leq b) = P(a < X < b) = P(a \leq X < b) = \\ &= F(b) - F(a) = \int_a^b f(u) du, \end{aligned}$$

iar

$$P(X > a) = P(X \geq a) = 1 - F(a) = \int_a^{+\infty} f(u) du.$$

Similar cu cazul discret, și în cazul (absolut) continuu d.d. $f(x)$ este o alternativă la f.d. $F(x)$ a v.a. privita ca model matematic ce descrie comportamentul probabilist al v.a. X . Concluzia aceasta se poate trage din următoarea

Propoziție (Legătura dintre d.d. și f.d., proprietățile d.d.). Dacă X este o v.a. de tip (absolut) continuu, atunci f.d. $F(x) = P(X \leq x)$ și d.d. corespunzătoare $f(x)$ au proprietățile

1°. Cunoșcând f.d. $F(x) = P(X \leq x)$ a v.a. X putem restabili d.d. corespunzătoare $f(x)$, conform relației

$$f(x) = \frac{dF(x)}{dx};$$

2°. Cunoșcând d.d. $f(x)$ a v.a. X putem restabili f.d. corespunzătoare $F(x)$, conform relației

$$F(x) = \int_{-\infty}^x f(u) du;$$

3°. Orice funcție $f(x)$ nenegativă, integrabilă Riemann pe \mathbb{R} și pentru care

$$\int_{-\infty}^{+\infty} f(x) dx = 1$$

reprezintă o d.d. probabilistă a unei v.a. X de tip (absolut) continuă.

Remarca 3. Multimea $\mathcal{X} = \{x \in \mathbb{R} : \text{d.d. } f(x) \text{ a v.a. } X \text{ are proprietatea că } f(x) > 0\}$ arată că valorile posibile ale v.a. X sunt concentrate pe \mathcal{X} . Într-adevar:

$$1 = \int_{-\infty}^{+\infty} f(x) dx = \int_{\mathcal{X}} f(x) dx + \int_{\mathbb{R} \setminus \mathcal{X}} f(x) dx = \int_{\mathcal{X}} f(x) dx + \int_{\mathbb{R} \setminus \mathcal{X}} 0 dx = \int_{\mathcal{X}} f(x) dx.$$

Exemplul 1. Multe modele legate de cercetarea unor fenomene aleatoare ce apar în domeniile tehnicii, asigurărilor, demografiei, finanțelor, etc., vizează v.a. de tip absolut continuu ale caror comportament probabilist este descris prin intermediul densității de distribuție de forma

$$f(x) = \begin{cases} 0, & \text{dacă } x < 0, \\ \lambda e^{-\lambda x}, & \text{dacă } x \geq 0. \end{cases} = \lambda e^{-\lambda x} \cdot I_{[0, +\infty)}(x)$$

unde parametrul $\lambda > 0$.

Cum $f(x) \geq 0$ și

$$\int_{-\infty}^{+\infty} f(x) dx = \int_{-\infty}^0 0 dx + \int_0^{+\infty} \lambda e^{-\lambda x} dx = (1 - e^{-\lambda x}) \Big|_0^{+\infty} = (1 - 0) - (1 - 1) = 1,$$

din proprietatea **3^o** formulată în propoziția anterioară, conchidem că $f(x)$ reprezintă un model matematic a unei v.a. Echivalentul acestui model este f.d. corespunzătoare d.d. $f(x)$, adică

$$F(x) = \int_{-\infty}^x f(u) du = \begin{cases} 0, & \text{dacă } x < 0, \\ 1 - e^{-\lambda x}, & \text{dacă } x \geq 0. \end{cases} = (1 - e^{-\lambda x}) \cdot I_{[0, +\infty)}(x).$$

Acest model poartă denumirea de *distribuție probabilistă exponențială* cu parametrul $\lambda > 0$ (a se vedea exemplele 9, din p.1.3 și 3, din p.2.2.).

Probleme propuse.

1. Care din următoarele funcții pot fi considerate ca fiind d.d. (modele probabiliste) ale unor v.a.:

a) $f(x) = e^{-0.25x} I_{[0, +\infty)}(x)$; **b)** $f(x) = 0.25e^{-0.25x} I_{[0, +\infty)}(x)$; **c)** $f(x) = x$,

pentru orice $x \in \mathbb{R}$;

d) $f(x)=2xI_{[0,1]}(x)$. În acele cazuri în care $f(x)$ este d.d. trasați graficul ei, determinații f.d. corespunzătoare și trasați graficul ei.

2. Pentru fiecare din următoarele exemple să se determine valoarea constantei k astfel încât funcția $f(x)$ să fie o d.d., după care determinați f.d. $F(x)$ corespunzătoare d.d. $f(x)$ și trasați graficele lor:

a) $f(x)=k(x-1)I_{[1,3]}(x)$; b) $f(x)=kxI_{[0,3]}(x)$; c) $f(x)=kx^3I_{[0,2]}(x)$; d) $f(x)=k(\sin x)I_{[0,\pi/2]}(x)$; e) $f(x)=k(\cos x)I_{[-\pi/2,\pi/2]}(x)$; f) $f(x)=k\sqrt{x}I_{[0,4]}(x)$.

3. Pentru fiecare din următoarele exemple de f.d. $F(x)$ să se determine d.d. $f(x)$ corespunzătoare: a) $F(x)=[1-e^{-\lambda x}(1+x)]I_{[0,+\infty)}(x)$; b)

$F(x)=\frac{1}{2\pi} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$; c) $F(x)=\frac{(x+2)}{4} I_{[-2,2]}(x)$.

4. Pentru estimarea cantității zilnice Q de combustibil, solicitate unei firme de aprovizionare cu combustibil auto, firma în cauză folosește formula $Q=100-10p+E$, unde $p \in [0,8]$. E reprezentând eroarea de estimare, aceasta este o v.a. dată de d.d. $f(x) = I_{[-20,20]}(x)$. Cantitatea Q solicitată se masoară în mii de galoane, iar prețul p este exprimat în dolari per galon.

a) Cu ce este egală probabilitatea că va fi solicitată o cantitate de combustibil mai mare decât 70000 de galoane, dacă prețul este egal cu 3\$ galonul? Dar pentru prețul de 4\$ per galon? **b)** Dacă știm costul mediu variabil de furnizare a benzinei fără plumb, că este dat de formula $C(Q)=\sqrt{Q}/2$, ce v.a. poate fi utilizată pentru a reprezenta profitul zilnic în funcție de prețul variabil p ? **c)** Cu ce este egală probabilitatea că profitul zilnic va fi unul pozitiv dacă prețul de vânzare este stabilit să fie 4\$ per galon/ Dar dacă prețul este egal cu 5\$? Dar 3\$?

2.5. Variabile aleatoare mixate discrete-continue

Multitudinea de modele probabiliste ale v.a. des întâlnite în cercetarea fenomenelor aleatoare din lumea înconjurătoare cuprinde, în afară de v.a. de tip discret și de tip (absolut) continuu, v.a. distribuțiile probabiliste ale cărora întrunesc proprietăți ale ambelor tipuri de v.a. Este vorba de variabilele aleatoare mixate discrete-continue.

Exemplu. Considerăm v.a. X ce reprezintă durata vieții, măsurată în mii de ore, a memoriei hard a unui laptop marca HP. În Exemplul 9 din p.1.3. distribuția probabilistică propusă corespundea unei v.a. aleatoare de

tip (absolut) continue cu f.d. $F(x) = (1 - e^{-10^{-6}x})I_{[0,+\infty)}(x)$. Prin urmare probabilitatea evenimentului $A = \{\text{unui laptop nou nou\c{u} \u00e2i va pica memoria hard chiar la prima lui pornire}\} = \{X = 0\}$ este egală cu zero (a se vedea Remarca 2 din paragraful anterior). În realitate, însă, aceasta probabilitate chiar dacă este mică este, totuși, nenulă. Urmatorul exemplu de f.d. a duratei vieții memoriei hard a unui laptop marca HP ia în calcul o astfel de posibilitate. Sa zicem, că

$$F(x) = 0.05 \cdot I_{[0,+\infty)}(x) + 0.95 \cdot (1 - e^{-10^{-6}x})I_{(0,+\infty)}(x).$$

Proprietățile caracteristice f.d. fiind respectate, rezulta ca funcția $F(x)$ poate fi luată în calitate de model matematic ce descrie mai adecvat comportamentul probabilist al v.a. X în cazul că serviciul cu controlul calității constată că în 5% din astfel de calculatoare hardul a picat chiar de la prima pornire. Într-adevăr, $P(X = 0) = 0.05$, iar probabilitatea că memoria hard va avea o durată de viață mai mare decât T este egală cu $P(X > T) = 1 - P(X \leq T) = 1 - [0.05 \cdot I_{[0,+\infty)}(x) + 0.95 \cdot (1 - e^{-10^{-6}x})I_{(0,+\infty)}(t)]$. Astfel, pentru $T = 0$, $P(X > 0) = 1 - 0.05 = 0.95$, iar pentru $T = 43800$ ore, adică aproximativ 5 ani, $P(X > 43800) = 1 - P(X \leq 43800)$. Dar

$$\begin{aligned} P(X \leq 43800) &= P(X = 0) + P(0 < X \leq 43800) = 0.05 + (F(43800) - F(0)) = \\ &= 0.05 + [0.05 + 0.95(1 - e^{-10^{-6} \cdot 43800}) - 0.05] = \\ &= 0.05 + 0.95 \cdot (1 - e^{-10^{-6} \cdot 43800}) = 9.0712 \times 10^{-2}. \end{aligned}$$

Prin urmare, $P(X > 43800) = 1 - 9.0712 \times 10^{-2} = 0.90929$.

Comparând probabilitatea $P(X > 43800) = 0.95715$, că durata vieții memoriei hardului va fi mai mare de 5 ani, în cazul în care nu este luat în calcul faptul că serviciul cu controlul calității constată că în 5% din astfel de calculatoare hardul a picat chiar de la prima pornire, cu aceeași probabilitate în caz ca informația în cauză este surprinsă în model, probabilitate care este egală cu 0.90929, deducem ca în ultimul caz șansele de supraviețuire a unui astfel de calculator mai mult de 5 ani s-au diminuat cu peste 4%. \square

Funcția de distribuție analizată în exemplul de mai sus se încadrează în următoarea

Definiție. Vom numi v.a. *mixată discretă-continuă* orice v.a. X dacă funcția ei de distribuție $F(x)$ are forma

$$F(x) = pF_d(x) + (1 - p)F_{ac}(x),$$

unde $0 \leq p \leq 1$, $F_d(x)$ reprezentând o f.d. de tip discret iar $F_{ac}(x)$ o f.d. de tip (absolut) continuă.

Astfel, în exemplul de mai sus avem că $p = 0.05$,

$$F_d(x) = I_{[0,+\infty)}(x) = \begin{cases} 0, & \text{dacă } x < 0, \\ 1, & \text{dacă } x \geq 0, \end{cases}$$

iar

$$F_{ac}(x) = (1 - e^{-10^{-6}x})I_{(0,+\infty)}(x) = \begin{cases} 0, & \text{dacă } x \leq 0, \\ (1 - e^{-10^{-6}x}), & \text{dacă } x > 0. \end{cases}$$

Remarcă. Apelând la forma specifică de scriere a f.d. în cazurile discret și absolut continuu, deducem ca

$$F_d(x) = \sum_{i \geq 1} p_i I_{[x_i, +\infty)}(x),$$

și corespunde unei v.a. discrete X_1 date de distribuția

$$X_1 : \begin{pmatrix} x_1 & x_2 & \dots & x_n & \dots \\ p_1 & p_2 & \dots & p_n & \dots \end{pmatrix}, \text{ unde } p_i = P(X = x_i) \geq 0, i \geq 1, \sum_{i \geq 1} p_i = 1,$$

iar

$$F_{ac}(x) = \int_{-\infty}^x f(u) du$$

și corespunde unei v.a. (absolut) continue X_2 cu d.d. $f(x)$. În acest sens, v.a. X dată de f.d.

$$F(x) = p \cdot \sum_{i \geq 1} p_i I_{[x_i, +\infty)}(x) + (1 - p) \cdot \int_{-\infty}^x f(u) du$$

prezintă un amestec (mixaj) de proprietăți ale v.a. discrete, dar și a v.a. (absolut) continue.

Probleme propuse.

1. Dacă presupunem că autobusul de ruta dată circulă la intervale de timp egale cu T_0 , $T_0 > 0$, atunci în calitate de model matematic ce descrie timpul de așteptare X a unui pasager sosirea în stație a acestui autobuz,

putem lua v.a. X de tip (absolut) continue cu d.d. $f(x) = \frac{1}{T_0} I_{[0, T_0]}(x)$. Cu alte cuvinte f.d. a timpului X de așteptare este dată de

$$F(x) = P(X \leq x) = \int_{-\infty}^x \frac{1}{T_0} I_{[0, T_0]}(t) dt,$$

adică v.a. T este *uniform distribuită* pe intervalul închis $[0, T_0]$. Conform acestui model, avem că probabilitățile $P(X = 0) = P(X = T_0) = 0$. În realitate, însă, se poate întâmpla (cei drept cu o probabilitate foarte mică) că pasagerul sa prindă autobusul chiar odată cu sosirea sa în stație atunci durata lui de așteptare $X=0$ sau, dimpotriva, ca pasagerul sa sosească chiar în momentul când autobusul a închis ușile și pleacă, atunci acesta va trebui să aștepte autobusul o durată de timp $X=T_0$. Să zicem că aceste probabilități sunt egale, respectiv cu $P(X = 0) = P(X = T_0) = 0.025$, în rest comportamentul probabilist al v.a. X , fiind descris de modelul descris mai sus.

a) Arăți că, în realitate, timpul de așteptare este o v.a. X^* dată de f.d.

$$F^*(x) = 0.025[I_{[0, +\infty)}(x) + I_{[T_0, +\infty)}(x)] + 0.95 \cdot \int_{-\infty}^x \frac{1}{T_0} I_{[0, T_0]}(t) dt;$$

b) Trasați grafic, funcțiile $F(x)$ și $F^*(x)$;

c) Calculați și comparați perechile de probabilități $P(0 < X \leq T_0)$, $P(0 < X^* \leq T_0)$ și $P(0 < X \leq T_0/2)$, $P(0 < X^* \leq T_0/2)$.

2.6. Variabilă aleatoare multidimensională (vectorială), funcția ei de distribuție, funcții de distribuție marginale

În paragrafele anterioare din acest capitol am introdus conceptul de variabilă aleatoare, ca fiind o singură aplicație/funcție cu valori numerice definită pe mulțimea evenimentelor elementare ce pot fi observate într-un experiment aleator, dar problemele practice vizează și cazuri când interes prezintă comportamentul probabilistic a două sau chiar mai multe astfel de (v.a.) aplicații/funcții legate de unul și același experiment aleator (câmp de probabilitate). De exemplu, din punctul de vedere al unei cercetări statistice, interes prezintă comportamentul probabilistic a unei perechi de variabile aleatoare (X_1, X_2) , unde X_1 reprezintă nivelul de studii iar X_2 venitul anual al unui IT-ist, angajat în câmpul muncii, ales la întâmplare din Republica

Moldova. Dealtfel, orice cercetare bazată pe un sondaj statistic, ce implica un eșantion (x_1, x_2, \dots, x_n) de observații făcute asupra unei v.a. X , pentru a aplica Statistica Matematică, se admite punctul de vedere al matematicianului, conform căruia (x_1, x_2, \dots, x_n) pot fi privite ca fiind n v.a. independente, identic distribuite ca și v.a. X și aceasta, spre deosebire de punctul de vedere al statisticianului, care a înregistrat (colectat) niște valori numerice concrete x_1, x_2, \dots, x_n . Toate aceste exemple conduc la notiunea de v.a. multidimensionale sau vectoriale.

Definiția 1. Vom numi *variabilă aleatoare n -dimensională (vectorială)* orice vector $X = (X_1, X_2, \dots, X_n)$, unde X_1, X_2, \dots, X_n sunt v.a. definite pe unul și același câmp de probabilitate (Ω, \mathcal{F}, P) .

Următoarea Teoremă ne arată că putem formula o definiție echivalentă a noțiunii de v.a. multidimensională.

Teorema 1. *Vectorul $X = (X_1, X_2, \dots, X_n)$ este o v.a. n -dimensională definită pe câmpul de probabilitate (Ω, \mathcal{F}, P) dacă și numai dacă mulțimea*

$$\{\omega \in \Omega : X_1(\omega) \leq x_1, X_2(\omega) \leq x_2, \dots, X_n(\omega) \leq x_n\} \in \mathcal{F}, \forall (x_1, x_2, \dots, x_n) \in \mathbb{R}^n.$$

Remarca 1. Aceasta teoremă garantează, de fapt, că pentru orice v.a. $X = (X_1, X_2, \dots, X_n)$ mulțimile invocate sunt evenimente aleatoare și prin urmare putem calcula probabilitățile lor, care pe scurt pot fi scrise ca fiind

$$P(X \leq x) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n), \text{ unde } x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n.$$

Mai mult, drept consecința, se poate arăta că evenimente aleatoare sunt și mulțimile, de exemplu, de forma

$$(a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2, \dots, a_n < X_n \leq b_n),$$

$$(a_1 < X_1, a_2 < X_2, \dots, a_n < X_n),$$

pentru orice $(a_1, a_2, \dots, a_n), (b_1, b_2, \dots, b_n) \in \mathbb{R}^n, a_i \leq b_i, i = \overline{1, n}$, etc. Dar anume astfel de mulțimi și prezintă interes din punct de vedere practic. Mai observăm că noțiunea de v.a. *multidimensională* generalizează noțiunea de v.a. studiată anterior și care poate fi privită ca v.a. *unidimensională*.

Ca și pentru v.a. unidimensionale, drept alternativă la noțiunile de vector aleator $X = (X_1, X_2, \dots, X_n)$ și de câmp de probabilitate (Ω, \mathcal{F}, P) pe care este definit acesta, în ansamblu, privite ca model matematic ce descrie comportamentul probabilist al v.a. X , se folosește pe larg, inclusiv în Statistica

Matematică, noțiunea de funcție de distribuție a unei v.a. multidimensionale X .

Definiția 2. Fie (Ω, \mathcal{F}, P) un câmp de probabilitate și $X = (X_1, X_2, \dots, X_n) : \Omega \longrightarrow \mathbb{R}^n$ o variabilă aleatoare n -dimensională definită pe el. Atunci funcția $F : \mathbb{R}^n \longrightarrow \mathbb{R}$ definită prin relația

$$F(x) = P(X \leq x) = F(x_1, x_2, \dots, x_n) = \\ = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n), \forall x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$$

se numește *funcție de distribuție (f.d.) a v.a. n -dimensionale X sau funcție de distribuție (în ansamblu) a v.a. X_1, X_2, \dots, X_n .*

Exemplul 1. În legătură cu experimentul aleator ce constă în aruncarea unei perechi de zaruri "perfecte" considerăm v.a. X și Y , ce coincid, respectiv, cu indicatorul evenimentului $A = \{\text{suma punctelor apărute va fi pară}\}$ și $B = \{\text{produsul punctelor apărute va fi par}\}$. Să se determine funcția de distribuție a v.a. bidimensionale (X, Y) .

Soluție. Deoarece mulțimea de valori posibile al v.a. X sau ale v.a. Y este mulțimea $\{0, 1\}$, rezultă că v.a. bidimensională (X, Y) ia valori din mulțimea $\{(0, 0), (0, 1), (1, 0), (1, 1)\}$. Folosind definiția clasică aflăm că $P(X = 0, Y = 0) = P(\bar{A} \bar{B}) = 0$, $P(X = 0, Y = 1) = P(\bar{A} B) = 1/2$, $P(X = 1, Y = 0) = P(A \bar{B}) = 1/4$, $P(X = 1, Y = 1) = P(AB) = 1/4$. F.d. $F(x, y)$ a v.a. (X, Y) este o funcție definită pentru orice $(x, y) \in \mathbb{R}^2$. Dar, ținând cont de probabilitățile anterioare, deducem că $P(X \leq x, Y \leq y) = 0$, pentru orice

$$(x, y) \in (-\infty, 0) \times (-\infty, +\infty) \cup [0, 1) \times (-\infty, 1) \cup [1, +\infty) \times (-\infty, +0),$$

$P(X \leq x, Y \leq y) = 1/2$, pentru orice $(x, y) \in [0, 1) \times [1, +\infty)$, $P(X \leq x, Y \leq y) = 1/4$, pentru orice $(x, y) \in [1, +\infty) \times [0, 1)$, $P(X \leq x, Y \leq y) = 1$, pentru orice $(x, y) \in [1, +\infty) \times [1, +\infty)$. Cu alte cuvinte, am definit f.d. $F(x, y)$ pe întreg spațiul \mathbb{R}^2 .

Teorema 2 (Proprietățile caracteristice f.d. ale v.a. multidimensionale). O funcție de n -variabile $F(x_1, x_2, \dots, x_n)$ definită pe \mathbb{R}^n cu valori în \mathbb{R} poate fi considerată f.d. a unui vector aleator n -dimensional dacă și numai dacă aceasta posedă următoarele proprietăți

$$\mathbf{1}^0. \sum_{\varepsilon_1=0}^1 \sum_{\varepsilon_2=0}^1 \dots \sum_{\varepsilon_n=0}^1 (-1)^{\varepsilon_1+\varepsilon_2+\dots+\varepsilon_n} F(\varepsilon_1 a_1 + (1-\varepsilon_1)b_1, \varepsilon_2 a_2 + (1-\varepsilon_2)b_2, \\ \dots, \varepsilon_n a_n + (1-\varepsilon_n)b_n) \geq 0 \text{ deîndată ce } (a_1, a_2, \dots, a_n), (b_1, b_2, \dots, b_n) \in \mathbb{R}^n, \\ a_1 \leq b_1, a_2 \leq b_2, \dots, a_n \leq b_n;$$

2⁰. $F(x_1, x_2, \dots, x_n)$ este continuă la dreapta pentru fiecare variabilă x_i în parte, adică,

$$\lim_{a_k \downarrow x_i} F(x_1, x_2, \dots, a_k, \dots, x_n) = F(x_1, x_2, \dots, x_i, \dots, x_n)$$

pentru orice șir monoton descrescător de valori a_k care tinde la x_i , atunci când k tinde la $+\infty$, fapt ce se notează, pe scurt,

$$F(x_1, x_2, \dots, x_i + 0, \dots, x_n) = F(x_1, x_2, \dots, x_i, \dots, x_n);$$

3⁰. $\lim_{x_i \rightarrow -\infty} F(x_1, x_2, \dots, x_n) = F(x_1, x_2, \dots, -\infty, \dots, x_n) = 0$ pentru $\forall i = \overline{1, n}$ și $F(+\infty, +\infty, \dots, +\infty) = \lim_{x_i \rightarrow +\infty, i=\overline{1, n}} F(x_1, x_2, \dots, x_n) = 1$.

Remarca 2. Ca și în cazul f.d. ale v.a. unidimensionale, pentru f.d. ale v.a. multidimensionale este valabila Remarca 3, p.2.2., ceea ce arată că orice funcție de n variabile care satisface proprietățile **1⁰ – 3⁰** de mai sus poate fi considerată un model probabilist, adică f.d. a unei v.a. multidimensionale. Mai mult, are loc următoarea

Propoziție. Dacă v.a. $X = (X_1, X_2, \dots, X_n)$ este dată de f.d. $F(x_1, x_2, \dots, x_n)$, atunci:

a)

$$P(a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2, \dots, a_n < X_n \leq b_n) =$$

$$\sum_{\varepsilon_1=0}^1 \sum_{\varepsilon_2=0}^1 \dots \sum_{\varepsilon_n=0}^1 (-1)^{\varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_n} F(\varepsilon_1 a_1 + (1 - \varepsilon_1) b_1, \varepsilon_2 a_2 + (1 - \varepsilon_2) b_2, \dots, \varepsilon_n a_n + (1 - \varepsilon_n) b_n),$$

$(a_1, a_2, \dots, a_n), \in \mathbb{R}^n, a_1 \leq b_1, a_2 \leq b_2, \dots, a_n \leq b_n;$

b) f.d. $F_i(x)$ a fiecărei v.a. X_i în parte poate fi refăcută din f.d. în ansamblu a v.a. (X_1, X_2, \dots, X_n) după formula

$$F_i(x) = F(+\infty, +\infty, \dots, x, \dots, +\infty).$$

Consecință. Dacă v.a. $X = (X_1, X_2)$ este dată de f.d. $F(x_1, x_2)$, atunci:

$$P(a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2) = F(b_1, b_2) + F(a_1, a_2) - F(a_1, b_2) - F(b_1, a_2),$$

deîndată ce $a_1 \leq b_1, a_2 \leq b_2$.

Definiția 3. Funcțiile de distribuție determinate conform p.b) din propoziția anterioară se numesc f.d. marginale ale v.a. $X_i, i = \overline{1, n}$.

Exemplul 2. Considerăm următoarea funcție de doua variabile

$$F(x, y) = \begin{cases} 0, & \text{dacă } \min(x, y) < 0, \\ \min(x, y), & \text{dacă } 0 \leq \min(x, y) < 1, \\ 1, & \text{dacă } 1 \leq \min(x, y). \end{cases}$$

Să se arate ca $F(x, y)$ este o f.d. a unei v.a. bidimensionale (X, Y) și să se afle f.d. a fiecărei v.a.

Soluție. Proprietățile $\mathbf{1}^0 - \mathbf{3}^0$ caracteristice f.d. de mai multe variabile se verifică cu ușurința, ținând cont și de faptul că $\min(x, y)$ este o funcție continuă ca funcție de două variabile x, y , deoarece $\min(x, y) = (x + y - |x - y|)/2$. Prin urmare $F(x, y)$ poate fi privita ca fiind o f.d. a unei v.a. (X, Y) . Atunci, conform punctului b) din propoziția de mai sus distribuțiile (*marginale*) a fiecărei v.a. în parte sunt egale, respectiv, cu

$$F_X(x) = F(x, +\infty) = \begin{cases} 0, & \text{dacă } x < 0, \\ x, & \text{dacă } 0 \leq x \leq 1, \\ 1, & \text{dacă } 1 < x, \end{cases}$$

$$F_Y(y) = F(+\infty, y) = \begin{cases} 0, & \text{dacă } y < 0, \\ y, & \text{dacă } 0 \leq y \leq 1, \\ 1, & \text{dacă } 1 < y. \end{cases}$$

Concluzie: ambele v.a. X și Y sunt (a se vedea remarca 1 din p.2.4.) *uniform, identic distribuite pe segmentul* $[0, 1]$.

2.7. Tipurile de variabile aleatoare multidimensionale (bidimensionale), distribuții, densități de distribuție, independența v.a.

Pentru simplitate, fără a afecta cazul general, vorbind despre variabile aleatoare vectoriale de dimensiunea n , vom considera cazul $n = 2$, adică vom considera cazul v.a. bidimensionale. Așadar, fie (X, Y) o v.a. bidimensională definită pe câmpul de probabilitate (Ω, \mathcal{F}, P) sau, ceea ce este echivalent, dată (governată) de funcția de distribuție $F(x, y)$, notat pe scurt $(X, Y) \sim (\Omega, \mathcal{F}, P)$ sau $(X, Y) \sim f.d. F(x, y)$. Ca și în cazul v.a. unidimensionale, din punct de vedere al aplicațiilor lor practice, cel mai des sunt întâlnite v.a. multidimensionale de tip discret și de tip (absolut) continue, inclusiv în varianta lor mixată discret-continue.

Dacă o v.a. este de tip discret se poate afla folosind

Definiția 1. Vom spune că v.a. (X, Y) este o v.a. de tip discret dacă fiecare din v.a. (unidimensionale) X, Y este v.a. de tip discret.

Din definiție rezultă că v.a. (X, Y) este o v.a. de tip discret dacă mulțimile de valori posibile \mathcal{X}, \mathcal{Y} ale v.a. X, Y au formele $\mathcal{X} = \{x_1, x_2, \dots, x_n\}$ sau $\mathcal{X} = \{x_1, x_2, \dots, x_n, \dots\}$ și respectiv $\mathcal{Y} = \{y_1, y_2, \dots, y_n\}$ sau $\mathcal{Y} = \{y_1, y_2, \dots, y_n, \dots\}$, cu proprietatea $P((X, Y) \in \mathcal{X} \times \mathcal{Y}) = 1$, unde $x_1 < x_2 < \dots < x_n < \dots$ iar $y_1 < y_2 < \dots < y_n < \dots$, $\mathcal{X} \times \mathcal{Y}$ fiind produsul cartezian ale mulțimilor \mathcal{X}, \mathcal{Y} .

Definiția 2. Vom numi distribuție probabilista a v.a. bidimensionale de tip discret (X, Y) orice set de forma $\{(x_i, y_j), p_{ij}\}_{i,j \geq 1}$ sau orice tabel de forma

$X \setminus Y$	y_1	y_2	\dots	y_j	\dots
x_1	p_{11}	p_{12}	\dots	p_{1j}	\dots
x_2	p_{21}	p_{22}	\dots	p_{2j}	\dots
\dots	\dots	\dots	\dots	\dots	\dots
x_i	p_{i1}	p_{i2}	\dots	p_{ij}	\dots
\dots	\dots	\dots	\dots	\dots	\dots

unde $p_{ij} = P(X = x_i, Y = y_j) \geq 0$, $\sum_{i,j \geq 1} p_{ij} = 1$.

Tabelul din exemplul analizat mai sus este, din câte vedem, un exemplu de repartiție probabilista, iar scrierea f.d. $F(x, y)$ arată că, știind distribuția v.a. (X, Y) putem scrie f.d., folosind formula

$$F(x, y) = \sum_{i,j: x_i \leq x, y_j \leq y} p_{ij}.$$

Mai mult, știind distribuția probabilista a v.a. (X, Y) , putem afla distribuția probabilistă a fiecărei dintre v.a. X, Y . Astfel,

$$\begin{aligned} P(X = x_i) &= P(\{X = x_i\} \cap \Omega) = P(\{X = x_i\} \cap \bigcup_{j \geq 1} \{Y = y_j\}) = \\ &= P(\bigcup_{j \geq 1} (\{X = x_i\} \cap \{Y = y_j\})) = P(\bigcup_{j \geq 1} \{X = x_i, Y = y_j\}) = \\ &= \sum_{j \geq 1} P(X = x_i, Y = y_j) = \sum_{j \geq 1} p_{ij} = p_{i.}, \text{ pentru orice } i \geq 1. \end{aligned}$$

Analogic, $P(Y = y_j) = \sum_{i \geq 1} p_{ij} = p_{.j}$, pentru orice $j \geq 1$. Aceste două distribuții poartă denumirea de *distribuții marginale*, prin analogie cu noțiunea de f.d. marginală.

Exemplul 1 (Continuare). În condițiile Exemplului 1 din paragraful anterior sa se determine distribuția v.a. (X, Y) , unde X, Y reprezintă indicatorii parității sumei și respectiv produsului punctelor apărute. Să afle distribuția fiecărei v.a. X, Y în parte.

Soluție. Așa cum am văzut, v.a.

$$(X, Y) \in \mathcal{X} \times \mathcal{Y} = \{0, 1\} \times \{0, 1\} = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$$

și $P((X, Y) \in \mathcal{X} \times \mathcal{Y}) = 1$. iar valorile $p_{ij} = P(X = i, Y = j)$, $i, j = 0, 1$, valori care au stat la baza aflării f.d. $F(x, y)$, reprezintă distribuția centralizată în următorul tabel

$X \setminus Y$	0	1
0	0	0.5
1	0.25	0.25

după care scrierea f.d. $F(x, y)$ a v.a. (X, Y) devine simplă

$$F(x, y) = P(X \leq x, Y \leq y) = \begin{cases} 0, & \text{dacă } (x, y) \in (-\infty, 0) \times (-\infty, +\infty) \cup \\ & \cup [0, 1) \times (-\infty, 1) \cup [1, +\infty) \times (-\infty, +\infty), \\ 0.5, & \text{dacă } (x, y) \in [0, 1) \times [1, +\infty), \\ 0.25, & \text{dacă } (x, y) \in [1, +\infty) \times [0, 1), \\ 1, & \text{dacă } (x, y) \in [1, +\infty) \times [1, +\infty). \end{cases}$$

Acest exemplu sugerează, că și în cazul v.a. multidimensionale de tip discret, ca și în cazul v.a. unidimensionale de tip discret, o alternativă la f.d., în calitate de model probabilist, poate servi noțiunea de distribuție probabilistă.

Având distribuția (în ansamblu) a v.a. X și Y scrisă mai sus și folosind formulele pentru aflarea distribuțiilor marginale, găsim că indicatorul X a parității sumei punctelor apărute la aruncarea a două zaruri perfecte are distribuția

$$X : \begin{pmatrix} 0 & 1 \\ 0.5 & 0.5 \end{pmatrix},$$

iar indicatorul Y a parității produsului punctelor apărute la aruncarea a două zaruri perfecte are distribuția următoare

$$Y : \begin{pmatrix} 0 & 1 \\ 0.25 & 0.75 \end{pmatrix},$$

fapt ce coincide și cu rezultatele calculelor directe.

Un alt experiment aleator (fie și imaginar) care conduce la identificarea unui nou tip de v.a. este descris în

Exemplul 2. Considerăm aruncarea (sau alegerea) unui punct la întâmplare în patratul de latură 1, adică în domeniul $[0, 1] \times [0, 1] \subset \mathbb{R}^2$.

Notăm prin (X, Y) coordonatele acestui punct. Date fiind condițiile experimentului, acest vector este un v.a. bidimensional definit pe câmpul de probabilitate geometrica (Ω, \mathcal{F}, P) , unde $\Omega = [0, 1] \times [0, 1]$, $\mathcal{F} = \{A \subseteq \Omega : \text{există } mesA\}$ iar $P(A) = \frac{mesA}{mes\Omega} = mesA$, pentru orice eveniment $A \in \mathcal{F}$, deoarece $mes\Omega = mes([0, 1] \times [0, 1]) = 1$. Observăm că aplicația (X, Y) este o aplicație definită pe $\Omega = [0, 1] \times [0, 1]$ cu valori în (X, Y) conform regulii

$$(X, Y) : (\omega_1, \omega_2) \longrightarrow (X(\omega_1, \omega_2), Y(\omega_1, \omega_2)) = (\omega_1, \omega_2)$$

pentru orice $(\omega_1, \omega_2) \in [0, 1] \times [0, 1]$. Evenimentul aleator

$$\begin{aligned} A = \{(\omega_1, \omega_2) \in [0, 1] \times [0, 1] : X(\omega_1, \omega_2) \leq x, Y(\omega_1, \omega_2) \leq y\} = \\ \{(\omega_1, \omega_2) \in [0, 1] \times [0, 1] : X(\omega_1, \omega_2) \leq x, Y(\omega_1, \omega_2) \leq y\} = \\ \{(\omega_1, \omega_2) \in [0, 1] \times [0, 1] : \omega_1 \leq x, \omega_2 \leq y\}, \end{aligned}$$

iar

$$mesA = mes([0, 1] \times [0, 1] \cap (-\infty, x] \times (-\infty, y]).$$

Putem, așadar, determina f.d. $F(x, y)$ a v.a. bidimensionale (X, Y) :

$$F(x, y) = P(X \leq x, Y \leq y) = \frac{mesA}{mes\Omega} = \begin{cases} 0, & \text{dacă } X < 0 \text{ sau } Y < 0, \\ x, & \text{dacă } 0 \leq x \leq 1, y > 1, \\ y, & \text{dacă } x > 1, 0 \leq y \leq 1, \\ xy, & \text{dacă } 0 \leq x \leq 1, 0 \leq y \leq 1, \\ 1, & \text{dacă } x > 1, y > 1. \end{cases}$$

Pentru noi acest exemplu este remarcabil prin faptul că funcția de distribuție obținută are proprietatea că există o funcție $f(x, y) \geq 0$, pentru orice $(x, y) \in \mathbb{R}^2$, integrabilă, astfel încât

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(x, y) dudv,$$

unde

$$f(x, y) = I_{[0,1] \times [0,1]}(x) = \begin{cases} 0, & \text{dacă } (x, y) \notin [0, 1] \times [0, 1], \\ 1, & \text{dacă } (x, y) \in [0, 1] \times [0, 1]. \end{cases}$$

Această proprietate este în consens cu

Definiția 3. Vom spune ca v.a. bidimensională (X, Y) este de tip (absolut) continue dacă funcția ei de distribuție are proprietatea că există o funcție $f(x, y) \geq 0$, pentru orice $(x, y) \in \mathbb{R}^2$, integrabilă, astfel încât

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv.$$

Funcția $f(x, y)$ cu aceste proprietăți se numește *densitate de distribuție (d.d.) a v.a. (X, Y)* .

Propoziția 1. Dacă (X, Y) este o v.a de tip (absolut) continue, atunci f.d. $F(x, y)$ și d.d. $f(x, y)$ a acestei variabile au următoarele proprietăți:

1. $f(x, y) \geq 0$, pentru orice $(x, y) \in \mathbb{R}^2$;
2. $F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv$, pentru orice $(x, y) \in \mathbb{R}^2$;
3. Cu excepția unei mulțimi A de puncte din \mathbb{R}^2 , aria (măsura) căreia $mesA = 0$, există $\frac{\partial^2 F(x, y)}{\partial x \partial y}$ și

$$\frac{\partial^2 F(x, y)}{\partial x \partial y} = f(x, y);$$

$$4. \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1;$$

$$5. P(a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2) = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x, y) dx dy, \text{ pentru orice}$$

$(a_1, b_1), (a_2, b_2) \in \mathbb{R}^2, a_1 \leq b_1, a_2 \leq b_2$;

6. d.d. $f_1(x)$ a v.a. X și d.d. $f_2(y)$ a v.a. Y pot fi aflate, știind, d.d. $f(x, y)$ a v.a. (X, Y) din formulele respective,

$$f_1(x) = \int_{-\infty}^{+\infty} f(x, y) dy, \quad f_2(y) = \int_{-\infty}^{+\infty} f(x, y) dx.$$

Definiția 4. D.d. $f_1(x), f_2(y)$ determinate prin formulele de la p.6. al Propoziției anterioare se numesc *d.d. marginale*.

Remarcă. Din pp. 2-3 a propoziției 1 rezultă ca f.d. $F(x, y)$ și d.d. $f(x, y)$ ale v.a. bidimensionale, privite ca *modele matematice ce descriu comportamentul probabilist al v.a. (X, Y)* , sunt echivalente în sens că, știind f.d.

putem restabili d.d. și viceversa. Proprietatea 2 arată că în caz (absolut) continuu f.d. $F(x, y)$ este continuă ca funcție de doua variabile. Reciproca nu este valabilă, adică se poate aduce un exemplu de f.d. $F(x, y)$ continuă ca funcție de doua variabile, dar nu și (absolut) continuă ca f.d..

Exemplul 3. Intr-adevăr, f.d.

$$F(x, y) = \begin{cases} 0, & \text{dacă } \min(x, y) < 0, \\ \min(x, y), & \text{dacă } 0 \leq \min(x, y) < 1, \\ 1, & \text{dacă } 1 \leq \min(x, y), \end{cases}$$

adusă drept exemplu în paragraful precedent, este continuă, dar nu există nicio funcție cu proprietățile 1 – 4 ale d.d. prezentate în propoziția anterioară. Dealtfel, acesta este un exemplu tipic de f.d. a unei v.a. (X, Y) de tip *singular*, acest tip de v.a. nefiind abordat în cursul nostru. Mai mult, observăm că f.d. marginale $F_1(x)$, $F_2(y)$ ce corespund acestora reprezintă f.d. ale unor v.a. uniform distribuite, mai exact

$$F_1(x) = \begin{cases} 0, & \text{dacă } x < 0, \\ x, & \text{dacă } x \in [0, 1], \\ 1, & \text{dacă } x > 1, \end{cases}, \quad F_2(y) = \begin{cases} 0, & \text{dacă } y < 0, \\ y, & \text{dacă } y \in [0, 1], \\ 1, & \text{dacă } y > 1. \end{cases}$$

Surprinzător, dar aceleași f.d. marginale corespund și f.d. $F(x, y)$ din Exemplul 2 de mai sus. Mai mult, în Exemplul 2, f.d. $F(x, y) = F_1(x) \cdot F_2(y)$, pe când în exemplul nostru f.d. $F(x, y) \neq F_1(x) \cdot F_2(y)$. Explicația vine din

Definiția 5. Vom spune că v.a. X_1, X_2, \dots, X_n definite pe unul și același câmp de probabilitate sunt *independente* dacă f.d. $F(x_1, x_2, \dots, x_n)$ a vectorului aleator (X_1, X_2, \dots, X_n) are proprietatea că

$$F(x_1, x_2, \dots, x_n) = F_1(x_1) \cdot F_2(x_2) \cdot \dots \cdot F_n(x_n),$$

unde $F_1(x_1), F_2(x_2), \dots, F_n(x_n)$ reprezintă f.d. marginale ce corespund v.a. X_1, X_2, \dots, X_n . În caz contrar vom spune că X_1, X_2, \dots, X_n sunt *dependente*.

Or, explicația promisă rezidă în faptul că v.a. X, Y ce reprezintă coordonatele unui punct aruncat la întâmplare în patratul $[0, 1] \times [0, 1]$ sunt, conform definiției 5, v.a. independente. În schimb v.a. X, Y ce au f.d. invocată în exemplul din paragraful anterior sunt dependente. De remarcat că aplicarea directă a definiției 5 este greoaie, de aceea, putem aplica variantele ei echivalente ce țin cont de specificul tipului de v.a. și ce rezultă din

Propoziția 2. *V.a. (X, Y) sunt independente atunci și numai atunci când*

a) *în caz discret: distribuția lor în ansamblu $\{P(X = x_i, Y = y_j)\}_{i, j \geq 1}$ are proprietatea că $P(X = x_i, Y = y_j) = P(X = x_i) \cdot P(Y = y_j)$ pentru orice $i, j \geq 1$;*

b) *în caz (absolut) continuu: densitatea ei de distribuție $f(x, y)$ și densitățile marginale $f_1(x), f_2(y)$ au proprietatea că $f(x, y) = f_1(x) \cdot f_2(y)$ pentru orice $x, y \in \mathbb{R}$.*

Exemplul 1. *(Continuare). Pentru a verifica dacă v.a. X, Y analizate în acest exemplu sunt sau nu independente putem aplica afirmația a) din Propoziția 2. Constatăm cu ușurință, de exemplu, ca $P(X = 0, Y = 0) = 0 \neq P(X = 0)P(Y = 0) = 0.5 \cdot 0.25$, ceea ce este suficient să afirmăm că v.a. X și Y sunt dependente.*

Exemplul 3. *Compania Ocean Fish are deschise în Republica Moldova două întreprinderi de procesare a peștelui despre care se știe că proporțiile de procesare a acestora, raportate la întreaga capacitate de procesare zilnică a companiei, reprezintă o v.a. bidimensională (X, Y) cu d.d. $f(x, y) = (x + y)I_{[0,1]}(x)I_{[0,1]}(y)$. a) Să se calculeze $P(X \leq 0.5, Y \leq 0.5)$. b) Să se afle d.d. a proporției capacității de procesare zilnică la care operează fiecare întreprindere în parte. c) Să se calculeze $P(X \leq 0.5), P(Y \leq 0.5)$. d) Sunt oare proporțiile X și Y independente?*

Soluție. a) Din definiția f.d. a v.a. de tip (absolut) continuu deducem ca

$$\begin{aligned}
 P(X \leq 0.5, Y \leq 0.5) &= F(0.5, 0.5) = \int_{-\infty}^{0.5} \int_{-\infty}^{0.5} f(x, y) dx dy = \\
 &= \int_{-\infty}^{0.5} \int_{-\infty}^{0.5} (x + y) I_{[0,1]}(x) I_{[0,1]}(y) dx dy = \\
 &= \int_0^{0.5} \left(\int_0^{0.5} (x + y) dx \right) dy = \int_0^{0.5} \left[\frac{x^2}{2} + xy \right]_0^{0.5} dy = \int_0^{0.5} \left(\frac{1}{8} + \frac{y}{2} \right) dy = \frac{2}{16} = 0.125.
 \end{aligned}$$

b) Folosind formula 6 din Propoziția 1 aflăm d.d. marginală $f_1(x)$ a v.a. X :

$$f_1(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \int_0^1 (x + y) I_{[0,1]}(x) dy = \left(x + \frac{1}{2} \right) I_{[0,1]}(x).$$

Din considerente de simetrie, observăm că d.d. marginală $f_2(y)$ a v.a. Y este dată de egalitatea $f_2(y) = (y + \frac{1}{2}) I_{[0,1]}(y)$. Prin urmare, răspunsul la p. c) este dat de egalitățile

$$P(X \leq 0.5) = \int_0^{0.5} (x + \frac{1}{2}) dx = P(Y \leq 0.5) = \int_0^{0.5} (y + \frac{1}{2}) dy = \frac{1}{8} + \frac{1}{4} = 0.375.$$

d) Apelând la Definiția 3 a independenței v.a. multidimensionale, deducem că v.a. X și Y sunt *dependente* deoarece

$$F(0.5, 0.5) = P(X \leq 0.5, Y \leq 0.5) = 0.125 \neq F_1(0.5)F_2(0.5) = P(X \leq 0.5)P(Y \leq 0.5) = 0.375^2 = 0.14063.$$

Probleme propuse.

1. Care din următoarele tabele reprezintă o distribuție probabilistă a unui vector aleator?

$$1) \begin{array}{c|ccc} X \backslash Y & -1 & 0 & 1 \\ \hline 0 & 1/8 & 0 & 1/8 \\ 1 & 1/4 & 1/8 & 1/4 \end{array}, \quad 2) \begin{array}{c|ccc} X \backslash Y & -1 & 0 & 1 \\ \hline 0 & 1/8 & 0 & 1/4 \\ 1 & 1/4 & 1/8 & 1/4 \end{array}, \quad 3) \begin{array}{c|ccc} X \backslash Y & -1 & 0 & 1 \\ \hline 0 & 1/8 & 0 & 1/8 \\ 1 & 1/4 & 1/2 & 1/4 \end{array}.$$

În caz că avem de a face cu o distribuție probabilistă aflați: a) $P(X < 1, -1 \leq Y < 1)$; b) distribuția fiecărei v.a. în parte; c) dacă v.a. X, Y sunt sau nu independente.

2. O pepinieră funcționează în baza unei echipe de 7 angajați din care 3 raspund de sectorul vânzări iar ceilalți 4 sunt responsabili de sectorul grădinarit. Evident, numărul, relativ mic, de angajați poate crea dificultăți legate de absenteism. Numărul de persoane absente în ziua dată din sectorul vânzări, dar și a acelor absente din sectorul grădinarit reprezintă un vector aleator (X, Y) distribuit conform cu tabelul următor:

$X \backslash Y$	0	1	2	3	4
0	0.75	0.025	0.01	0.01	0.03
1	0.06	0.03	0.01	0.01	0.003
2	0.025	0.01	0.005	0.005	0.002
3	0.005	0.004	0.003	0.002	0.001

a) Calculați probabilitatea că în ziua dată vor absenta mai mult de doi angajați (indiferent de profilul lor); b) Aflați distribuția absentarii angajaților din sectorul grădinarit. Calculați probabilitatea că în ziua dată vor absenta mai mult de doi angajați din sectorul grădinarit. c) Sunt oare v.a. X, Y independente?

3. Considerăm doua v.a. bidimensionale (X, Y) și (U, V) independente. Sunt oare v.a. Y și U independente? Argumentați de ce sau de ce nu.

4. F.d. $F(x, y)$ a v.a. bidimensionale (X, Y) este dată de formula

$$F(x, y) = (1 - e^{-x/10} - e^{-y/2} + e^{-(x+5y)/10})I_{[0,+\infty)}(x)I_{[0,+\infty)}(y).$$

Aflați: a) d.d. a v.a. (X, Y) , folosind formula 3 din Propozitia 1 anterioară; b) d.d. marginală a v.a. X ; c) f.d. a v.a. X ; d) dacă v.a. X și Y sunt independente?

3. Caracteristici numerice ale variabilelor aleatoare

3.1. Parametri de poziție: valoarea medie, moda, mediana, cuantile

În directă legătură cu cercetarea comportamentului probabilist al unei v.a. (caracteristici statistice) X nu întotdeauna prezintă interes modelul matematic exhaustiv, fie sub formă de camp de probabilitate (Ω, \mathcal{F}, P) pe care este definită X , fie sub forma de f.d. $F(x)$ a acestei v.a., fie sub forma de distribuție, dacă X este o v.a. de tip discret, fie sub forma de d.d. $f(x)$, dacă X este o v.a. de tip (absolut) continuu, ci doar unele caracteristici numerice sumare, numite și parametri. Printre ei se află și parametri numiți *parametri de poziție sau parametri ai tendinței centrale* și care reprezintă niște valori numerice sumare, de referință cu care putem compara valorile (posibile) individuale ale v.a. X . Sa luăm un exemplu concret de experiment aleator care ne va conduce la noțiunea de valoare medie, pe scurt, medie.

Exemplul 1. Este vorba de o loterie în care sunt emise și vândute, să zicem, 10000 de bilete și în care un participant cu un singur bilet de loterie poate câștiga una din sumele bănești x_1, x_2, \dots, x_k , unde $x_1 < x_2 < \dots < x_k$. Evident, suma totală acordată biletelor câștigătoare nu poate fi mai mare decât costul tuturor biletelor. Dacă vom considera "câștig" și valoarea x_0 ce coincide cu prețul unui bilet luat cu semnul minus, chiar și atunci când acesta este necâștigător, atunci câștigul X a unui jucător, ce a cumpărat un singur bilet de loterie, este o v.a. de tip discret ale cărei valori posibile fac parte din mulțimea $\mathcal{X} = \{x_0, x_1, x_2, \dots, x_k\}$. Participând de n ori la această loterie, cumpărând de fiecare dată un singur bilet de loterie și câștigând de n_i ori valoarea x_i , $i=0, 1, \dots, k$, $n_0 + n_1 + \dots + n_k = n$, putem calcula media câștigului după formula de calcul, cunoscută în Statistică sub denumirea de medie de selecție:

$$\bar{x} = \frac{1}{n} \sum_{i=0}^k n_i \cdot x_i = \sum_{i=0}^k \frac{n_i}{n} \cdot x_i = \sum_{i=0}^k f_n(X = x_i) \cdot x_i,$$

unde $f_n(X = x_i)$ este frecvența relativă a încasării câștigului x_i jucând de n ori, de fiecare dată cu un singur bilet de loterie. Dar, conform Principiului Regularității Statistice, frecvența relativă $f_n(X = x_i)$ se apropie, odată cu

creșterea lui n , tot mai mult și mai mult de probabilitatea (teoretică) $P(X = x_i)$, $i = \overline{1, k}$, ceea ce înseamnă că pentru n suficient de mare:

$$\bar{x} = \sum_{i=0}^k f_n(X = x_i) \cdot x_i \simeq \sum_{i=0}^k P(X = x_i) \cdot x_i.$$

Dar, $\{P(X = x_i)\}_{i=\overline{1, n}}$ reprezintă distribuția probabilistă a v.a. X . Prin urmare, dacă am cunoaște, apriori, această distribuție, am putea evalua o medie a câștigului nostru, pentru a decide, de exemplu, dacă merită sau nu să participăm la această loterie. Considerentele de mai sus reprezintă unul din motivele pentru care putem da

Definiția 1. Vom numi *valoare medie a v.a. X* numărul $\mathbb{E}X$ calculat după formula

$$\mathbb{E}X = \begin{cases} \sum_{i \geq 1} x_i P(X = x_i), & \text{în caz discret,} \\ \int_{-\infty}^{+\infty} x f(x) dx, & \text{în caz (absolut) continuu,} \end{cases}$$

considerând că *valoarea medie există*, dacă în caz discret

$$\sum_{i \geq 1} |x_i| P(X = x_i) < +\infty,$$

iar în caz (absolut) continuu

$$\int_{-\infty}^{+\infty} |x| f(x) dx < +\infty,$$

cu alte cuvinte, *dacă suma sau integrala Riemann respectivă converge absolut*.

Remarca 1. Din Definiția 1 deducem că, atunci când v.a. (în caz discret) ia valori dintr-o mulțime finită de valori sau (în caz (absolut) continuu) d.d. $f(x)$ este nemulă pe un număr finit de intervale de lungimi finite din \mathbb{R} , valoarea medie există întotdeauna.

Exemplul 1 (Continuare). Dacă se știe că v.a. X are distribuția $\{(x_i, p_i)\}_{i=\overline{1, k}}$, $p_i = P(X = x_i) \geq 0$, $\sum_{i \geq 1} p_i = 1$, atunci valoarea medie există întotdeauna și este egală cu

$$\mathbb{E}X = \sum_{i=1}^k x_i p_i .$$

Pe de altă parte, convergența absolută, nu este o condiție de prisos impusă pentru a garanta existența valorii medii. Aceasta o demonstrează

Exemplul 2 (O v.a. de tip discret care nu posedă valoare medie). Considerăm v.a. X dată de distribuția $P(X = (-2)^{n+1}) = \frac{1}{2^n}$, $n = 1, 2, \dots$. Observăm că seria (suma)

$$\sum_{n=1}^{+\infty} x_n P(X = x_n) = \sum_{i=1}^{+\infty} (-2)^{n+1} \frac{1}{2^n} = \sum_{i=1}^{+\infty} (-2)$$

asociată cu valoarea medie $\mathbb{E}X$, evident, nu converge absolut, prin urmare nu există valoarea medie a acestei v.a.

Exemplul 3. Considerăm v.a. X ce corespunde coordonatei unui punct ales la întâmplare din intervalul $[0, 1]$, adică v.a. dată de d.d. $f(x) = I_{[0,1]}(x)$. Atunci valoarea ei medie $\mathbb{E}X$ există și

$$\mathbb{E}X = \int_{-\infty}^{+\infty} x f(x) dx = \int_{-\infty}^{+\infty} x I_{[0,1]}(x) dx = \int_{-\infty}^0 0 dx + \int_0^1 x dx + \int_1^{+\infty} 0 dx = \frac{1}{2}.$$

Concluzie: Coordonata unui punct ales la întâmplare din intervalul $[0, 1]$ va coincide, în medie, cu mijlocul acestui interval.

Propoziția 1. (Proprietățile valorii medii). Valoarea medie a unei v.a. posedă următoarele proprietăți:

a) Dacă v.a. X este nenegativă cu probabilitatea 1 și există valoarea ei medie $\mathbb{E}X$, atunci $\mathbb{E}X \geq 0$ și $\mathbb{E}X = 0$ dacă și numai dacă $P(X = 0) = 1$;

b) Dacă valoarea medie a v.a. X există, atunci există valoarea medie a v.a. $aX + b$ și $\mathbb{E}(aX + b) = a\mathbb{E}X + b$, pentru orice $a, b \in \mathbb{R}$;

c) Dacă valoarea medie a v.a. X există, atunci există valoarea medie a v.a. $|X|$ și $|\mathbb{E}X| \leq \mathbb{E}|X|$;

d) Dacă valoarea medie a v.a. X și Y există, atunci există și valoarea medie a v.a. $X + Y$ și $\mathbb{E}(X + Y) = \mathbb{E}X + \mathbb{E}Y$;

e) Dacă v.a. X și Y sunt independente și există valorile lor medii $\mathbb{E}X$, $\mathbb{E}Y$, atunci există și valoarea medie a v.a. XY și $\mathbb{E}(X \cdot Y) = \mathbb{E}X \cdot \mathbb{E}Y$.

Exemplul 4. Considerăm două v.a. independente X și Y pentru care $\mathbb{E}X = 1$, $\mathbb{E}Y = 2$. Să se calculeze valoarea medie a v.a. $3(X + 1)Y + 5$.

Soluție. Folosind proprietățile valorii medii găsim că

$$\begin{aligned}\mathbb{E}[3(X + 1)Y + 5] &= \mathbb{E}(3XY + 3Y + 5) = 3\mathbb{E}XY + 3\mathbb{E}Y + 5 = \\ &= 3\mathbb{E}X\mathbb{E}Y + 3\mathbb{E}Y + 5 = 3 \cdot 1 \cdot 2 + 3 \cdot 2 + 5 = 15.\end{aligned}$$

Propoziția 2. (Formula de transport). a) Dacă $g : \mathbb{R} \rightarrow \mathbb{R}$ este o funcție reală de o singură variabilă și X o v.a., astfel încât $g(X)$ este o v.a. pentru care există $\mathbb{E}g(X)$, atunci are loc următoarea formulă de calcul

$$\mathbb{E}g(X) = \begin{cases} \sum_{i \geq 1} g(x_i)P(X = x_i), & \text{în caz discret,} \\ \int_{-\infty}^{+\infty} g(x)f(x)dx, & \text{în caz (absolut) continuu;} \end{cases}$$

b) Dacă $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ este o funcție reală de două variabile și (X, Y) o v.a. bidimensională, astfel încât $g(X, Y)$ este o v.a. pentru care există $\mathbb{E}g(X, Y)$, atunci are loc următoarea formulă de calcul

$$\mathbb{E}g(X, Y) = \begin{cases} \sum_{i \geq 1} \sum_{j \geq 1} g(x_i, y_j)P(X = x_i, Y = y_j), & \text{în caz discret,} \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y)f(x, y)dxdy, & \text{în caz (absolut) continuu.} \end{cases}$$

Din această propoziție, drept caz particular al proprietății b), rezultă următoarea

Consecință. Dacă v.a. bidimensională este data de distribuția ei (în caz discret) sau densitatea ei de distribuție (în caz absolut continuu), atunci putem calcula valorile medii $\mathbb{E}X$ și $\mathbb{E}Y$ conform formulelor:

$$\begin{aligned}\mathbb{E}X &= \begin{cases} \sum_{i \geq 1} \sum_{j \geq 1} x_i P(X = x_i, Y = y_j), & \text{în caz discret,} \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x f(x, y) dxdy, & \text{în caz (absolut) continuu;} \end{cases} \\ \mathbb{E}Y &= \begin{cases} \sum_{i \geq 1} \sum_{j \geq 1} y_j P(X = x_i, Y = y_j), & \text{în caz discret,} \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y f(x, y) dxdy, & \text{în caz (absolut) continuu;} \end{cases}\end{aligned}$$

Remarca 2. Sensul și utilitatea formulei de transport rezidă în faptul că, atunci când avem de a face cu o v.a. privită ca o funcție $g(X)$ de v.a. X , indiferent de dimensionalitatea v.a. X , pentru calcularea valorii medii a v.a. compuse $g(X)$ nu este obligatorie cunoașterea distribuției sau densității ei de distribuție.

În plus, dacă v.a. X este de tip discret, atunci și v.a. $g(X)$ va fi de tip discret, dar dacă v.a. X este de tip (absolut) continuu, atunci v.a. $g(X)$ poate fi sau de tip discret sau de tip (absolut) continuu sau una de tip mixt discret-continuu, aceasta în dependență de specificul funcției $g(\cdot)$.

Exemplul 5. Considerăm o v.a. X dată de distribuția $P(X = 0) = 1 - p$, $P(x = 1) = p$, $0 < p < 1$. Atunci, folosind formula de transport pentru $g(x) = X^r$, $r \geq 1$, găsim că $\mathbb{E}X^r = 0^r \cdot (1 - p) + 1^r \cdot p = p^r$. Dealtfel, anticipând, valorile de tipul $\mathbb{E}X^r$, dacă acestea există, se numesc *momente de ordinul r* , valoarea medie $\mathbb{E}X$ fiind un caz particular pentru $r = 1$.

Exemplul 6. Considerăm v.a. bidimensională (X, Y) analizată în exemplul 2, p. 2.7., iar pentru $g(x, y) = \min(x, y)$, v.a. $Z = \min(X, Y)$. Observăm că v.a. (X, Y) are d.d. $f(x, y) = I_{[0,1] \times [0,1]}(x)$, prin urmare, folosind formula de transport, găsim că

$$\mathbb{E}Z = \mathbb{E} \min(X, Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \min(x, y) f(x, y) dx dy =$$

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \min(x, y) I_{[0,1] \times [0,1]}(x) dx dy = \int_0^1 \int_0^1 \min(x, y) dx dy = 1/3.$$

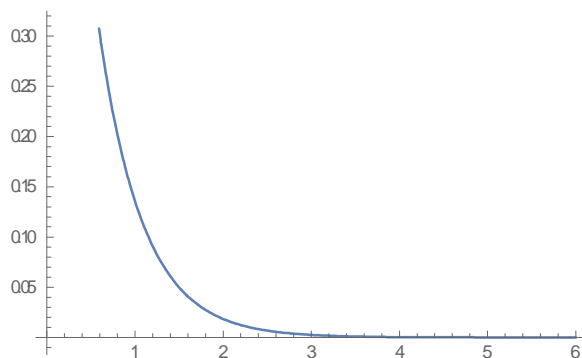
Un alt parametru de poziție utilizat este *moda*, acesta nefiind un parametru care face parte din categoria parametrilor de tip momente.

Definiția 2. Vom numi *modă* a v.a. X numărul

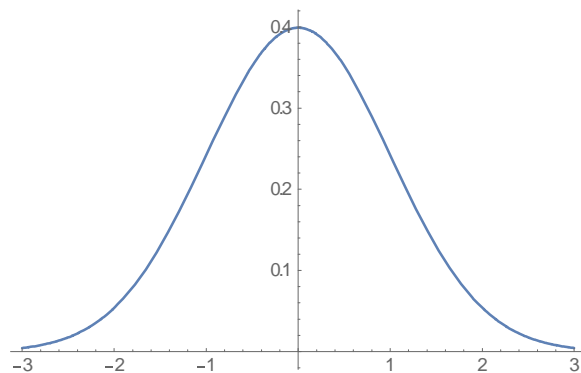
$$\text{mod}(X) = \begin{cases} x_{i_0}: \max_{i \geq 1} P(X = x_i) = P(X = x_{i_0}), \text{ în caz discret,} \\ x_0: \max_{x \in \mathbb{R}} f(x) = f(x_0), \text{ în caz (absolut) continuu.} \end{cases}$$

dacă acesta există, iar în cazul când $\text{mod}(X)$ există și este unic, atunci spunem că *distribuția* $\{P(X = x_i)\}_{i \geq 1}$ sau d.d. $f(x)$ a v.a. X este *unimodală*, altfel, dacă $\text{mod}(X)$ există și nu este unic, atunci spunem că *distribuția sau d.d. a v.a. X este multimodală*.

Exemplul 7. (Exemplu când moda nu există). Considerăm v.a. X dată de d.d. $f(x) = 2e^{-2x} \cdot I_{[0,+\infty)}(x)$. Graficul d.d. $f(x)$ de mai jos arată în mod evident că $f(x)$ nu posedă o valoare a argumentului x pentru care d.d. să ia valoare maximală.



Exemplul 8. (Exemplu de d.d. unimodală). Considerăm v.a. X dată de d.d. $f(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$, pentru orice $x \in \mathbb{R}$. Graficul d.d. $f(x)$ de mai jos arată în mod evident că $f(x)$ posedă o singură valoare a argumentului x , $x = 0$ pentru care d.d. să ia valoare maximală. Prin urmare $\text{mod}(X) = 0$.



Exemplul 9. (Exemplu de distribuție multimodală). Considerăm v.a. X dată de distribuția $P(X = 0) = P(x = 1) = 1/2$. Conform definiției, $\text{mod}(X) \in \{0, 1\}$, cu alte cuvinte *distribuția* în cauză *este bimodală*.

Un al parametru de poziție care nu face parte din categorii parametrilor de tip momente de ordinul r este prezentat în

Definiția 3. Vom numi *mediană* a v.a. X numărul a determinat din condiția că $P(X \leq a) \geq 1/2$ și totodată $P(X \geq a) \geq 1/2$. Mediana se notează $\text{Med}(X)$.

Remarca 3. Spre deosebire de *modă* și chiar spre deosebire de *valoarea medie*, mediana există întotdeauna și, în caz discret, $Med(X)$ coincide cu acea valoare reală a pentru care distribuția $\{P(X = x_i)\}_{i \geq 1}$ a v.a. X are proprietatea că

$$\sum_{i: x_i \leq a} P(X = x_i) \geq 1/2 \text{ și } \sum_{i: x_i \geq a} P(X = x_i) \geq 1/2,$$

iar în caz (absolut) continuu, cu valoarea a pentru care d.d. $f(x)$ a v.a. X are proprietatea

$$\int_{-\infty}^a f(x) dx = 1/2, \quad \int_a^{+\infty} f(x) dx = 1/2.$$

Mai mult, în caz (absolut) continuu soluția este unică, dacă f.d.

$$F(x) = \int_{-\infty}^x f(u) du$$

este continuă și strict monoton crescătoare. Această proprietate arată că valorile posibile ale v.a. X , poziționate la stânga și la dreapta de mediana ei, au aceeași masă (pondere) probabilistă. Dealtfel, folosind interpretarea geometrică a integralei și faptul că în caz (absolut) continuu d.d. $f(x)$ a v.a. X posedă proprietatea

$$\int_{-\infty}^{+\infty} f(x) dx = 1,$$

rezulta ca aria figurii cuprinse între graficul d.d. $f(x)$ și axa Ox de la $-\infty$ până la $Med(X)$ coincide cu aria figurii cuprinse între graficul d.d. $f(x)$ și axa Ox de la $Med(X)$ până la $+\infty$, ambele fiind egale cu $1/2$. A se vedea, în acest sens, v.a. X din exemplul 9 de mai sus, exemplu pentru care $Med(X) = Mod(X) = 0$, graficul d.d. $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ confirmând interpretarea noastră.

Exemplul 10 (O v.a. X pentru care mediana nu este unică). Considerăm v.a. X ce coincide cu numărul de puncte apărute la aruncarea unui zar "perfect" o singură dată. Atunci distribuția v.a. X fiind $P(X =$

$i) = 1/6, i = \overline{1, 6}$, rezultă că $Med(X)$ nu este unică deoarece inegalitățile

$$\sum_{i:i \leq a} \frac{1}{6} \geq 1/2 \text{ și } \sum_{i:i \geq a} \frac{1}{6} \geq 1/2$$

au loc pentru orice $3 \leq a \leq 4$.

Mediana reprezintă, la rândul ei, un caz particular a unei noțiuni mai generale numite *quantilă*.

Definiția 4. Vom numi *quantilă de ordinul α* , pe scurt, α -*quantilă* a v.a. X (sau $\alpha \cdot 100\%$ procentilă a distribuției/d.d. a v.a. X) acea valoare b pentru care $P(X \leq b) \geq \alpha$ și totodată $P(X \geq b) \geq 1 - \alpha$, unde $0 < \alpha < 1$.

Or, vedem că *mediana coincide* cu 0.5 -*quantila* v.a. X (50% procentila distribuției/d.d. a v.a. X). Ca și mediana, α -*quantila* întotdeauna există, dar nu pentru orice v.a. este unică. Excepție fac v.a. de tip absolut continuu.

Probleme propuse.

1. Consideră în calitate de v.a. X numărul de steme apărute la aruncarea unei monede "imperfecte" pentru care probabilitatea apariției stemei este egală cu $p, 0 < p < 1$. Aflați valoarea ei medie.

2. Fie X o v.a. dată de distribuția $P(X = k) = p(1-p)^{k-1}, k = 1, 2, 3, \dots$, unde $0 < p < 1$. Dealtfel, această distribuție descrie comportamentul probabilist al numărului de aruncări a unei monede "imperfecte" (pentru care probabilitatea apariției stemei este egală cu $p, 0 < p < 1$) până la prima apariție a stemei. Aflați valoarea medie a v.a. X .

3. La Las Vegas, roata de ruletă are un 0 și un 00 și apoi numerele de la 1 până la 36 marcate pe sloturi egale; roata este rotită și o bilă se oprește la întâmplare într-unul din sloturi. Când un jucător pariază 1 dolar pe un anumit număr, el primește 36 de dolari dacă bila se oprește la acest număr, având, astfel, un câștig net de 35 de dolari; în caz contrar, el își pierde pariul în dolari. Aflați valoarea medie pentru câștigul său.

4. Într-o a doua versiune de ruletă la Las Vegas, un jucător pariază pe roșu sau negru. Jumătate din numerele de la 1 la 36 sunt roșii, iar jumătate sunt negre. Dacă un jucător pariază un dolar pe negru și dacă bila se oprește pe un număr negru, atunci acesta își primește dolarul înapoi și încă un dolar. Dacă bila se oprește pe un număr roșu sau pe 0 sau 00, el își pierde dolarul. Găsiți valoarea medie a câștigului jucătorului în această versiune a jocului.

5. În legătură cu aruncarea unui zar "perfect" de două ori succesiv considerăm v.a. bi-dimensională (X, Y) , unde X reprezintă suma punctelor apărute

iar Y diferența lor (numărul de puncte apărute la primul zar minus numărul de puncte apărute la al doilea). Arătați că $\mathbb{E}XY = \mathbb{E}X\mathbb{E}Y$ și că X, Y sunt dependente. Cu alte cuvinte, reciproca proprietății $e)$ din Propoziția 1 nu este valabilă.

6. O companie de asigurare de viață oferă unui bărbat în vârstă de 50 de ani o valoare nominală de 1000 USD, pentru o poliță de asigurare de viață pe un an pentru o primă de 14 dolari. Tabelele de mortalitate standart indică faptul că probabilitatea ca un bărbat din această categorie de vârstă să moară în cursul anului este egală cu 0.006. Care este valoarea așteptată (medie) a câștigului pentru fiecare poliță vândută de către această companie de asigurări de viață?

7. Un mare producător auto efectuează sondaje trimestriale privind satisfacția clienților care au achiziționat automobile noi în ultimii trei ani. Proporția celor care nu au dat răspuns într-un trimestru dat este privita, din punct de vedere matematic, ca fiind un rezultat al unei v.a. X cu densitatea de distribuție $f(x) = 3x^2 I_{[0,1]}(x)$. Cu ce este egală, în medie, proporția celor care nu au dat răspuns într-un trimestru dat? Dar mediana acestei proporții, cu ce este egală?

8. Profitul unei firme este o v.a. de forma $\Pi(X) = pq(X) - rX$, unde X e o v.a. ale carei valori posibile reprezintă cantitatea zilnică de marfă (exprimată în kg), sub formă de produs agricol perisabil, furnizată firmei spre procesare, unde $p = 5$ reprezintă prețul produsului procesat per kg , $r = 2$ este prețul de cumparare per kg de către firmă a mărfii, iar funcția $q(x) = x^{0.9}$ arată care este cantitatea de produs, rezultat în urma procesării unei cantități x de marfă. În presupunerea că X este o v.a. a cărui comportament probabilist este guvernat de densitatea de distribuție $f(x) = \frac{1+2x}{110} I_{[0,10]}(x)$, aflați valoarea medie a profitului zilnic.

9. O firmă producătoare de calculatoare vinde calculatoare cu o garanție de 3 ani, ce constă în înlocuirea calculatorului cu unul nou dacă în acest rastimp iese din funcțiune memoria hard. Presupunem că durata vieții memoriei hard, exprimată în ani, este o v.a. X cu densitatea de distribuție $f(x) = 0.005e^{-0.005x} I_{[0,+\infty)}(x)$. a) Cu ce este egală probabilitatea că memoria hard a unui astfel de calculator va avea o dura de viața de cel mult 3 ani; b) Calculați durata medie a funcționalității memoriei hard pentru un astfel de calculator; c) Calculați mediana v.a. X ; d) Este oare v.a. X unimodală sau multimodală?

10. Presupunem ca durata vieții unui bec electric de tip Super Light este o v.a. X dată de densitatea de distribuție $f(x) = \lambda^2 x e^{-\lambda x} I_{[0,+\infty)}(x)$ cu $\lambda =$

0.05. Aflați durata medie a vieții a acestui tip de bec electric.

11. Dealerul unei firme de vânzare a autoturismelor este specializat în vânzarea a două mărci de mașini de teren: Range Rover și Dacia Duster. Numărul de autoturisme vândute săptămânal, corespunzător acestor mărci, reprezintă o realizare a unei v.a. bidimensionale (X, Y) guvernată de următoarea distribuție probabilistică în ansamblu:

X / Y	0	1	2	3	4
0	0.20	0.15	0.075	0.05	0.03
1	0.10	0.075	0.04	0.03	0.02
2	0.05	0.03	0.02	0.01	0.01
3	0.04	0.03	0.02	0.01	0.01

La salariul sau de 100 USD săptămânal dealerul mai încasează prime a câte 200 USD pentru fiecare mașină Range Rover vândută și 100 USD pentru fiecare mașină Dacia Duster vândută.

a) Sunt oare v.a. X și Y independente?

b) Aflați valoarea medie a comisionului săptămânal încasat de dealer pentru vânzarea autoturismelor. Dar valoarea totală a plății încasate de dealer săptămânal pentru activitatea sa de vânzător auto cu ce este egală?

c) Aflați valoarea medie a comisionului săptămânal încasat de dealer pentru vânzarea autoturismelor Range Rover. Dar valoarea medie a comisionului săptămânal încasat de dealer pentru vânzarea autoturismelor Dacia Duster cu ce este egală?

3.2. Dispersia (varianța), abaterea standard, covarianța, coeficientul de corelație, regresie liniară

Vom începe cu un exemplu care demonstrează necesitatea introducerii unui parametru sau a unei caracteristici numerice a v.a. X pentru a măsura gradul de împrăștiere a valorilor ei individuale în jurul unui parametru de poziție cum ar fi, de exemplu, valoarea medie a v.a. X .

Exemplul 1. Din câte se știe, nu există aparate de măsurare absolut exacte, rezultatele măsurărilor fiind influențate de o mulțime de factori (temperatura, umiditatea mediului înconjurător, materialul din care este produs aparatul, cine este producătorul, etc.). Or, modelele matematice corespunzătoare sunt de natură probabilistică. Considerăm, de pildă, că v.a. X , ce reprezintă rezultatul unei măsurări executate cu un aparat de măsurare, este dată de distribuția

$$X : \left(\begin{array}{ccc} a - \varepsilon & a & a + \varepsilon \\ (1 - p)/2 & p & (1 - p)/2 \end{array} \right), 0 < p < 1,$$

a fiind valoarea exactă a mărimii măsurate iar $\varepsilon > 0$, eroarea de măsurare.

Valoarea ei medie a v.a. X este egală cu

$$\mathbb{E}X = (a - \varepsilon) \cdot \frac{1 - p}{2} + a \cdot p + (a + \varepsilon) \cdot \frac{1 - p}{2} = a.$$

Cu alte cuvinte, un astfel de aparat ne arată "în medie" valoarea exactă a mărimii măsurate, dar această constatare nu ne ajută cu nimic la caracterizarea calității aparatului nostru de măsurat. Noțiunea următoare vine să compenseze această lacună.

Definiția 1. Vom numi *dispersie* sau *variantă* a v.a. X numărul $\mathbb{D}X = \mathbb{E}(X - \mathbb{E}X)^2$, dacă această valoare medie există, bineînțeles.

Exemplul 1 (Continuare). În exemplul nostru pentru a calcula dispersia v.a. X putem folosi Formula de transport din p.3.1 și faptul că $\mathbb{D}X = \mathbb{E}g(X)$, unde $g(x) = (x - \mathbb{E}X)^2 = (x - a)^2$. Prin urmare

$$\mathbb{D}X = (a - \varepsilon - a)^2 \cdot \frac{1 - p}{2} + (a - a)^2 \cdot p + (a + \varepsilon - a)^2 \cdot \frac{1 - p}{2} = \varepsilon^2 \cdot (1 - p).$$

Distribuția v.a. X arată că instrumentul corespunzător de măsurare este cu atât mai bun cu cât valoarea erorii de măsurare ε este mai mică iar probabilitatea $p = P(X = a)$ este mai aproape de 1, dar aceasta are loc atunci și numai atunci când valoarea dispersiei este mai aproape de 0. Cu alte cuvinte, cu cât dispersia este mai mică, cu atât este mai mic gradul de împrăștiere a valorilor individuale a v.a. X față de valoarea ei medie $\mathbb{E}X = a$.

Pentru a simplifica, uneori, calculul dispersiei este utilă

Propoziția 1. *Dispersia v.a. X poate fi calculată după formula*

$$\mathbb{D}X = \mathbb{E}X^2 - (\mathbb{E}X)^2.$$

Exemplul 2. Calculați dispersia v.a. X date de d.d. $f(x) = I_{[0,1]}(x)$, adică X fiind uniform distribuită pe $[0, 1]$.

Soluție. Deoarece,

$$\mathbb{E}X = \int_0^1 x dx = 1/2, \mathbb{E}X^2 = \int_0^1 x^2 dx = 1/3,$$

rezultă $\mathbb{D}X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = 1/3 - (1/2)^2 = 1/12$.

Remarca 1. Din definiția dispersiei v.a. X desprindem, că dispersia nu are aceeași unitate de măsură ca și v.a. X sau valoarea ei medie. Astfel, în exemplul 1, dacă X se masoară în cm valoarea ei medie se masoară, la fel, în cm , dar dispersia ei se masoară în cm^2 . Notiunea introdusă în definiția următoare păstrează calitățile dispersiei ca măsură a gradului de împrăștiere, dar este, în schimb, exprimată în aceeași unitate de măsură ca și v.a. X .

Definiția 2. Vom numi *abatere standard* a v.a. X numărul $\sigma = \sqrt{\mathbb{D}X}$.

La calcularea dispersiei, prin urmare și a abaterii standard, sunt utile proprietățile dispersiei centralizate în

Propoziția 2. *Dispersia posedă următoarele proprietăți:*

- a) $\mathbb{D}X \geq 0$, iar $\mathbb{D}X = 0$ atunci și numai atunci, când $P(X = \mathbb{E}X) = 1$;
- b) Dacă există dispersia $\mathbb{D}X$ a v.a., atunci pentru orice $a, b \in \mathbb{R}$ există dispersia v.a. $aX \pm b$ și $\mathbb{D}(aX \pm b) = a^2 \mathbb{D}X$;
- c) Dacă există dispersiile $\mathbb{D}X$ și $\mathbb{D}Y$ ale v.a. X și Y , atunci există dispersia v.a. $X \pm Y$ și

$$\mathbb{D}(X \pm Y) = \mathbb{D}X + \mathbb{D}Y \pm 2 \cdot \mathbb{E}(X - \mathbb{E}X)(Y - \mathbb{E}Y);$$

- d) Dacă v.a. X și Y sunt independente și există dispersiile lor $\mathbb{D}X$ și $\mathbb{D}Y$, atunci există dispersia v.a. $X \pm Y$ și

$$\mathbb{D}(X \pm Y) = \mathbb{D}X + \mathbb{D}Y.$$

Definiția 3. Vom numi *covarianța* a două v.a. X și Y numărul $Cov(X, Y) = \mathbb{E}(X - \mathbb{E}X)(Y - \mathbb{E}Y)$. Dacă această valoare medie există, atunci, din Formula de Transport rezultă ca, ea poate fi calculată după formula

$$Cov(X, Y) == \begin{cases} \sum_{i \geq 1} \sum_{j \geq 1} (x_i - \mathbb{E}X)(y_j - \mathbb{E}Y)P(X = x_i, Y = y_j), & \text{in caz} \\ \text{discret, } \{P(X = x_i, Y = y_j)\}_{i, j \geq 1} \text{ fiind distrib. v.a. } (X, Y); \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x_i - \mathbb{E}X)(y_j - \mathbb{E}Y)f(x, y)dxdu, & \text{in caz (absolut)} \\ \text{continuu, } f(x, y) \text{ fiind d.d. a v.a. } (X, Y). \end{cases}$$

Din definiția covarianței, dar și din pp. c)-d) ale Propoziției 2, deducem următoarea

Consecință. a) Dacă covarianța v.a. X și Y există, atunci aceasta poate fi calculată după formula

$$Cov(X, Y) = \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y;$$

b) Dacă v.a. X și Y sunt independente atunci $Cov(X, Y) = 0$.

Remarca 2. Reciproca afirmației b) din Consecință nu are loc, așa cum arată și următorul

Contraexemplu (Două v.a. de covariință nula, dar dependente).

Considerăm v.a. X, Y independente cu valorile medii $\mathbb{E}X = \mathbb{E}Y = 0$ și pentru care există $\mathbb{E}X^2$.

Atunci v.a. $Z = XY$, fiind dependentă de v.a. X , va avea $\mathbb{E}Z = \mathbb{E}XY = \mathbb{E}X\mathbb{E}Y = 0$, dar aceste două v.a. X și Z au, totuși,

$$Cov(X, Z) = \mathbb{E}(X - \mathbb{E}X)(Z - \mathbb{E}Z) = \mathbb{E}XZ = \mathbb{E}X^2Y = \mathbb{E}X^2\mathbb{E}Y = \mathbb{E}X^2 \cdot 0 = 0.$$

Concluzie. Consecința formulată mai sus, împreună cu contraexemplul nostru, arată, de fapt, că *orice două v.a. X, Y care au $Cov(X, Y) \neq 0$ sunt dependente.*

Propoziția 3 (Inegalitatea Cauchy-Buniakovski). Dacă există dispersiile $\mathbb{D}X$ și $\mathbb{D}Y$ ale v.a. X și Y , atunci există $Cov(X, Y)$ și are loc inegalitatea

$$|Cov(X, Y)| \leq \sqrt{\mathbb{D}X\mathbb{D}Y}.$$

Exemplul 1, p.2.6 (Continuare). Așa cum am văzut, v.a. (X, Y) , unde X reprezintă indicatorul parității sumei punctelor apărute iar Y - indicatorul parității produsului punctelor apărute la aruncarea unui zar "perfect" de două ori succesiv, are distribuția (în ansamblu)

$X \setminus Y$	0	1
0	0	1/2
1	1/4	1/4

cu distribuțiile marginale

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

Atunci,

$$\mathbb{E}X = 1/2, \mathbb{D}X = 1/4, \sigma_X = 1/2, \mathbb{E}Y = 3/4, \mathbb{D}Y = 3/16, \sigma_Y = \sqrt{3}/4,$$

iar

$$\begin{aligned} \mathbb{E}XY &= 0 \cdot 0 \cdot P(X = 0, Y = 0) + 0 \cdot 1 \cdot P(X = 0, Y = 1) + \\ &+ 1 \cdot 0 \cdot P(X = 1, Y = 0) + 1 \cdot 1 \cdot P(X = 1, Y = 1) = 1/4 \end{aligned}$$

$$\text{Cov}(X, Y) = \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y = 1/4 - (1/2)(3/4) = -1/8.$$

Inegalitatea Cauchy-Buniakovski se verifică cu ușurință:

$$|\text{Cov}(X, Y)| = 1/8 = 0.125 \leq \sigma_X\sigma_Y = \sqrt{3}/8 = 0.21651.$$

În altă ordine de idei, la sfârșitul p.2.7, am arătat, reeșind din definiția independenței (dependenței) v.a., ca v.a. X, Y din exemplul de mai sus sunt dependente, dar același lucru rezultă și din faptul că $\text{Cov}(X, Y) \neq 0$. Într-adevar, presupunem contrariul, că X, Y sunt independente. Atunci, drept consecință, $\text{Cov}(X, Y) = 0$. Dar în exemplul nostru $\text{Cov}(X, Y) = -1/8$. Contradicție, ce arată că nenulitatea covarianței a două v.a. poate fi considerată o condiție suficientă, dar nu și necesară (vezi contraexemplul de mai sus) ca acestea să fie dependente. Doar atât, nu putem, însă, spune nimic despre gradul lor de dependență sau asociere. În acest scop servește noțiunea din

Definiția 4. Vom numi *coeficient de corelație* a două v.a. X și Y numărul $\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\mathbb{D}X\mathbb{D}Y}} = \frac{\text{Cov}(X, Y)}{\sigma_X\sigma_Y}$.

Așa cum arată Inegalitatea Cauchy-Buniakovski, coeficientul de corelație există, deîndată ce există dispersiile $\mathbb{D}X, \mathbb{D}Y$. Mai mult, din aceeași inegalitate și consecința de mai sus rezultă

Propoziția 4. *Coeficientul de corelație $\rho(X, Y)$ a două v.a. X și Y are proprietățile:*

- a) Dacă v.a. X și Y sunt independente, atunci $\rho(X, Y) = 0$;
- b) $-1 \leq \rho(X, Y) \leq 1$.

În exemplul de mai sus $\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\mathbb{D}X\mathbb{D}Y}} = \frac{-1/8}{\sqrt{3}/8} = -0.57735$. Despre ce fel de dependență și în ce grad putem vorbi în baza valorii acestea? Pentru a putea răspunde la această întrebare luăm, pentru comparație, cazul dependenței liniare: $Y = aX + b$, care se mai numește *regresie liniară*, unde $a, b \in \mathbb{R}, a \neq 0$. Cum $\mathbb{D}Y = a^2 \mathbb{D}X$,

$$\begin{aligned} \text{Cov}(X, Y) &= \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y = \mathbb{E}X(aX + b) - \\ &- \mathbb{E}X(a\mathbb{E}X + b) = a[\mathbb{E}X^2 - (\mathbb{E}X)^2] = a\mathbb{D}X, \end{aligned}$$

rezultă că

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\mathbb{D}X\mathbb{D}Y}} = \frac{a\mathbb{D}X}{\sqrt{a^2(\mathbb{D}X)^2}} = \frac{a}{|a|} = \begin{cases} 1, & \text{dacă } a > 0, \\ -1, & \text{dacă } a < 0. \end{cases}$$

Așadar, dacă v.a. Y depinde liniar de v.a. X , atunci $|\rho(X, Y)| = 1$. Important e că are loc și reciproca acestei afirmații. Ma exact are loc

Propoziția 5. *Coeficientul de corelație $|\rho(X, Y)| = 1$, atunci și numai atunci când există $a, b \in \mathbb{R}$, $a \neq 0$, astfel încât probabilitatea $P(Y = aX + b) = 1$.*

Remarca 3. Dacă $\rho(X, Y) = 1$, atunci *regresia liniară este pozitivă*, deoarece în acest caz coeficientul $a > 0$, ceea ce atrage după sine creșterea lui Y odată cu creșterea lui X și invers, dacă $\rho(X, Y) = -1$, atunci *regresia liniară este negativă*, deoarece în acest caz coeficientul $a < 0$, ceea ce atrage după sine descreșterea lui Y odată cu creșterea lui X .

În concluzie, dacă revenim la exemplul nostru de mai sus, faptul că $\rho(X, Y) = -0.57735$ arată, că regresia dintre Y și X este *neliniară*.

Are loc următoarea

Teoremă (*Predicția liniară cea mai bună pentru valoarea posibilă a v.a. Y*). *Dacă v.a. X și Y posedă, cel puțin, momentele inițiale de ordinul 2, atunci există două valori $a, b \in \mathbb{R}^2$, $a = \text{cov}(X, Y)/\mathbb{D}X$, $b = \mathbb{E}Y - a\mathbb{E}X$, pentru care valoarea medie a patratului distanței dintre Y și $\hat{Y} = aX + b$, fiind egală cu $\mathbb{E}(Y - \hat{Y})^2 = \mathbb{E}[Y - (aX + b)]^2$, ia valoarea cea mai mică.*

Exemplu 3. Prețul X în *Euro per litru* al combustibilului și al cantității Y de combustibil exprimat în mii de tone vândute de către o companie de carburanți într-o zi pot fi modelate, din punct de vedere matematic ca fiind o v.a. (X, Y) guvernată de o legitate probabilistă dată de densitatea de distribuție

$$f(x, y) = 2xe^{-xy}I_{[0.5, 1]}(x)I_{(0, +\infty)}(y).$$

a) Determinați predicția liniară cea mai bună pentru cantitatea Y în funcție de prețul X ;

b) Cu ce va fi egală cantitatea medie de combustibil vândut dacă prețul combustibilului este dat de valoarea $x = 0.95$ *Euro*?

c) Care va fi suma medie în euro a încasării zilnice pentru combustibilul vândut?

Soluție. Calculăm, mai întâi, valorile necesare pentru a determina coeficienții a, b care definesc predicția liniară cea mai bună.

$$\mathbb{E}X = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xf(x, y)dx dy = \int_{0.5}^1 \int_0^{+\infty} x \cdot 2xe^{-xy}dx dy =$$

$$\int_{0.5}^1 2x^2 \left(\frac{-1}{x} e^{-xy} \Big|_0^{+\infty} \right) dx = \int_{0.5}^1 2x dx = x^2 \Big|_{0.5}^1 = 1 - 0.25 = 0.75.$$

$$\mathbb{E}X^2 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^2 f(x, y) dx dy = \int_{0.5}^1 \int_0^{+\infty} x^2 \cdot 2xe^{-xy} dx dy = 0.583333.$$

$$\mathbb{D}X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = 0.583333 - 0.75^2 \simeq 0.021.$$

$$\mathbb{E}Y = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y f(x, y) dx dy = \int_{0.5}^1 \int_0^{+\infty} y \cdot 2xe^{-xy} dx dy =$$

$$\int_0^{+\infty} 2y \left(\int_{0.5}^1 xe^{-xy} dx \right) dy = \int_0^{+\infty} 2y(1 - e^{-xy}) \Big|_{x=0.5}^1 dy \simeq 1.39$$

$$\mathbb{E}XY = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy f(x, y) dx dy = \int_{0.5}^1 \int_0^{+\infty} xy \cdot 2xe^{-xy} dx dy = 1$$

$$\text{Cov}(X, Y) = \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y \simeq 1 - 0.75 \cdot 1.39 = -0.0425$$

Prin urmare:

a) Din Teorema anterioară deducem ca $a = \text{cov}(X, Y) / \mathbb{D}X = -0.0425 / 0.021 = -2.0238$, iar $b = \mathbb{E}Y - a\mathbb{E}X \simeq 1.39 - ((-0.0425) / 0.021) \times 0.75 = 2.9079$. Prin urmare cea mai buna predicție liniară pentru v.a. Y în funcție de X este v.a. $\hat{Y} = aX + b = -2.0238X + 2.9079$.

b) Atunci când prețul $X = 0.95$ Euro/litru predicția cantității zilnice de combustibil vândut $\hat{Y} = -2.0238 \times 0.95 + 2.9079 = 0.98529$ mii tone de combustibil;

c) Suma medie în euro a încasării zilnice pentru combustibilul vândut conform predicției va fi egală cu

$$\begin{aligned} \mathbb{E}(X \times (1000)^2 \hat{Y}) &= (1000)^2 \mathbb{E}X\hat{Y} = (1000)^2 \mathbb{E}(2.9079X - 2.0238X^2) = \\ &= (1000)^2 \times (2.9079\mathbb{E}X - 2.0238\mathbb{E}X^2) = \\ &= (1000)^2 \times (2.9079 \times 0.75 - 2.0238 \times 0.583333) = 1.0004 \times 10^6 \text{ Euro} \end{aligned}$$

Probleme propuse.

1. În condițiile problemelor 1 – 4, 6 – 10 din lista de Probleme propuse în paragraful anterior, calculați și dispersia v.a. abordate la calcularea valorii medii.

2. Un joc de noroc este considerat ca fiind unul "echitabil" sau "corect" dacă valoarea medie a sumei câștigate de către jucător este nulă. Sa se verifice dacă următorul joc este "echitabil" și să se calculeze dispersia sumei câștigate. Jocul constă în următoarele: jucătorul aruncă o perche de zaruri "perfecte", plătind suma de N unități monetare (u.m.). Dacă suma punctelor apărute este egală cu 7 sau 11, atunci jucătorului i se returnează suma egală cu N u.m., plus o sumă în valoare de $2N$ u.m., în caz contrar jucătorul pierde suma introdusă.

3. Fie X un număr ales la întâmplare din mulțimea de numere $\{-1, 0, 1\}$. Calculați valoarea medie, dispersia și abaterea standard a v.a. X .

4. Considerăm v.a. X pentru care valoarea medie $\mathbb{E}X = 100$ și abaterea standard $\sigma = 15$. Aflați cu ce sunt egale valorile: a) $\mathbb{E}X^2$; b) $\mathbb{E}(3X + 10)$; c) $\mathbb{E}(-X)$; d) $\mathbb{D}X$; e) $\mathbb{D}(3X + 10)$; f) $\mathbb{D}(-3X - 10)$; g) $\mathbb{D}(3X - 10)$.

5. Un sondaj statistic organizat de către guvern interoga populația privind proiectul de construire a unei stații electrice atomice, având la bază un eșantion ce includea 2400 de persoane alese la întâmplare. Drept rezultat au fost înregistrate 40% de răspunsuri favorabile, restul fiind nefavorabile. Aflați dispersia și abaterea standard pentru numărul S_{2400} ce coincide cu numărul persoanelor incluse în eșantion care s-au exprimat în favoarea construirii unei stații atomice.

5. Este dată o v.a. X cu valoarea medie $\mathbb{E}X = \mu$ și abaterea standard $\sqrt{\mathbb{D}X} = \sigma$. Construim o nouă v.a. $Y = (X - \mu)/\sigma$ numită *normare* sau *standardizare* a v.a. X . Arătați că $\mathbb{E}Y = 0$ iar $\mathbb{D}Y = 1$.

6. Considerăm două v.a. X, Y independente, identic distribuite cu d.d. $f(x) = I_{[0,1]}(x)$. Introducem v.a. $Z = X + Y$ și $W = X - Y$. Calculați coeficienții de corelație: a) $\rho(X, Y)$; b) $\rho(X, Z)$; c) $\rho(Y, W)$; d) $\rho(Z, W)$.

7. Prețul estimativ X exprimat în USD și cantitatea totală Y de pixuri de o anumită marcă realizată într-o rețea de magazine de specialitate, având ca unitate de măsură 10000 de pixuri, pe durata unei anumite perioade de vânzare, reprezintă o v.a. bidimensională (X, Y) cu d.d. $f(x, y) = 10xe^{-ps} I_{[0,20]}(x) I_{[0,+\infty]}(y)$.

a) Determinați cea mai bună predicție liniară \hat{Y} pentru cantitatea Y în funcție de prețul X ;

b) Cu ce va fi egală cantitatea medie de pixuri vândute dacă prețul unui

pix este dat de valoarea $x = 0.12$ USD?

c) Cu ce va fi egală valoarea medie în USD a încasării totale pentru vânzările de pixuri pentru perioada dată, i.e. valoarea $\mathbb{E}(10000X\hat{Y})$ USD.

3.3. Momente ale variabilei aleatoare (inițiale, centrale), asimetria, boltirea (aplatizarea)

Atunci când sunt puse în discuție astfel de aspecte ca tendința centrală, gradul de înrăștiere a unei v.a. X , forma distribuției sau d.d. a acestei v.a., dar și în cadrul aplicării unor proceduri statistice, sunt utile nu numai valoarea medie și dispersia, dar și valorile medii ale v.a. X^r sau $(X - \mathbb{E}X)^r$, calculate și pentru valori ale lui r diferite de 1 sau de 2. Toate aceste valori medii poartă denumirea de *momente*, cu ele întâlnindu-ne în exemplul 5, p.3.1. După cum vom vedea, deosebim *momente inițiale* și *momente centrale* conform cu

Definiția 1. Vom numi *moment inițial de ordinul $r \geq 0$ a v.a. X* numărul $\alpha_r = \mathbb{E}X^r$, dacă această valoare medie există, bineînțeles.

Definiția 2. Vom numi *moment central de ordinul $r \geq 0$ a v.a. X* numărul $\mu_r = \mathbb{E}(X - \mathbb{E}X)^r$, dacă această valoare medie există, bineînțeles.

Observăm ca $\alpha_0 = \mu_0 = 1$, iar $\alpha_1 = \mathbb{E}X$, $\mu_1 = 0$, $\mu_2 = \mathbb{D}X$. În plus, atunci când α_r , μ_r există, aplicarea Formulei de transport conduce la următoarele formule de calcul direct, care în cazul discret, adică în cazul când v.a. este dată de distribuția $\{P(X = x_i)\}_{i \geq 1}$, sunt formulele

$$\alpha_r = \sum_{i \geq 1} x_i^r P(X = x_i), \quad \mu_r = \sum_{i \geq 1} (x_i - \mathbb{E}X)^r P(X = x_i),$$

iar în cazul (absolut) continuu, adică în cazul când v.a. este dată de d.d. $f(x)$, sunt formulele

$$\alpha_r = \int_{-\infty}^{+\infty} x^r f(x) dx, \quad \mu_r = \int_{-\infty}^{+\infty} (x - \mathbb{E}X)^r f(x) dx.$$

Înloc de formula directă, la calcularea momentelor centrală, devine utilă formula ce exprimă legătura dintre momentele inițiale și cele centrale din

Propoziția 1 (Momentele centrale ca funcție de momentele inițiale). Dacă pentru valoarea $r \in \{1, 2, \dots, n, \dots\}$ există momentul central μ_r de ordinul r al v.a. X , atunci

$$\mu_r = \sum_{i=0}^r (-1)^i \mathbb{C}_r^i \alpha_i \alpha_{r-i}.$$

Privitor la existența momentelor este utilă și

Propoziția 2. Dacă, pentru $r > 0$, există momentul inițial $\mathbb{E}X^r$, respectiv, momentul central $\mathbb{E}(X - EX)^r$, atunci pentru orice $s \in [0, r]$ există momentul inițial $\mathbb{E}X^s$, respectiv, momentul central $\mathbb{E}(X - EX)^s$.

Exemplul 1. Presupunem că pentru v.a. X există momentul central de ordinul $r = 4$. Atunci, din propozițiile 1,2 rezultă că există și momentele inițiale și centrale de ordinele 1, 2, 3 și

$$\begin{aligned} \mu_1 &= \alpha_1, \\ \mu_2 &= \alpha_2 - \alpha_1^2, \\ \mu_3 &= \alpha_3 - 3\alpha_1\alpha_2 + 2\alpha_1^3, \\ \mu_4 &= \alpha_4 - 4\alpha_1\alpha_3 + 6\alpha_1^2\alpha_2 - 3\alpha_1^4. \end{aligned}$$

Aceste din urmă formule sunt cele care se aplică cel mai des în Statistică, mai ales în Statistica Descriptivă, pentru a calcula, în baza datelor statistice cei doi parametri ce descriu forma distribuției sau d.d. a v.a. (caracteristicii statistice) X : coeficientul de Asimetrie și coeficientul de Aplatizare (Boltire).

Definiția 3. Vom numi *coeficient de asimetrie a distribuției/d.d. a v.a. X* numărul $As(X) = \mu_3/\sigma^3$, unde μ_3 este momentul central de ordinul 3, iar σ abaterea standard a v.a. X , dacă μ_3 există, bineînțeles. Vom spune că *distribuția/d.d. v.a. X este: simetrică* dacă $As(X) = 0$ sau ($\mu_3 = 0$), *pozitiv asimetrică* dacă $As(X) > 0$ sau ($\mu_3 > 0$), *negativ asimetrică* dacă $As(X) < 0$ sau ($\mu_3 < 0$).

Exemplul 2 (O v.a. cu d.d. simetrică). Astfel, v.a. X (standard normal distribuită) cu d.d. $f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ este simetrică. A se vedea graficul acestei d.d. în exemplul 8, p.3.1. Mai mult, funcțiile $\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, $\frac{1}{\sqrt{2\pi}}x^3e^{-\frac{x^2}{2}}$ fiind funcții impare pe domeniul $(-\infty, +\infty)$, rezultă că

$$\mathbb{E}X = \int_{-\infty}^{+\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0, \mu_3 = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} x^3 e^{-\frac{x^2}{2}} dx = 0,$$

ceea ce înseamnă ca $As(X) = 0$. Apropo, în acest exemplu $\mathbb{E}X = Med(X) = Mod(X) = 0$.

Pentru caracterizarea formei distribuției/d.d. a unei v.a., înafară de coeficientul de asimetrie, se mai folosește și coeficientul de Boltire (Aplatizare).

Definiția 4. Vom numi *coeficient de boltire (aplatizare) a distribuției/d.d. a v.a. X* numărul $Ap(X) = \mu_4/\sigma^4$, unde μ_4 este momentul central de ordinul 4, iar σ abaterea standard a v.a. X , dacă μ_4 există, bineînțeles.

Exemplul 2 (Continuare). Pentru v.a. X (*standard normal distribuită*), folosind formula integrării prin părți, găsim că momentele centrale de ordinele 2 și 4 este sunt egale, respectiv, cu

$$\mu_2 = \mathbb{D}X = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} x^2 e^{-\frac{x^2}{2}} dx = 1, \quad \mu_4 = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} x^4 e^{-\frac{x^2}{2}} dx = 3.$$

Prin urmare abaterea standard a v.a. X (acum putem spune, a v.a. *standard normal distribuită cu media 0 și dispersia 1*) este egală cu 1, iar coeficientul de aplatizare $Ap(X) = \mu_4/\sigma^4 = 3$.

Remarca 1. Comparând definiția coeficientului de boltire și faptul, ca pentru distribuția normală cu media 0 și dispersia 1, acesta este egal cu 3, deacum înainte gradul de boltire a distribuției/d.d. se va compara cu gradul de boltire a d.d. a unei v.a. standard normal distribuite, spunând ca *boltirea unei d.d. e negativă*, dacă $\mu_4/\sigma^4 < 3$, adică *boltirea ei e mai mică (graficul ei este mai plat) decât boltirea d.d. a v.a. standard normal distribuite cu media 0 și dispersia 1*, ori că *boltirea unei d.d. e pozitivă*, dacă $\mu_4/\sigma^4 > 3$, adică *boltirea ei e mai mare (graficul ei este mai ascuțit) decât boltirea d.d. a v.a. standard normal distribuite cu media 0 și dispersia 1*.

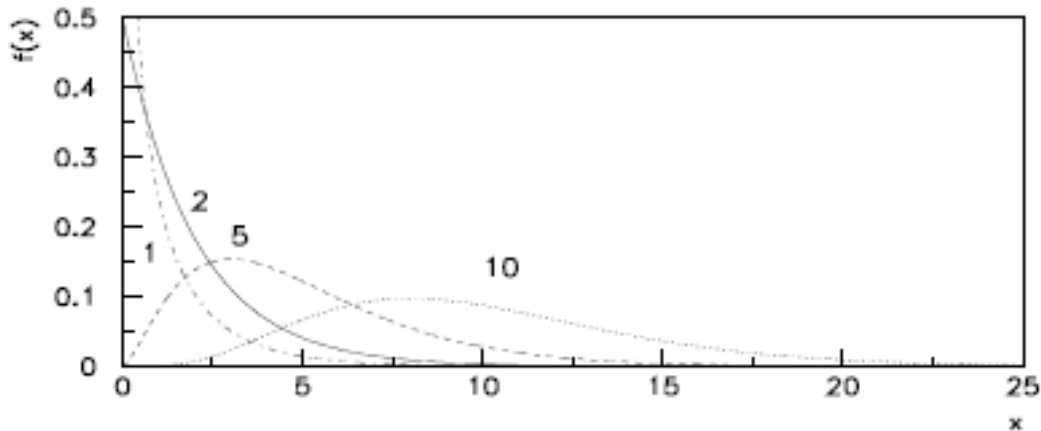
Exemplul 3 (V.a. cu d.d. pozitiv asimetrică în diverse variante de boltire (aplatizare)). Considerăm v.a. X cu d.d.

$$f(x) = \frac{x^{(n-2)/2} \exp(-x/2)}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \cdot I_{[0,+\infty)}(x),$$

unde $\Gamma(u) = \int_0^{+\infty} x^{u-1} e^{-x} dx$ este funcția gamma.

Modelul acesta, adică d.d., coincide cu d.d. a v.a. $X_1^2 + X_2^2 + \dots + X_n^2$, unde X_1, X_2, \dots, X_n sunt variabilele aleatoare, independente, identic distribuite (v.a.i.i.d.) normal cu media 0 și dispersia 1, d.d. ce poartă denumirea de *distribuție Hi-patrat cu n grade de libertate*. Conform cărții Forbes C.,

Evans M., et alls, *Statistical Distributions*, Ed. John Willey&Sons, USA, 2010, $\mathbb{E}X = n$, $\mathbb{D}X = 2n$, $Med(X) \simeq n - \frac{2}{3}$, $Mod(X) = n - 2$, pentru $n \geq 2$, coeficientul de asimerie $As(X) = \frac{2^{3/2}}{\sqrt{n}}$, coeficientul de aplatizare $Ap(X) = 3 + 12/n$. Iata cum arata pe unul si același grafic d.d. Hi-patrat pentru valorile parametrului $\nu=1, 2, 5, 10$.



Or, din grafice se vede cum arată *d.d. pozitiv asimetrică*. Cum $0 < As(X) = \frac{2^{3/2}}{\sqrt{n}} \xrightarrow{n \rightarrow \infty} 0$, $Ap(X) = 3 + 12/n \xrightarrow{n \rightarrow \infty} 3$, se vede cum graficul d.d. tinde să devină simetric, gradul de aplatizare apropiindu-se de cel al d.d. normale. Cu alte cuvinte, odată cu creșterea numărului gradelor de libertate ν d.d. Hi-patrat se apropie de d.d. normală, lucru folosit pe larg in Statistica Matematică.

In opoziție cu *asimetria pozitivă*, devine clar cum arata grafic o *distribuție/d.d. negativ asimetrică*.

Remarca 2. Exemplele aduse mai scot in evidență următoarele situații:

- daca distribuția/d.d. este simetrică ($As(X) = 0$), atunci $Mod(X) = Med(X) = \mathbb{E}X$;
- daca distribuția/d.d. este pozitiv asimetrică ($As(X) > 0$), atunci $Mod(X) < Med(X) < \mathbb{E}X$;
- daca distribuția/d.d. este negativ asimetrică ($As(X) < 0$), atunci $\mathbb{E}X < Med(X) < Mod(X)$.

Probleme propuse.

1. Pentru următoarele v.a. date de distribuția probabilistă sau d.d. determinați dacă există valoarea medie, dispersia. În plus, aflați modul, mediana și stabiliți care distribuție este simetrică:

a) $P(X = x) = \mathbb{C}_4^x(0.2)^x(0.8)^{4-x}I_{\{0,1,2,3,4\}}(x)$;

b) $f(x) = 3x^2I_{[0,1]}(x)$;

c) $f(x) = 2x^{-3}I_{[1,+\infty)}(x)$;

d) $f(x) = [\pi(1+x^2)]^{-1}I_{(-\infty,+\infty)}(x)$.

2. Primele trei momente inițiale ale v.a. X sunt egale cu $\alpha_1 = 0.5$, $\alpha_2 = 0.5$, $\alpha_3 = 0.75$.

a) Determinați primele trei momente centrale μ_i , $i = \overline{1,3}$, ale v.a. X ;

b) Stabiliți dacă distribuția/d.d. este antisimetrică. De ce sau de ce nu?

4. Modele (distribuții, d.d.) probabiliste uzuale, inegalități, Legea Numerelor Mari, Teorema Limită Centrală

4.1. Distribuții probabiliste uzuale în caz discret (Uniformă, Bernoulli, Binomială, Geometrică, Poisson, Multinomială, Hypergeometrică)

Cum distribuția în caz discret sau d.d. a v.a. în caz (absolut) continuu reprezintă, de fapt, adevărate modele matematice ale unor variabile aleatoare ce prezintă interes din punct de vedere practic noi vom scoate în evidență câteva dintre cele mai răspândite modele probabiliste.

Distribuția Uniformă (caz discret).

Definiția 1. Vom spune că v.a. X este distribuită uniform pe mulțimea de valori posibile $\{0,1, 2, \dots, N\}$, pe scurt $X \sim U\{0,1, 2, \dots, N\}$, dacă probabilitățile

$$P(X = k) = \frac{1}{N+1}, \quad k = 0, 1, \dots, N,$$

numărul N având rol de parametru, unde N este un număr întreg nenegativ.

Se arată că dacă $X \sim U\{0,1, 2, \dots, N\}$, atunci $\mathbb{E}X = N/2$, $\mathbb{D}X = (N^2 - 1)/12$, $\mu_3 = 0$.

Analogic, vom spune că $X \sim U\{1, 2, \dots, N\}$, dacă $P(X = k) = \frac{1}{N}$, $k = 1, \dots, N$ sau $X \sim U\{a_1, a_2, \dots, a_N\}$, $P(X = a_k) = \frac{1}{N}$, $k = 1, \dots, N$, unde $a_1, a_2, \dots, a_N \in \mathbb{R}$, $a_1 < a_2 < \dots < a_N$, iar parametrul N este un număr întreg pozitiv.

Remarca 1. Din punct de vedere matematic distribuția uniformă modelează, de exemplu, alegerea la întâmplare a unui element din multimea de N elemente diferite. Experimentele aleatoare de acest gen pot fi simulate și pe calculator deoarece orice limbaj de programare evoluat (C++, Java, etc.) are în biblioteca sa de funcții program funcția *random*, corespunzătoare cazului discret, accesarea căreia are drept rezultat generarea unui număr ales la întâmplare din mulțimea $\{0, 1, 2, \dots, N\}$. Acest model este utilizat și în Sondajele Statistice, atunci când se urmărește ca valoarea aleasă și inclusă în eșantion dintr-o Populație Statistică să fie aleasă aleator.

Exemplul 1. Presupunem ca un PC consta din N blocuri, calculatorul ieșind din funcțiune deîndată ce iese din funcțiune unul din blocuri, ieșirea simultană din funcțiune a două sau mai multe calculatoare fiind exclusă. Depanatorul de calculatoare verifică, luând la întâmplare, unul după altul câte un bloc, până când va depista blocul defectat. Cu ce este egală probabilitatea că depanatorul va depista blocul defectat la încercarea cu numărul de ordine k , $k = 1, 2, \dots, N$.

Soluție. Introducem evenimentele aleatoare $A_k = \{\text{depanatorul va depista blocul defectat la încercarea cu numărul de ordine } k\}$, $k = 1, 2, \dots, N$. Considerăm v.a. $X = \text{numărul de ordine al încercării la care va fi depistat blocul defect}$, unde $X \in \{1, 2, \dots, N\}$. Atunci,

$$P(X=1)=P(A_1)=\frac{1}{N},$$

$$P(X=2)=P(\overline{A_1} A_2)=P(\overline{A_1})P(\overline{A_1} / A_2)=\frac{N-1}{N} \frac{1}{N-1}=\frac{1}{N},$$

...

$$P(X=N)=P(\overline{A_1} \overline{A_2} \dots \overline{A_{N-1}} A_N)=P(\overline{A_1})P(\overline{A_2} / \overline{A_1})P(\overline{A_3} / \overline{A_1 A_2}) \dots P(A_N / \overline{A_1 A_2} \dots A_{N-1}) = \frac{N-1}{N} \frac{N-2}{N-1} \frac{N-3}{N-2} \dots \frac{1}{N-(N-1)} = \frac{1}{N}.$$

În concluzie, v.a. $X \sim U\{1, 2, \dots, N\}$.

Distribuția Bernoulli.

Definiția 2. Vom spune că v.a. X este distribuită Bernoulli cu parametrul p , $0 \leq p \leq 1$, pe scurt $X \sim \text{Bernoulli}(p)$, dacă mulțimea ei de valori posibile este $\{0, 1\}$, iar probabilitățile

$$P(X = 0) = 1 - p, \quad P(X = 1) = p.$$

Caracteristicile numerice de bază ale acestei v.a. sunt $\mathbb{E}X=p$, $\mathbb{D}X=p(1-p)$, $\mu_3=2p^3 - 3p^2 + p$.

Remarca 2. Din punct de vedere matematic distribuția Bernoulli modelează comportamentul probabilist al numărului de "succese" într-o singură probă Bernoulli, aceasta din urmă noțiune fiind introdusă în

Definiția 3. Vom numi *probă Bernoulli* orice experiment aleator ce posedă proprietățile:

1. Spațiul de evenimente elementare $\Omega = \{ "succes", "insucces" \}$;
2. Probabilitatea $p = P\{ "succes" \}$, $0 \leq p \leq 1$, nu variază de la probă la probă;
3. Rezultatele unei probe nu influențează rezultatele celorlalte probe (independența probelor).

Drept exemplu de probă Bernoulli poate servi aruncarea unei monede o singură dată, "succes" considerând $\{aparitia\}$ stemei}. În genere, orice experiment aleator poate fi considerat probă Bernoulli de îndată ce acesta poate fi repetat de fiecare dată independent de celelalte repetări, iar "succes" fiind considerat orice eveniment A asociat experimentului, unde $p=P\{ "succes" \}=P(A)$. Observăm, astfel, că numărul X de "succese" într-o singură probă Bernoulli are distribuția Bernoulli. De altfel, observăm că pentru $p = \frac{1}{2}$ distribuția *Bernoulli*(p) coincide cu distribuția $U\{0,1\}$.

Exemplul 2. Pe raftul unui magazin de calculatoare sunt expuse spre vânzare 100 de laptopuri de același model, din care 3 laptopuri au defecte ascunse, dar care nimeresc sub incidența garanției. Un client cumpără, alegând la întâmplare, unul din laptopuri. Codificăm rezultatul alegerii cumpărătorului prin X , unde $X = 0$ dacă laptopul va fi returnat în urma depistării unui defect ascuns și $X = 1$ în caz contrar. Atunci v.a. $X \sim Bernoulli(p)$, unde parametrul $p=P\{laptopul\ cumpărat\ nu\ va\ avea\ defecte\ ascunse\}=0.97$.

Distribuția Binomială.

Definiția 4. Vom spune că v.a. X este distribuită Binomial cu parametrul n , p , $n \in \{1, 2, \dots\}$, $0 \leq p \leq 1$, pe scurt $X \sim Bi(n; p)$, dacă mulțimea ei de valori posibile este $\{0, 1, 2, \dots, n\}$, iar probabilitățile

$$P_n(k) = P(X = k) = \mathbb{C}_n^k p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \dots, n.$$

Caracteristicile numerice de bază ale acestei v.a. sunt $\mathbb{E}X=np$, $\mathbb{D}X=np(1-p)$, $\mu_3=np(1-p)(1-2p)$.

Pentru $n = 1$ avem că $Bi(1; p) \equiv Bernoulli(p)$.

Remarca 3. Din punct de vedere matematic distribuția binomială modelează comportamentul probabilist al numărului total de "succese" în n probe Bernoulli deoarece este valabilă

Propoziția 1. Numărul total de "succese" în n probe Bernoulli cu probabilitatea "succesului" p în fiecare probă este o v.a. $X \sim Bi(n; p)$.

Exemplul 3. În baza înregistrărilor de la casele de marcaj a centrului comercial Mall-Dova s-a constatat ca doar 30% din cumpărători își achită cumpărăturile cu cardul de credit, restul sub formă de cash. Cu ce este egală probabilitatea că din următorii 5 cumpărători 3 își vor achita cumpărăturile cu cardul de credit.

Soluție. Modul de plata al fiecărui cumpărător poate fi interpretat ca rezultat al unei probe Bernoulli: plata cu cardul="succes", plata cash="insucces", unde probabilitatea $p = P\{\text{"succes"}\} = 30/100 = 0.3$. Prin

urmare numărul platitorilor cu cardul din numărul total de $n = 5$ cumpărători este o v.a. $X \sim Bi(5; 0.3)$. Deci $P\{\text{din următorii 5 cumpărători 3 își vor achita cumpărăturile cu cardul de credit}\} = P(X = 3) = \mathbb{C}_5^3 (0.3)^3 (1 - 0.3)^{5-3} = 0.1323$.

Distribuția Geometrică.

Definiția 5. Vom spune că v.a. X este distribuită Geometric cu parametrul p , $0 < p < 1$, pe scurt $X \sim Geom(p)$, dacă mulțimea ei de valori posibile este $\{0, 1, 2, \dots, n, \dots\}$, iar probabilitățile

$$p_k = P(X = k) = p(1 - p)^k, \quad k = 0, 1, 2, \dots$$

sau dacă mulțimea ei de valori posibile este $\{1, 2, \dots, n, \dots\}$, iar probabilitățile

$$P(X = k) = p(1 - p)^{k-1}, \quad k = 1, 2, \dots,$$

ultima fiind numită și distribuție geometrică trunchiată în zero.

Dealtfel, dacă v.a. $X \sim Geom(p)$ trunchiată în zero, atunci v.a. $X - 1 \sim Geom(p)$.

Caracteristicile numerice de bază ale v.a. $X \sim Geom(p)$ trunchiată în zero sunt $\mathbb{E}X = 1/p$, $\mathbb{D}X = (1 - p)/p^2$, $\mu_3 = (1 - p)(2 - p)/p^3$.

Distribuția geometrică (trunchiată în zero) este unica v.a. de tip discret ce posedă proprietatea remarcabilă enunțată în

Propoziția 2 (Proprietatea "lipsei memoriei" a distribuției geometrice). V.a. X de tip discret, pentru care mulțimea de valori posibile este

multimea numerelor naturale, este distribuită geometric trunchiată în zero, adică

$$P(X = k) = p(1 - p)^{k-1}, k = 1, 2, \dots,$$

dacă și numai dacă aceasta posedă proprietatea "lipsei memoriei":

$$P(X = n + k / X > n) = P(X = k), \text{ pentru orice } n, k \geq 1.$$

Consecință. Dacă v.a. $X \sim \text{Geom}(p)$ trunchiată în zero, atunci

$$P(X > n + k / X > n) = P(X > k), \text{ pentru orice } n, k \geq 1.$$

Remarca 4. Din punct de vedere matematic distribuția Geometrică apare în contextul experimentului aleator ce constă în repetarea unei probe Bernoulli cu probabilitatea "succesului" p până la prima apariție a "succesului" așa cum se vede din

Propoziția 3. Numărul total de probe Bernoulli, cu probabilitatea "succesului" p în fiecare probă, efectuate până la prima apariție a "succesului" este o v.a. $X \sim \text{Geom}(p)$ trunchiată în zero, iar numărul total de "insuccese" înregistrate în același experiment este o v.a. $Y \sim \text{Geom}(p)$, unde $0 < p < 1$.

Exemplul 4 (Continuare la exemplul 3). În condițiile enunțate anterior, ne interesează cu ce este egală probabilitatea că de la începutul următoarelor zile de lucru a centrului comercial până la prima apariție a unui cumpărător plătitor cu cardul de credit va fi înregistrat un număr mai mare decât 10 cumpărători ce-și vor face cumpărături? Dar probabilitatea că, știind că primii, cel puțin, 10 cumpărători și-au achitat cumpăraturile cash, după aceasta, până la prima apariție a unui cumpărător plătitor cu cardul de credit, cel puțin, încă 10 cumpărători își vor achita cumpăraturile cash?

Soluție. De această dată, conform Propoziției 3, numărul de cumpărători înregistrați până la apariția primului cumpărător plătitor cu cardul este o v.a. $X \sim \text{Geom}(0.3)$ trunchiată în zero, iar răspunsul la prima întrebare se reduce la calcularea

$$P(X > 10) = P(X = 11) + P(X = 12) + \dots = 0.3(0.7)^{10} + 0.3(0.7)^{11} + \dots =$$

$$0.3 \cdot 0.7^{10} [1 + 0.7 + 0.7^2 + \dots] = 0.3 \cdot 0.7^{10} \cdot \frac{1}{1 - 0.7} = 0.028248.$$

Răspunsul la cea de a doua întrebare vizează calcularea probabilității condiționate $P(X > 10 + 10 / X > 10) = P(X > 20 / X > 10)$. Dar, conform consecinței de mai sus $P(X > 20 / X > 10) = P(X > 10) = 0.028248$.

Distribuția Poisson.

Definiția 6. Vom spune că v.a. X este distribuită Poisson cu parametrul λ , $\lambda > 0$, pe scurt $X \sim \text{Poisson}(\lambda)$, dacă mulțimea ei de valori posibile este $\{0, 1, 2, \dots, n, \dots\}$, iar probabilitățile

$$p_k = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots$$

Caracteristicile numerice de bază ale v.a. $X \sim \text{Poisson}(\lambda)$ $\mathbb{E}X = \mathbb{D}X = \mu_3 = \lambda$. Dacă λ este număr întreg, atunci p_k își atinge valoarea maximă pentru $k_0 = \lambda$ și $k_0 = \lambda + 1$. Dacă λ este fracționar, atunci $\text{Mod}(X) = [\lambda] + 1$.

Un context aparte în care apare distribuția Poisson este legat de noțiunea dată de

Definiția 7. Vo numi *flux de evenimente* un șir de evenimente aleatoare, care se produc în momente aleatoare de timp. Un flux de evenimente se numește *flux Poisson* dacă el are proprietățile:

a) este staționar, adică probabilitatea că într-un anume interval de timp se vor realiza exact k evenimente depinde numai de numărul k și de lungimea (durata) intervalului de timp și nu depinde de începutul lui;

b) probabilitatea realizării a k evenimente într-un anume interval de timp nu depinde de numărul de evenimente care s-au realizat înainte de începerea acestui interval;

c) realizarea a două sau mai multe evenimente într-un interval mic de timp are, practic, probabilitate nulă.

Numărul mediu de evenimente dintr-un flux Poisson care se realizează într-o unitate de timp se numește *intensitate a fluxului*. Vom nota intensitatea fluxului cu λ . Atunci are loc

Propoziția 4. Numărul total de evenimente produse pe un interval de timp de durată t într-un flux Poisson de intensitate λ , $\lambda > 0$, este o v.a. $X(t)$ distribuită Poisson cu parametrul

$$p_k(t) = P(X(t) = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, 2, \dots$$

Remarca 5. Mulțimea de v.a. $\{X(t), t \geq 0\}$, unde $X(t)$ reprezintă numărul total de evenimente produse pe un interval de timp de durată t într-un flux Poisson de intensitate λ , $\lambda > 0$ este un proces aleator ce numește proces Poisson cu parametrul λ .

Drept exemple de v.a. $X(t)$ ce apar în contextul fluxului Poisson putem, cu o anumită doză de exactitate, considera:

1. Numărului de cutremure de pământ care au loc într-o regiune seismică într-un interval de timp;
 2. Numărului de accidente rutiere produse într-un oraș, într-un interval de timp;
 3. Numărului de decese printre asigurații unei companii de asigurare într-un interval de timp;
 4. Numărului de particule (alfa) emise de o substanță radioactivă într-un anumit interval de timp;
 5. Numărului de automobile care vin la o stație de alimentare cu benzină într-un interval de timp;
 6. Numărului de clienți care se adresează la un oficiu poștal într-o zi;
 7. Numărului de apeluri la un post telefonic într-un interval de timp.
- etc., etc.

Remarca 6. Din punct de vedere matematic distribuția Poisson mai modelează, de exemplu, comportamentul probabilistic ale unor astfel de v.a. ca:

1. Numărului de erori de programare comise de un programator într-un soft de o anumită lungime;
2. Numărului de bacterii descoperite într-o picătură de apă;
3. Numărului de erori de tipar care se conțin pe o pagină (sau un grup de pagini) dintr-o carte;
4. Numărului de 3 gemeni noi născuți în decurs de un an într-o țară anume;
5. Numărului de oameni dintr-o anumită țară care au depășit vârsta de 100 de ani, etc.

Printre exemplele enumerate sunt și exemple ce nimeresc sub incidența distribuției Binomiale, dar cate poate fi aproximată cu distribuția Poisson conform

Propoziția 5 (Teorema Limită Poisson). Dacă v.a. $X \sim Bi(n; p)$ și $np \xrightarrow{n \rightarrow +\infty} \lambda$, $\lambda > 0$, deîndată ce $p \rightarrow 0$, atunci probabilitățile

$$P_n(k) = P(X = k) = \mathbb{C}_n^k p^k (1-p)^{n-k} \xrightarrow{n \rightarrow +\infty} \frac{(\lambda)^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots$$

Exemplul 5. Conform normelor tehnologice la 1000 de cozonaci cu stafide sunt prevăzute 10000 de stafide. Cu ce este egală probabilitatea ca

într-un cozonac ales la întâmplare nu vom descoperi nicio stafidă? Dar care este numărul cel mai probabil de stafide care vor nimeri în cozonacul ales?

Soluție. Dacă sunt respectate normele tehnologice, atunci pentru fiecare din cele 10000 stafide probabilitatea de a nimeri în cozonacul ales ("succes") este egală cu $p = 1/1000$, astfel, avem de a face cu $n = 10000$ de probe Bernoulli cu probabilitatea succesului $1/1000$ în fiecare probă. prin urmare Numărul stafidelor numerite în cozonacul ales este o v.a. $X \sim Bi(10000; 1/1000)$. Cum n este suficient de mare, iar p suficient de mic, putem aplica propoziția 5, considerând că $\lambda = np = 10000 \cdot 1/1000 = 10$. Prin urmare probabilitatea, de exemplu, că într-un cozonac ales la întâmplare nu vom descoperi nicio stafidă este egală cu

$$\begin{aligned} P_{10000}(0) &= P(X = 0) = \mathbb{C}_{10000}^0 (1/1000)^0 (1 - 1/1000)^{10000-0} = \\ &= (999/1000)^{10000} \simeq \frac{\lambda^k}{k!} e^{-\lambda} = \frac{10^0}{0!} e^{-10} = 4.54 \times 10^{-5} \simeq 0. \end{aligned}$$

Or, dacă în urma unui control de rutină, se constată că într-un cozonac ales la întâmplare nu a fost descoperită nicio stafidă, atunci ipoteza cea mai verosimilă e că... au fost încălcate normele tehnologice anunțate mai sus. În plus, deoarece $\lambda=10$, acesta fiind un număr întreg, rezultă că pentru $k_0 = \lambda$ și $k_0 = \lambda + 1$, probabilitățile $P(X = 10) = P(X = 11)$ au valoarea cea mai mare, egală cu $\frac{10^{10}}{10!} e^{-10} = 1.2511 \times 10^{-11}$.

Distribuția Multinomială.

Este vorba despre distribuția unei v.a. n -dimensionale, $n \geq 2$, dar care pentru $n = 2$ coincide cu distribuția $Bi(n; p)$.

Definiția 8. Vom spune că v.a. $X = (X_1, X_2, \dots, X_n)$ este *multinomial distribuită cu parametrii m și $p_1, p_2, \dots, p_n > 0, p_1 + p_2 + \dots + p_n = 1$* , pe scurt $X \sim Multi(m; p_1, p_2, \dots, p_n)$, dacă mulțimea de valori posibile a vectorului X este $\mathcal{X} = \{(m_1, m_2, \dots, m_n) : m_i = \overline{0, m}, i = \overline{1, n}, m_1 + m_2 + \dots + m_n = m\}$, iar

$$\begin{aligned} p(m_1, m_2, \dots, m_n) &= P(X_1 = m_1, X_2 = m_2, \dots, X_n = m_n) = \\ &= \frac{m!}{m_1! m_2! \dots m_n!} p_1^{m_1} p_2^{m_2} \dots p_n^{m_n} \end{aligned}$$

pentru orice $m_i = \overline{0, m}, i = \overline{1, n}, m_1 + m_2 + \dots + m_n = m$.

Caracteristicile numerice marginale pentru fiecare v.a. $X_i, i=\overline{1, n}$, sunt $\mathbb{E}X_i=mp_i, \mathbb{D}X_i=mp_i(1-p_i), \mu_{3,i}=mp_i(1-p_i)(1-2p_i)$, analogia cu cazul distribuției binomiale fiind concludentă.

Remarca 7. Este util să știm că fiecare v.a. în parte $X_i \sim Bi(m; p_i), i=\overline{1, n}$. Mai mult, pentru orice subset de indici $1 \leq i_1 < i_2 < \dots < i_k, 1 \leq k \leq n$ vectorul aleator $(X_{i_1}, X_{i_2}, \dots, X_{i_k}) \sim Multi(m; p_{i_1}, p_{i_2}, \dots, p_{i_k})$. Din punct de vedere matematic distribuția multinomială descrie toate situațiile conforme cu

Propoziția 6. Fie X_i numărul de înregistrări a evenimentului $A_i, i = \overline{1, n}$, în urma a m repetări independente a experimentului aleator \mathcal{E} în care se poate produce de fiecare dată doar unul din evenimentele A_i , unde $\bigcup_{i=1}^n A_i = \Omega, A_i \cap A_j = \emptyset$, pentru orice $i \neq j, i, j = \overline{1, n}$, iar probabilitățile $p_i = P(A_i) > 0, p_1 + p_2 + \dots + p_n = 1$ sunt cunoscute. Atunci $X = (X_1, X_2, \dots, X_n)$ este o v.a. n -dimensională și $X \sim Multi(m; p_1, p_2, \dots, p_n)$.

Exemplul 6. Presupunem ca în finala campionatului mondial la șah au ajuns doi șahiști, din care primul este campionul mondial *en titre*, despre care se mai știe că, în baza partidelor de șah jucate anterior între aceștia, 20% din partide le-a câștigat primul, 20% al doilea șahist, iar restul partidelor s-au terminat la egalitate. Meciul pentru titlul de campion consta din 12 partide, care, în caz de, de acumulare unui număr egal de puncte la sfârșit campionul *en titre* își păstrează titlul. Cu ce ste egală probabilitatea că primul jucător va câștiga 4 partide, al doilea 5, rezultatul partidelor terminându-se la egalitate? Dar probabilitatea că cel de al doilea jucător va deveni campion mondial?

Soluție. Putem considera fiecare partidă de șah un experiment aleator \mathcal{E} în care se poate produce de fiecare dată doar unul din evenimentele $A_1 = \{\text{partida dată va fi câștigată de primul jucător}\}, A_2 = \{\text{partida dată va fi câștigată de al doilea jucător}\}, A_3 = \{\text{partida dată se va termina la egalitate}\}$, unde $\bigcup_{i=1}^3 A_i = \Omega, A_i \cap A_j = \emptyset$, pentru orice $i \neq j, i, j = \overline{1, 3}$, iar probabilitățile $p_1 = P(A_1) = p_2 = P(A_2) = 0.2, p_3 = P(A_3) = 0.6$. Atunci, fiind aplicabilă propoziția 6, deducem: probabilitatea că primul jucător va câștiga 4 partide, al doilea 5, rezultatul partidelor terminându-se la egalitate este egală cu

$$p(4, 5, 3) = \frac{12!}{4!5!3!} 0.2^4 0.2^4 0.6^3 = 0.015328.$$

Ambii jucători, având aceeași putere de joc, pentru a-și păstra titlul de campion mondial, primul jucător este avantajat, totuși, de un meci terminat la egalitate. Dar probabilitatea ca meciul se va termina la egalitate este egală

cu

$$\begin{aligned}
 & p(6, 6, 0) + p(5, 5, 2) + p(4, 4, 4) + p(3, 3, 6) + p(2, 2, 8) + p(1, 1, 10) + p(0, 0, 12) = \\
 & \frac{12!}{6!6!0!} 0.2^6 0.2^6 0.6^0 + \frac{12!}{5!5!2!} 0.2^5 0.2^5 0.6^2 + \frac{12!}{4!4!4!} 0.2^4 0.2^4 0.6^4 + \frac{12!}{3!3!6!} 0.2^3 0.2^3 0.6^6 + \\
 & \frac{12!}{2!2!8!} 0.2^2 0.2^2 0.6^8 + \frac{12!}{1!1!10!} 0.2^1 0.2^1 0.6^{10} + \frac{12!}{0!0!12!} 0.2^0 0.2^0 0.6^{12} = 0.18121.
 \end{aligned}$$

Or, probabilitatea ca cel de al doilea jucator va deveni noul campion mondial este egala cu $(1 - 0.18121)/2 = 0.40940$, fapt ce i-a determinat pe organizatorii campionatelor mondiale la șah să renunțe la condiția enunțată în problemă, înlocuind-o cu condiția că în caz că cele 12 partide nu vor desemna un câștigător net, atunci meciul continuă pâna la prima partidă câștigată de unul din jucători, acesta și fiind declarat învingător.

Distribuția Hipergeometrică.

Definiția 9. Vom spune că v.a. X este *hipergeometric distribuită cu parametrii* N, M, n , pe scurt $X \sim \text{Hypergeom}(N, M, n)$, unde (N, M, n) : $N=1, 2, \dots$; $M=0, 1, \dots, N$; $n=1, 2, \dots, N$, dacă mulțimea de valori posibile a v.a. X este $\mathcal{X} = \{k : \min[0, n - (N - M)] \leq k \leq \min(n, M)\}$, iar

$$P_{N,M,n}(k) = P(X = k) = \frac{\mathbb{C}_M^k \mathbb{C}_{N-M}^{n-k}}{\mathbb{C}_N^n},$$

$$\min[0, n - (N - M)] \leq k \leq \min(n, M).$$

Caracteristicile numerice de bază ale v.a. X sunt: $\mathbb{E}X = n \frac{M}{N}$, $\mathbb{D}X = n \frac{M}{N} \cdot \frac{N-M}{N} \cdot \frac{N-n}{N-1}$, $\mu_3 = n \frac{M}{N} \cdot \frac{NM}{N} \cdot \frac{N-2M}{N} \cdot \frac{N-n}{N-1} \cdot \frac{N-2n}{N-2}$.

Remarca 8. Din punct de vedere matematic distribuția hipergeometrică descrie extragerea la întâmplare, fără întoarcere/repetare a n elemente dintr-o mulțime de N elemente, din care M elemente sunt de tipul 1, restul $N - M$ elemente fiind de tipul 2. Atunci, folosind definiția clasică a probabilității, se arată că numărul de elemente de tipul 1 care se vor regăsi printre cele n elemente extrase este o v.a. $X \sim \text{Hypergeom}(N, M, n)$, $(N, M, n) \in \{(N, M, n) : N=1, 2, \dots; M=0, 1, \dots, N; n=1, 2, \dots, M\}$. Menționăm că modul de extragere, fără întoarcere sau cu întoarcere, este principal. Astfel, dacă extragerea succesivă a n bile este cu întoarcere, atunci, folosind definiția clasică a probabilității, se demonstrează că numărul de elemente de tipul 1 care se vor regăsi printre cele n elemente extrase este o v.a. $X \sim \text{Bi}(n; p)$,

unde $p = P\{\text{"succes"}\} = P\{\text{elementul extras la extragerea } i \text{ va fi unul de tipul } 1\} = M/N$, pentru orice $i = \overline{1, n}$. Aceasta este esential, de exemplu, la modelarea matematica a eșantionării aleatoare în funcție de tipul de selectare/extragere (cu întoarcere sau fără întoarcere) a elementelor din populația statistică.

Exemplul 7 (Loto 6 din 49). În România se bucură de popularitate printre amatorii de jocuri de noroc Loteria Națională 6 din 49. Interes prezintă probabilitatea că, jucând cu o singură variantă, vom incasa câștigul maxim, adică vom ghici $X = 6$ numere câștigătoare. Deoarece ne aflăm în situația aplicării distribuției $X \sim \text{Hypergeom}(N, M, n)$, unde $N=49$, $M=6$, $n=6$, deducem ca

$$P(X = 6) = P_{49,6,6}(6) = \frac{\mathbb{C}_6^{6k} \mathbb{C}_{49-6}^{6-6}}{\mathbb{C}_{49}^6} = \frac{1}{\mathbb{C}_{49}^6} = \frac{1}{13\,983\,816}.$$

Distribuția Hipergeometrică Multivariată.

Distribuția Hipergeometrică Multivariată generalizează distribuția Hipergeometrică definită anterior.

Definiția 10. Vom spune că v.a. $X = (X_1, X_2, \dots, X_k)$ este *hipergeometric multivariată distribuită cu parametrii* $(N, M_1, M_2, \dots, M_k, n)$, pe scurt $X \sim \text{Hypergeom}(N, M_1, M_2, \dots, M_k, n)$, unde $N \in \{1, 2, \dots\}$, $M_i \in \{0, 1, \dots, N\}$, $i = \overline{1, k}$, $\sum_{i=1}^k M_i = N$, $n \in \{1, 2, \dots, N\}$, dacă mulțimea de valori posibile a v.a. X este $\mathcal{X} = \{(x_1, x_2, \dots, x_k) : x_i \in \{0, 1, \dots, N\}, i = \overline{1, k}, \sum_{i=1}^k x_i = n\}$, iar

$$\begin{aligned} P_{N, M_1, M_2, \dots, M_k, n}(x_1, x_2, \dots, x_k) &= \\ &= P(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k) = \frac{\prod_{i=1}^k \mathbb{C}_{M_i}^{x_i}}{\mathbb{C}_N^n}. \end{aligned}$$

Caracteristicile numerice marginale pentru fiecare v.a. $X_i, i = \overline{1, k}$, sunt $\mathbb{E}X = n \frac{M_i}{N}$, $\mathbb{D}X = n \frac{M_i}{N} \cdot \frac{N-M_i}{N} \cdot \frac{N-n}{N-1}$, $\mu_3 = n \frac{M_i}{N} \cdot \frac{NM_i}{N} \cdot \frac{N-2M_i}{N} \cdot \frac{N-n}{N-1} \cdot \frac{N-2n}{N-2}$.

Remarca 9. Este util să știm că fiecare v.a. în parte $X_i \sim \text{Hypergeom}(N, M_i, n)$, $i = \overline{1, k}$. Mai mult, pentru orice subset de indici $1 \leq i_1 < i_2 < \dots < i_j$, $1 \leq j \leq k$ vectorul aleator $(X_{i_1}, X_{i_2}, \dots, X_{i_j}) \sim \text{Hypergeom}(N, M_{i_1}, M_{i_2}, \dots, M_{i_j}, n)$.

Remarca 10. Din punct de vedere matematic distribuția hipergeometrică Multivariată descrie extragerea la întâmplare, fără întoarcere/repetare a n elemente dintr-o mulțime de N elemente, din care M_i elemente sunt de tipul i , $i = \overline{1, k}$, $\sum_{i=1}^k M_i = N$. Atunci, folosind definiția clasică a probabilității, se arată că numărul de elemente x_i de tipul i care se vor regăsi printre cele n elemente extrase, $i = \overline{1, k}$, $\sum_{i=1}^k x_i = n$ este o v.a. k -dimensională $X \sim \text{Hypergeom}(N, M_1, M_2, \dots, M_k, n)$.

Exemplul 8. Dintr-o cutie în care se află 30 de mingi de tenis, din care 5 de culoare roșie, 10 de culoare galbenă și 15 de culoare albastră au fost extrase la întâmplare, fără întoarcere 9 mingi. Calculați probabilitatea că printre mingile extrase se vor afla câte 3 mingi de fiecare culoare.

Soluție. În conformitate cu ultima Remarcă această probabilitate este egală cu

$$P_{30,5,10,15;9}(3, 3, 3) = \frac{\mathbb{C}_5^3 \cdot \mathbb{C}_{10}^3 \cdot \mathbb{C}_{15}^3}{\mathbb{C}_{30}^9} \simeq 0.0381627$$

Probleme propuse.

Acolo unde e cazul, în problemele următoare sa se foloseasca, pentru calcule numerice, facilitățile site-ului *wolframalpha.com*.

1. Presupunem că probabilitatea statistică ca un copil nou născut va fi băiat este egală cu 0.51. Se cere: 1) să se determine distribuția v.a. X care reprezintă numărul de băieți printre 10000 de copii noi născuți; 2) să se calculeze probabilitatea ca printre 10000 de copii noi născuți numărul băieților va fi cuprins între 4500 și 5500.

2. Un țintaș trage în ținta până când va nimeri în ea, probabilitatea de a nimeri fiind egală cu 0.1. Se cere: 1) să se determine distribuția v.a. X care reprezintă numărul total de trageri în ținta; 2) să se calculeze probabilitatea ca v.a. X va depăși valoarea 15.

3. Considerăm în calitate de v.a. X numărul de bile roșii înregistrate printre 10 bile extrase la întâmplare, una câte una, cu întoarcere, dintr-o cutie cu 2 bile albe și 3 bile roșii. Se cere: 1) să se determine distribuția v.a. X ; 2) să se calculeze probabilitatea ca v.a. X va depăși valoarea 1.

4. Considerăm în calitate de v.a. X numărul de bile roșii înregistrate printre 10 bile extrase la întâmplare, una câte una, fără întoarcere, dintr-o cutie cu 200 bile albe și 300 bile roșii. Se cere: 1) să se determine repartiția v.a. X ; 2) să se calculeze probabilitatea ca v.a. X va depăși valoarea 1.

5. Presupunem ca avem o sală de laborator în care avem 30 de calculatoare, din care 5 sunt nefuncționale, în care fiecare student, al unei grupe de 20 de studenți, ocupă la întâmplare câte un calculator. Considerăm în calitate de v.a. X numărul studentilor care vor ocupa calculatoare nefuncționale. Se cere: 1) să se determine distribuția v.a. X ; 2) să se calculeze probabilitatea ca v.a. X va varia între valorile 1 și 4.

6. Considerăm aruncarea unui zar "perfect" de 10000 de ori succesiv, în calitate de v.a. X luând numărul de câte ori apare fața 6. Se cere: 1) să se determine distribuția v.a. X ; 2) să se calculeze probabilitatea ca v.a. X va varia între valorile 100 și 300.

7. Considerăm în calitate de v.a. X numărul de aruncări nereușite ale unui zar "perfect" până la prima apariție a unui număr de puncte multiplu lui 3. Se cere: 1) să se determine distribuția v.a. X ; 2) să se calculeze probabilitatea ca v.a. X va varia între valorile 100 și 300.

8. Considerăm în calitate de v.a. X numărul de peștișori roșii înregistrați printre cei 100 de peștișori scoși la întâmplare, fără întoarcere, dintr-un acvariu enorm cu 1000 de peștișori, din care doar 10 sunt de culoare roșie. Se cere: 1) să se determine distribuția v.a. X ; 2) să se calculeze probabilitatea ca v.a. X va depăși valoarea 2.

9. Considerăm în calitate de v.a. X numărul de PC-uri de același tip supuse unui control, unul câte unul, controlul terminându-se în momentul depistării unui calculator defect. Știind că un PC ales la întâmplare va fi defect este egală cu $p = 0.05$, se cere: 1) să se determine distribuția v.a. X ; 2) să se calculeze probabilitatea că v.a. X va depăși valoarea 1.

10. Considerăm în calitate de v.a. X numărul de numere ghicite, jucând cu o singură variantă la "LOTOSPORT 5 din 35". Scrieți distribuția v.a. X . Care este probabilitatea că, jucând cu o singură variantă la LOTOSPORT "5 din 35", nu vom fi în pierdere, adică vom câștiga ceva?

11. Considerăm în calitate de v.a. X numărul format în felul următor: dintr-o cutie cu 9 bile numerotate de la 1 la 9 sunt extrase la întâmplare, succesiv, cu întoarcere, 2 bile, formând astfel un număr din două cifre, prima cifră fiind numărul primei bile, iar cea de a doua cifră, fiind numărul celei de a doua bile extrase. 1) Determinați repartiția v.a. X ; 2) Calculați probabilitatea că numărul X nu va întrece valoarea 50.

12. Un dispozitiv electronic constă din 1000 de elemente, astfel încât fiecare element poate ieși din funcțiune, independent unul de celelalte, pe o durată de timp T , cu una și aceeași probabilitate $p = 5 \cdot 10^{-4}$. Notăm prin X numărul total de elemente ieșite din funcțiune pe durata de timp T . 1)

Aflati repartitia v.a. X ; 2) Calculati probabilitatea ca pe durata T va iesi din functiune, cel putin, un element.

13. Consideram in calitate de v.a. X numărul format in felul următor: dintr-o cutie cu 9 bile numerotate de la 1 la 9 sunt extrase la întâmplare, succesiv, fără întoarcere, 2 bile, formând astfel un număr din două cifre, prima cifră fiind numărul primei bile, iar cea de a doua cifră, fiind numărului celei de a doua bile extrase. 1) Determinați distribuția v.a. X ; 2) Calculati probabilitatea ca numărul X nu va întrece valoarea 90.

14. Considerăm in calitate de v.a. X numărul total de aruncări a două zaruri "perfecte " până când, prima data, *suma* punctelor apărute va fi un număr multiplu lui 3. 1) Aflat distribuția v.a. X ; 2) Calculați probabilitatea că valoarea lui X va depăși 1000.

15. Considerăm in calitate de v.a. X numărul total de "succese" la aruncarea a doua zaruri "perfecte " de 5000 de ori, considerând "succes" la fiecare aruncare faptul ca *produsul* punctelor apărute va fi un numar multiplu lui 3. 1) Aflat distribuția v.a. X ; 2) Calculati probabilitatea că valoarea lui X va depăși 1000.

16. Considerăm in calitate de v.a. X numărul total de aruncări a două zaruri "perfecte " până când, prima dată, *minimumul* punctelor apărute va fi un număr multiplu lui 3. 1) Aflați distribuția v.a. X ; 2) Calculati probabilitatea că valoarea lui X va depăși 1000.

17. Considerăm in calitate de v.a. X numărul total de "succese" la aruncarea a două zaruri "perfecte " de 5000 de ori, considerand "succes" la fiecare aruncare faptul că *maximumul* punctelor apărute va fi un număr multiplu lui 3. 1) Aflat distribuția v.a. X ; 2) Calculati probabilitatea că valoarea lui X va depăși 1000.

18. Considerăm în calitate de v.a. X numărul total de elevi care au fost depistați ca fiind infectați TBC, dintr-un număr total de 1000 de elevi supuși controlului medical , dacă se știe că de această infecție este afectată 0.1% din populație. 1) Aflat repartitia v.a. X ; 2) Calculați probabilitatea că valoarea lui X va depăși 100.

19. Presupunem ca un an va fi secetos, independent de anii precedenți, cu probabilitatea 0.05. Notam prin X numărul de ani nesecetoși înregistrați pana la înregistrarea pentru prima dată a unui an secetos. 1) Aflat distribuția v.a. X ; 2) Calculati probabilitatea ca valoarea lui X va depasi 100.

20. O companie de asigurări are incheiate 10000 de polițe de asigurare anuală de tip RCA. Considerăm in calitate de v.a. X numărul total de despăgubiri pe care le va plăti compania pe parcursul unui an. Știind că

probabilitatea unei avarii auto este egală cu $p = 0.001 : 1$) Aflați distribuția v.a. X ; 2) Calculați probabilitatea că valoarea lui X va depăși 100.

21. Numărul X de particule alfa emise de un gram de substanță radioactivă într-o secundă este o v.a. cu distribuția Poisson cu parametrul a , unde a este numărul mediu de particule alfa emise într-o secundă. 1) Să se determine distribuția v.a. X . 2) Să se calculeze probabilitățile evenimentelor: $A = \{\text{într-o secundă vor fi emise nu mai mult de două particule alfa}\}$ și $B = \{\text{într-o secundă vor fi emise cinci particule alfa}\}$, $C = \{\text{într-o secundă vor fi emise mai mult de zece particule alfa}\}$. Care este numărul de particule alfa care corespunde celei mai mari probabilități? Să se considere că $a = 0,25$.

22. Numărul X de apeluri telefonice înregistrate la Urgența Medicală în 24 de ore este o v.a. cu distribuția Poisson cu parametrul a , unde a este numărul mediu de apeluri telefonice înregistrate în 24 de ore. 1) Să se determine distribuția v.a. X . 2) Să se calculeze probabilitățile evenimentelor: $A = \{\text{în 24 de ore fi înregistrate cel mult 100 de apeluri telefonice}\}$ și $B = \{\text{în 24 de ore vor fi înregistrate exact 100 de apeluri telefonice}\}$, $C = \{\text{în 24 de ore vor fi înregistrate mai mult de 1000 de apeluri telefonice}\}$. Aflați numărul de apeluri telefonice înregistrate în 24 de ore, care corespunde celei mai mari probabilități? Să se considere că $a = 100$.

23. Numărul X de pene de curent înregistrate la un Centru de Calcul pe parcursul unui an este o v.a. cu distribuția Poisson cu parametrul a , unde a este numărul mediu de pene de curent electric înregistrate pe parcursul unui an. 1) Să se determine distribuția v.a. X . 2) Să se calculeze probabilitățile evenimentelor: $A = \{\text{într-un an vor fi înregistrate cel mult 10 pene de curent}\}$ și $B = \{\text{într-un an vor fi înregistrate exact 10 pene de curent}\}$, $C = \{\text{într-un an vor fi înregistrate mai mult de 100 pene de curent}\}$. Aflați numărul de pene de curent înregistrate într-un an, care corespunde celei mai mari probabilități? Să se considere că $a = 10$.

24. Numărul X de erori de programare comise de un programator la elaborarea unui soft de proporții este o v.a. cu distribuția Poisson cu parametrul a , unde a este numărul mediu de erori de programare comise de programator la elaborarea unui astfel de soft. 1) Să se determine distribuția v.a. X . 2) Să se calculeze probabilitățile evenimentelor: $A = \{\text{vor fi comise cel mult 10 erori de programare}\}$ și $B = \{\text{vor fi comise exact 10 erori de programare}\}$, $C = \{\text{numărul de erori de programare comise va varia între 100 și 500}\}$. Aflați numărul de erori de programare comise la elaborarea unui soft de proporții, care corespunde celei mai mari probabilități? Să se considere că $a = 50$.

25. Numărul X de bacterii descoperite într-o picătură de apă luată dintr-un lac este o v.a. cu distribuția Poisson cu parametrul a , unde a este numărul mediu de bacterii descoperite într-o picătură de apă. 1) Să se determine distribuția v.a. X . 2) Să se calculeze probabilitățile evenimentelor: $A = \{ \text{vor fi descoperite cel mult 10 bacterii} \}$ și $B = \{ \text{vor fi descoperite exact 10 bacterii} \}$, $C = \{ \text{numarul de bacterii va varia între 500 și 5000} \}$. Aflați numărul de erori de programare comise la elaborarea unui soft de proporții, care corespunde celei mai mari probabilități? Să se considere că $a = 1000$.

26. Numărul de accidente aviatice înregistrate anual la o anumită companie aviatică este o v.a. cu distribuția Poisson cu parametrul a , unde a este numărul mediu de accidente aviatice înregistrate anual. 1) Să se determine repartiția v.a. X . 2) Să se calculeze probabilitățile evenimentelor: $A = \{ \text{într-un an vor fi înregistrate cel mult 5 accidente} \}$ și $B = \{ \text{într-un an vor fi înregistrate exact 5 accidente} \}$, $C = \{ \text{numarul de accidente înregistrate într-un an va varia între 10 și 20 accidente} \}$. Aflați numărul de accidente înregistrate într-un an, care corespunde celei mai mari probabilități? Să se considere că $a = 0.1$.

27. Numărul de stafide descoperite într-un cozonac cu stafide ales la întâmplare dintr-o partida de astfel de cozonaci este o v.a. cu repartiția Poisson cu parametrul a , unde a este numărul mediu de stafide care pot fi descoperite într-un astfel de cozonac. 1) Să se determine distribuția v.a. X . 2) Să se calculeze probabilitățile evenimentelor: $A = \{ \text{vor fi descoperite cel mult 10 stafide} \}$ și $B = \{ \text{vor fi descoperite exact 10 stafide} \}$, $C = \{ \text{numărul de bacterii va varia între 500 și 1000} \}$. Aflați numărul de stafide descoperite într-un cozonac cu stafide ales la întâmplare, care corespunde celei mai mari probabilități? Să se considere că $a = 10$.

28. Numărul de grindine înregistrate per m^2 în timpul unei ploii cu grindină este o v.a. cu distribuția Poisson cu parametrul a , unde a este numărul mediu de grindine înregistrate pe m^2 . 1) Să se determine repartiția v.a. X . 2) Să se calculeze probabilitățile evenimentelor: $A = \{ \text{pe o suprafață de } 1 m^2 \text{ vor fi înregistrate cel mult 50 de grindine} \}$ și $B = \{ \text{vor fi descoperite exact 50 de grindine} \}$, $C = \{ \text{numarul de bacterii va varia între 100 și 200} \}$. Aflați numărul de stafide descoperite pe o suprafață de $1m^2$ aleasă la întâmplare, număr care corespunde celei mai mari probabilități? Să se considere că $a = 30$.

29. Numărul de accidente auto înregistrate în or. Chișinău în 24 de ore este o v.a. cu distribuția Poisson cu parametrul a , unde a este numărul mediu de accidente auto în acest interval de timp. 1) Să se determine distribuția

v.a. X . 2) Să se calculeze probabilitățile evenimentelor: $A = \{X \leq 5\}$ și $B = \{X = 5\}$, $C = \{5 \leq X \leq 10\}$. Aflați numărul de accidente auto, care corespunde celei mai mari probabilități? Să se considere că $a = 0.5$.

30. Numărul anual de decese printre asigurații unei companii de asigurare este o v.a. cu distribuția Poisson cu parametrul a , unde a este numărul mediu de decese anual. 1) Să se determine distribuția v.a. ξ . 2) Să se calculeze probabilitățile evenimentelor: $A = \{X \leq 10\}$ și $B = \{X = 10\}$, $C = \{10 \leq X \leq 20\}$. Aflați numărul anual de decese printre asigurații unei companii de asigurare, care corespunde celei mai mari probabilități? Să se considere că $a = 3$.

4.2. Distribuții probabiliste uzuale în caz (absolut) continuu (Uniformă, Exponențială, Normală, Hi-pătrat (χ^2), T -Student)

Distribuția Uniformă (caz (absolut) continuu).

Definiția 1. Vom spune că v.a. X este distribuită uniform pe segmentul $[a, b]$, pe scurt $X \sim U[a, b]$, $a, b \in \mathbb{R}$, $a < b$, dacă densitatea ei de distribuție $f(x) = \frac{1}{b-a} \cdot I_{[a,b]}(x)$.

Caracteristicile numerice de bază ale v.a. X sunt: $\mathbb{E}X = \frac{a+b}{2}$, $\mathbb{D}X = \frac{(b-a)^2}{12}$, $\mu_3 = 0$.

Remarca 1. Din punct de vedere matematic distribuția uniformă modelează, de exemplu, alegerea la întâmplare a unui punct de pe segmentul $[a, b]$. Experimentele aleatoare de acest gen pot fi simulate și pe calculator deoarece orice limbaj de programare evoluat (C++, Java, etc.) are în biblioteca sa de funcții program funcția Random, corespunzătoare cazului absolut, accesarea căreia are drept rezultat generarea unui număr ales la întâmplare din mulțimea $[0, 1]$. Dat fiind faptul că în memoria calculatorului numerele iraționale, de exemplu, sunt, din cauza lungimii limitate a cuvântului de calculator (32 sau 64 biți), approximate de numere raționale, generatorul corespunzător poartă denumirea de *Generator de Numere Pseudoaleatoare*. Acest generator în ambele sale ipostaze (discretă și (absolut) continuă) se afla la baza simulării pe calculator a oricărei legități probabiliste. De exemplu, având la bază generatorul de valori ale v.a. $X \sim U[0, 1]$, putem simula valori ale v.a. $Y \sim U[a, b]$ și viceversa. Algoritmul se desprinde din

Propoziția 1. Dacă v.a. $X \sim U[0, 1]$, atunci v.a. $Y = (b-a)X + a \sim U[a, b]$ pentru orice $a, b \in \mathbb{R}$, $a < b$. Dimpotrivă, dacă v.a. $Y \sim U[a, b]$, atunci v.a. $X = (Y - a)/(b - a) \sim U[0, 1]$.

Are loc și

Propoziția 2. *Daca v.a. $X \sim U[0, b]$, atunci și v.a. $Y = b - X \sim U[0, b]$.*

Exemplul 1. Un troleibus sosește în stație din 5 în 5 minute. Care este probabilitatea că un pasager, care vine în stație într-un moment aleator de timp, va aștepta troleibuzul cel mult 2 minute?

Soluție. Din faptul că momentul de sosire în stație este unul aleator, rezultă ca acest moment X este, de fapt, rezultatul alegerii la întâmplare a unui punct de pe un segment de lungime 5, cu alte cuvinte, $X \sim U[0, 5]$. Timpul de așteptare a următorului troleibus coincide cu v.a. $Y = b - X$, prin urmare, conform propoziției 2, $Y \sim U[0, 5]$, ceea ce înseamnă că probabilitatea că un pasager, care vine în stație într-un moment aleator de timp, va aștepta troleibuzul cel mult 2 minute coincide cu

$$P(0 \leq Y \leq 2) = \int_0^2 \frac{1}{5} dx = \frac{2}{5} = 0.4.$$

Distribuția exponențială.

Definiția 2. Vom spune că v.a. X este distribuită exponențial cu parametrul λ , $\lambda > 0$, pe scurt $X \sim \exp\{\lambda\}$, dacă mulțimea ei de valori posibile coincide cu $[0, +\infty)$, iar $f(x) = \lambda e^{-\lambda x} \cdot I_{[0, +\infty)}(x)$.

Caracteristicile numerice de bază ale v.a. X sunt: $\mathbb{E}X = \frac{1}{\lambda}$, $\mathbb{D}X = \frac{1}{\lambda^2}$, $\mu_3 = \frac{2}{\lambda^3}$. Ținând cont de legătura dintre f.d. și d.d. a v.a. X , găsim ca f.r. a unei v.a. $X \sim \exp\{\lambda\}$ este egală cu $F(x) = (1 - e^{-\lambda x}) \cdot I_{[0, +\infty)}(x)$.

Această distribuție apare în directă legătură cu fluxurile (procesele Poisson), legatură evidențiată în

Propoziția 3. *Durata intervalului de timp dintre două momente succesive în care se produc două evenimente într-un flux Poisson cu intensitatea $\lambda > 0$ este o v.a. $X \sim \exp\{\lambda\}$.*

Prin urmare, în toate exemplele de flux Poisson enumerate în paragraful anterior putem invoca și distribuția exponențială.

Remarca 2. Dacă în caz discret singura distribuție care posedă proprietatea "lipsei memoriei" este distribuția geometrică, apoi în caz (absolut) continuu singura v.a. ce posedă aceasta proprietate este v.a. $X \sim \exp\{\lambda\}$. Într-adevăr, are loc

Propoziția 4 (Proprietatea "lipsei memoriei" a distribuției exponențiale). *V.a. X de tip (absolut) continuu, pentru care mulțimea de valori posibile coincide $[0, +\infty)$, este o v.a. distribuită exponențial, adică*

$$F(x) = (1 - e^{-\lambda x}) \cdot I_{[0, +\infty)}(x),$$

dacă și numai dacă aceasta posedă proprietatea "lipsei memoriei":

$$P(X \leq x + h \mid X > x) = P(X \leq h), \text{ pentru orice } h \geq 0.$$

Consecință. *Dacă v.a. $X \sim \exp\{\lambda\}$, atunci*

$$P(X > x + h \mid X > x) = P(X > h), \text{ pentru orice } h \geq 0.$$

Proprietatea "lipsei memoriei" numită și "proprietate markoviană" face ca această distribuție să fie folosită pe larg acolo unde apare noțiunea de durată (a vieții): în Demografie, Teoria Fiabilității, Teoria Așteptării, Actuarial, etc. A se vedea, în acest sens, exemplele 9 din p.1.3 și 2 din p.2.2.

Exemplul 2. Timpul de așteptare la o stație de alimentare cu carburanți este o v.a. X cu o durată medie de așteptare egală cu T_0 . Să se afle probabilitățile următoarelor evenimente: $A = \{\frac{1}{2}T_0 < X \leq \frac{3}{2}T_0\}$, $B = \{X > 2T_0\}$.

Soluție. Cum $X \sim \exp\{\lambda\}$ iar $\mathbb{E}X = \frac{1}{\lambda} = T_0$, rezultă că $X \sim \exp\{\frac{1}{T_0}\}$. Prin urmare

$$\begin{aligned} P(A) &= (1 - \exp\{-\frac{1}{T_0} \cdot \frac{3}{2}T_0\}) - (1 - \exp\{-\frac{1}{T_0} \cdot \frac{1}{2}T_0\}) = \\ &= \exp\{-\frac{1}{2}\} - \exp\{-\frac{3}{2}\} = 0.3834, \\ P(B) &= 1 - (1 - \exp\{-\frac{1}{T_0} \cdot 2T_0\}) = \exp\{-2\} = 0.13534. \end{aligned}$$

Distribuția normală.

Definiția 3. Vom spune că v.a. X este distribuită normal cu parametrul m și σ , $m \in \mathbb{R}$, $\sigma > 0$, pe scurt $X \sim N(m; \sigma)$, dacă mulțimea ei de valori posibile coincide cu $(-\infty, +\infty)$ iar $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$. Atunci când $m=0$, $\sigma=1$ vom spune că X este standard normal distribuită sau Gauss distribuită, pe scurt, $X \sim N(0; 1)$ (vezi exemplul 8, p.3.1.)

Caracteristicile numerice de bază ale v.a. X sunt: $\mathbb{E}X=m$, $\mathbb{D}X=\sigma^2$, $\mu_3=0$. Putem, astfel, vorbi că $X \sim N(m; \sigma)$ este o v.a. normal distribuită cu media m și abaterea standard σ .

F.d. a v.a. $X \sim N(m; \sigma)$ este funcția $F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(u-m)^2}{2\sigma^2}} du$ iar în cazul când X este standard normal distribuită f.d. este $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du$

și d.d. $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. De remarcat faptul, ca valorile f.d. $\Phi(x)$, deci și $F(x)$ nu pot fi determinate decât prin metode numerice aproximative deoarece integrala $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du$ nu se calculează exact, dar în cadrul aplicațiilor concrete putem utiliza Tabelul valorilor f.d. $\Phi(x)$ (Vezi Anexa 1) sau calculatorul online pentru distribuția normală cu media și abaterea standard cunoscute, ce poate fi accesat la adresa, de exemplu,

[http : //onlinestatbook.com/2/calculators/normal_dist.html](http://onlinestatbook.com/2/calculators/normal_dist.html) .

Important e că Tabelul valorilor f.d. $\Phi(x)$ poate fi utilizat și la calcularea valorilor f.d. $F(x)$, conform formulei $F(x) = \Phi(\frac{x-m}{\sigma})$, dacă $F(x)$ corespunde distribuției $N(m; \sigma)$. Aceasta din urmă rezulta din

Propoziția 5. Dacă v.a. $X \sim N(0; 1)$, atunci v.a. $Y = \sigma X + m \sim N(m; \sigma)$. Invers, dacă v.a. $Y \sim N(m; \sigma)$, atunci v.a. $X = \frac{Y-m}{\sigma} \sim N(0; 1)$.

Transformarea v.a. $Y \sim N(m; \sigma)$ în v.a. $\frac{Y-m}{\sigma}$ se numește *standardizare sau normare*.

Propoziția 6. F.d. $\Phi(x)$ a v.a. $X \sim N(0; 1)$ posedă următoarele proprietăți: a) $\Phi(-x) + \Phi(x) = 1$ pentru orice $x \in \mathbb{R}$; b) $P(-3 \leq X \leq 3) = \Phi(3) - \Phi(-3) = 0.9973$; c) dacă $Y \sim N(m; \sigma)$, atunci $P(\alpha \leq Y \leq \beta) = P(\frac{\alpha-m}{\sigma} \leq \frac{Y-m}{\sigma} \leq \frac{\beta-m}{\sigma}) = \Phi(\frac{\beta-m}{\sigma}) - \Phi(\frac{\alpha-m}{\sigma})$ pentru orice $\alpha, \beta \in \mathbb{R}$, $\alpha \leq \beta$.

Proprietatea a) arată că Tabelul valorilor f.d. $\Phi(x)$ poate fi definit doar pentru valorile argumentului $x \geq 0$, iar proprietatea b) justifică de ce Tabelul din Anexa 1 este dat doar pentru valorile $-3.99 < x < 3.99$. Mai mult, din proprietatea c) deducem următoarea

Consecință (Regula "3 σ "). Pentru orice v.a. $Y \sim N(m; \sigma)$, indiferent de valoarea ei medie m și abaterea ei standard σ , probabilitatea

$$P(m - 3\sigma \leq Y \leq m + 3\sigma) = 0.9973.$$

Cu alte cuvinte, în peste 99% din cazuri, valorile v.a. $Y \sim N(m; \sigma)$ se vor situa în jurul valorii ei medii la o distanță de 3σ .

O alta proprietate utilă ne-o oferă

Propoziția 7. Dacă X_1, X_2, \dots, X_n sunt v.a.i.i.d. Normal cu valoarea medie $\mathbb{E}X_k = m$ și dispersia $\mathbb{D}X_k = \sigma^2$, atunci v.a. $Y = X_1 + X_2 + \dots + X_n \sim N(nm; \sqrt{n}\sigma)$.

Remarca 3. Exemple de v.a. distribuite normal sunt: cantitatea anuală de precipitații atmosferice dintr-o anumită regiune, eroarea care se obține în urma măsurării unei mărimi cu un aparat de măsurat, înălțimea unui bărbat luat la întâmplare, etc. Mai mult, în anumite condiții, sume de v.a. într-un număr suficient de mare pot fi considerate, cu o anumită doză de aproximare, ca fiind v.a. normal distribuite, ceea ce simplifică esențial analiza statistica a datelor.

Exemplul 3. În cadrul planificării producției de costume pentru bărbați destinate comercializării în Republica Moldova, Fabrica de Confecții "Ionel" din Chișinău urmează să determine cota parte de costume produse pentru bărbații ce au înălțimea ce variază între 164 și 170 cm., știind doar faptul că înălțimea unui bărbat ales la întâmplare din regiunea dată este o v.a. X normal distribuită cu media $m=170\text{cm}$ și abaterea standard $\sigma = 5\text{cm}$. Evident că această cotă exprimată în procente este egală cu $P(164 \leq X \leq 170) \cdot 100\%$. Dar

$$P(164 \leq X \leq 170) = \Phi\left(\frac{170 - 170}{5}\right) - \Phi\left(\frac{164 - 170}{5}\right) =$$

$$\Phi(0) - \Phi(-1.2) = 0.5 - 0.11505 = 0.38495.$$

Prin urmare, cota parte în cauză este egală cu aproximativ 38.5%.

Distribuția Hi-pătrat (χ^2).

Definiția 4. Vom spune că variabila v.a. $\chi^2(n)$ este repartizată hi-pătrat cu n grade de libertate dacă aceasta are aceeași d.r. ca și v.a. $X_1^2 + X_2^2 + \dots + X_n^2$, unde X_1, X_2, \dots, X_n sunt variabile aleatoare independente, identic standard normal distribuite, sau ceea ce este echivalent, dacă aceasta are d.d.

$$f(x) = \frac{x^{(n-2)/2} \exp(-x/2)}{2^{n/2} \Gamma(n/2)} \cdot I_{[0,+\infty)}(x),$$

$\Gamma(u) = \int_0^{+\infty} x^{u-1} e^{-x} dx$ fiind $\Gamma(u) = \int_0^{+\infty} x^{u-1} e^{-x} dx$. Distribuția hi-pătrat este tabelată și în funcție de n și x , și astfel putem afla valoarea probabilității

$P(\chi^2(n) \leq x)$. Aceasta din urma poate fi calculată și online la adresa

http://onlinestatbook.com/2/calculators/normal_dist.html .

Din exemplul 3, p. 3.3. putem vedea cum variaza grafic forma d.r. a v.a. $\chi^2(n)$ in functie de valoarea parametrului n .

Caracteristicile numerice de bază ale v.a. $\chi^2(n)$ sunt: $\mathbb{E}\chi^2(n)=n$, $\mathbb{D}\chi^2(n)=2n$, $\mu_3=8n$.

Remarca 4. Această distribuție este utilizată in calitate de model matematic pentru diverse proceduri de estimare de interval sau de verificare a ipotezelor în Statistica Matematică.

Exemplul 4. Considerăm doua v.a. independente $X \sim \chi^2(15)$ și $Y \sim \chi^2(5)$. Să se determine valorile medii și dispersia v.a. $U = X + Y$, dar și valorile $u_{0.05}$ și $u_{0.95}$ pentru care $P(U \leq u_{0.05}) = 0.05$, $P(U \leq u_{0.95}) = 0.95$.

Soluție. Din definiția distribuției $\chi^2(n)$ deducem că $X \sim X_1^2 + X_2^2 + \dots + X_{15}^2$ iar $Y \sim X_{16}^2 + X_{17}^2 + \dots + X_{20}^2$, unde $X_1, X_2, \dots, X_{15}, X_{16}, X_{17}, \dots, X_{20}$ sunt v.a.i.i.d. standard normal distribuite. Prin urmare $U = X + Y \sim X_1^2 + X_2^2 + \dots + X_{15}^2 + X_{16}^2 + X_{17}^2 + \dots + X_{20}^2 \sim \chi^2(20)$. Or, $\mathbb{E}(X + Y) = \mathbb{E}\chi^2(20) = 20$, $\mathbb{D}(X + Y) = \mathbb{D}\chi^2(20) = 40$.

Cum $U \sim \chi^2(20)$, din tabelul prezentat in Anexa 2, vedem că $P(U \leq u_\alpha) = P(\chi^2(20) \leq Q(\alpha, 20)) = \alpha$, $0 < \alpha < 1$. Cu alte cuvinte $u_{0.05} = 10.851$ iar $u_{0.95} = 31.41$.

Distribuția T -Student.

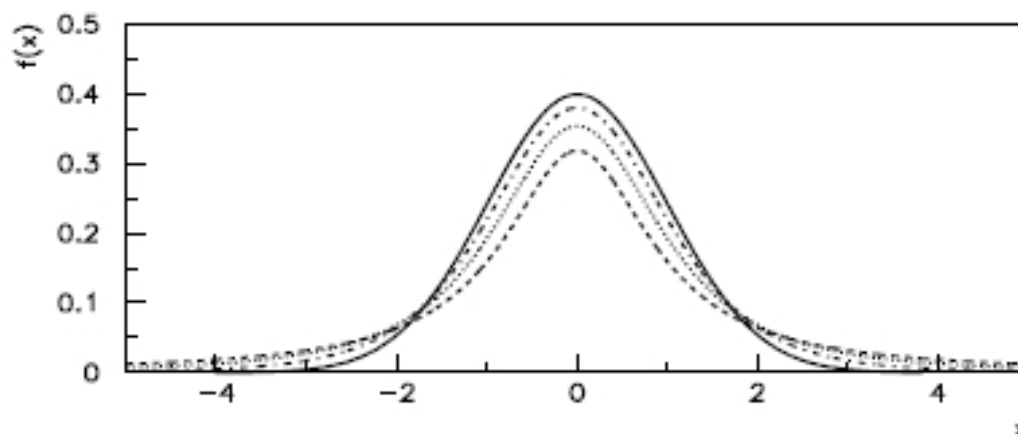
Definiția 5. Vom spune că variabila v.a. $T(n)$ este Student repartizată cu n grade de libertate, pe scurt, $T(n) \sim Student(n)$, daca aceasta are d.d. ce coincide cu d.d. a v.a. $\frac{X}{\sqrt{\chi^2(n)/n}}$, unde $X, \chi^2(n)$ sunt variabile aleatoare independente respectiv, standard normal distriuită și hi-pătrat distribuită cu n grade de libertate.

Caracteristicile numerice de bază ale v.a. $T(n)$ sunt: $\mathbb{E}T(n) = 0$, $\mathbb{D}T(n) = n/(n-2)$ pentru $n > 2$, $\mu_3=0$

La fel, aceasta distribuție se foloseste pe larg in Statistica Matematica la construirea unor estimatori de interval, dar și la verificarea ipotezelor statistice.

Pentru calculul probabilităților de forma $P(T(n) \leq x)$ se folosesc tabele care iau in calcul doar cazurile cand numarul de grade de libertate nu intrece 120, deoarece odata cu cresterea lui n aceasta distribuție se apropie de cea

normala standart, ceea ce se vede și din graficele densității de distribuție Student în funcție de n (graficul cu valoarea maximală cea mai mică pentru $n=1$, apoi $n = 2, 5, \infty$, ultimul fiind trasat cu linie îngroșată) :



Probleme propuse.

1. Înălțimea unui bărbat este o v.a. cu distribuția normală. Presupunem că această distribuția are parametrii $m = 175 \text{ cm}$ și $\sigma = 6 \text{ cm}$. Să se formeze programul de confecționate a costumelor bărbătești pentru o fabrică de confecții care se referă la asigurarea cu costume a bărbaților, înălțimile cărora aparțin intervalelor: $[150, 155)$, $[155, 160)$, $[160, 165)$, $[165, 170)$, $[170, 175)$, $[175, 180)$, $[180, 185)$, $[185, 190)$, $[190, 195)$, $[195, 200]$.

2. Presupunem că o convorbire telefonică durează în medie 5 minute și este o v.a. de distribuție exponențială. 1) Să se scrie f.dr. și să se traseze graficul ei. 2) Dacă vă apropiați de o cabină telefonică imediat după ce o persoană a intrat în ea atunci care este probabilitatea că o să așteptați nu mai mult de 2 minute?

3. Un autobus circulă regulat cu intervalul 30 minute. 1) Să se scrie d.d. a v.a.c. care reprezintă durata așteptării autobusului de către un pasager care sosete în stație într-un moment aleator de timp. 2) Să se traseze graficul d.d. 3) Să se determine f.d. și să se traseze graficul ei. 4) Care este probabilitatea că, sosind în stație, pasagerul va aștepta autobusul nu mai mult de 10 minute, unde numărul n coincide cu ultima cifra a numărului de ordine a problemei.

4. Cantitatea anuală de precipitații atmosferice are distribuția normală. Presupunem că anual, cantitatea de precipitații într-o anumită regiune este

o v.a. aleatoare de repartiție normală de parametrii $m = 500mm$ și $\sigma = 150mm$. Care este probabilitatea că în anul viitor cantitatea de precipitații va fi cuprinsă între 400 și 500 milimetri. Considerând că un an este secetos când cantitatea de precipitații nu depășește 300 mm, aflați probabilitatea că doi din viitorii zece ani vor fi secetoși?

5. Fie X o v.a. care urmează o distribuție normală cu media $m = -2$ și abaterea standard $\sigma = 3$.

- Calculați $P(-2.5 < X < 1.5)$;
- Calculați $P(-7 < X < -3.2)$;
- Calculați $P(X > 1.9)$;
- Identificați valoarea $x_{0.25}$ astfel încât $P(X \leq x_{0.25}) = 0,25$.

6. Diametrul mingilor de Ping-Pong (notat cu X) fabricate într-o fabrică mare este de așteptat să fie normal distribuite cu o medie de 3,3cm și o abatere standard de 0,6cm.

Care este probabilitatea ca o minge de Ping-Pong aleasă la întâmplare va avea un diametru

- Mai mic de 3,25cm?
- Între 3.25cm și 3.35cm?
- Să se determine cele două valori ale extremităților intervalului cu centrul în valoarea medie pentru care 60% din diametrele mingilor de Ping-Pong vor nimeri în acesta?

Dacă sunt selectate multe eșantioane aleatoare de 25 de mingi Ping-Pong

- Ce distribuție va urma media eșantionului?
- Ce proporție din eșantioane vor avea o medie mai mică de 3,25cm?
- Ce proporție din eșantioane vor avea o medie între 3.25cm și 3.35cm?
- Să se determine cele două valori ale extremităților intervalului cu centrul în valoarea medie a *mediei* diametrelor mingilor incluse în eșantion pentru care 60% din *mediile* diametrelor mingilor incluse în eșantion vor nimeri în acest interval?
- Ce este mai probabil: să apară o minge luată aparte cu diametrul între 3,25 și 3,35cm sau un eșantion cu o medie între 3,25 și 3,35cm? Explicați.

4.3. Inegalitatea Chebyshev, Legea Numerelor Mari (în formele Chebyshev, Bernoulli, Hincin), Teorema Limită Centrală

Modelele matematice ce-și găsesc aplicare în Statistica Matematică vizează, de fapt, v. a. n -dimensionale $X = (X_1, X_2, \dots, X_n)$ sau funcții de forma $g(X_1,$

X_2, \dots, X_n) privitye ca v.a., interes prezentând comportamentul lor probabilist atunci când $n \rightarrow \infty$. În acest scop sunt utile inegalități de tip Inegalitatea Chebyshev, care scot în evidență faptul că, în anumite condiții, cunoașterea doar a unor caracteristici sumare ale v.a. X , cum ar fi valoarea medie, momentul central de ordinul 2 sau dispersia, sunt suficiente pentru a evalua probabilități legate de această v.a.

Propoziția 1 (Inegalitatea Markov). *Dacă pentru v.a. X există $\mathbb{E}|X|$, atunci are loc inegalitatea*

$$P(|X| \geq \varepsilon) \leq \frac{\mathbb{E}|X|}{\varepsilon}, \forall \varepsilon > 0.$$

Consecința 1. *Dacă v.a. $X \geq 0$ și există $\mathbb{E}X$, atunci are loc inegalitatea*

$$P(X \geq \varepsilon) \leq \frac{\mathbb{E}X}{\varepsilon}, \forall \varepsilon > 0.$$

Consecința 2 (Inegalitatea Chebyshev). *Dacă pentru v.a. X există momentul inițial $\mathbb{E}X^2$, atunci are loc inegalitatea Chebyshev în formele*

$$P(|X| \geq \varepsilon) \leq \frac{\mathbb{E}X^2}{\varepsilon^2}, \forall \varepsilon > 0.$$

și

$$P(|X - \mathbb{E}X| \geq \varepsilon) \leq \frac{\mathbb{D}X}{\varepsilon^2}, \forall \varepsilon > 0.$$

Consecința 2 (Regula " $k\sigma$ "). *Dacă pentru v.a. X există momentul inițial $\mathbb{E}X^2$, atunci pentru $k=2, 3, 4$*

$$P(\mathbb{E}X - k\sigma < X < \mathbb{E}X + k\sigma) > 1 - \frac{1}{k^2}, .$$

Această consecință arată, de exemplu, pentru $k = 3$, că în cel puțin $(1 - \frac{1}{3^2})100\% \simeq 89\%$ din cazuri, valorile v.a. X se vor situa în intervalul $(\mathbb{E}X - 3\sigma, \mathbb{E}X + 3\sigma)$. Așa cum arată regula 3σ , formulată drept consecință din propoziția 6 pentru distribuția normală cu media m și abaterea standard σ , aceeași regulă dedusă din inegalitatea Chebyshev este mai puțin exactă decât prima, dar aceasta din urmă având avantajul de a fi aplicabilă pentru toate v.a. pentru care există media și dispersia.

Exemplul 1. În presupunerea că v.a. X are valoarea medie $\mathbb{E}X = 1$ și abaterea standard $\sigma = 0.2$ putem evalua probabilitățile (granițele lor de jos) evenimentelor

$$A = \{0.5 < X < 1.5\}, B = \{0.75 < X < 1.35\}, C = \{X < 2\}.$$

Într-adevăr, $A = \{1 - 0.5 < X < 1 + 0.5\} = \{|X - \mathbb{E}X| < 0.5\}$. Dar

$$P(\bar{A}) = P(|X - \mathbb{E}X| \geq 0.5) \leq \frac{\sigma^2}{0.5^2} = \frac{0.2^2}{0.5^2} = 0.16.$$

Prin urmare $P(A) = 1 - P(\bar{A}) \geq 1 - 0.16 = 0.84$. Analogic, deoarece $B \supseteq \{0.75 < X < 1.25\} = \{|X - \mathbb{E}X| < 0.25\}$, avem

$$P(B) = 1 - P(\bar{B}) \geq 1 - P(|X - \mathbb{E}X| < 0.25) \geq \frac{0.2^2}{0.25^2} = 0.64.$$

Cum $C \supseteq \{0 < X < 2\} = \{|X - \mathbb{E}X| < 1\}$, avem că

$$P(C) = 1 - P(\bar{C}) \geq 1 - P(|X - \mathbb{E}X| < 1) \geq 1 - \frac{0.2^2}{1^2} = 1 - 0.04 = 0.96.$$

Exemplul 2. Viteza vântului în $km/oră$ în orice punct al Pământului este o v.a. X . Să se evalueze granița de sus a probabilității evenimentului $\{X \geq 80 \text{ km/ora}\}$ în fiecare din următoarele cazuri: a) în urma unor măsurări facute mai mulți ani la rând s-a stabilit că $\mathbb{E}X = 16 \text{ km/oră}$; b) în urma unor măsurări facute suplimentar s-a stabilit că abaterea standard a v.a. X este egală cu $\sigma = 4 \text{ km/oră}$.

Soluție. a) Deoarece X este o v.a. nenegativa, putem aplica Inegalitatea Markov:

$$P\{X \geq 80 \text{ km/ora}\} \leq \frac{16 \text{ km/ora}}{80 \text{ km/ora}} = 0.2.$$

b) În acest caz, înainte să aplicăm Inegalitatea Cebyshev, observăm că

$$\{X \geq 80\} = \{X - \mathbb{E}X \geq 80 - \mathbb{E}X\} = \{X - \mathbb{E}X \geq 64\} \subseteq \{|X - \mathbb{E}X| \geq 64\}.$$

Prin urmare

$$P\{X \geq 80\} \leq P\{|X - \mathbb{E}X| \geq 64\} \leq \frac{\mathbb{D}X}{64^2} = \frac{4^2}{64^2} \simeq 0.004.$$

Remarca 1. Exemplul 2 este sugesitiv, în sensul că arată foarte clar efectul cunoașterii suplimentare a dispersiei v.a. X asupra marii exactității estimării probabilității necunoscute.

Definiția 1. Vom spune că șirul de v.a. $X_1, X_2, \dots, X_n, \dots$ este guvernat de Legea Numerelor Mari (LNM) dacă

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{1}{n} \sum_{k=1}^n X_k - \frac{1}{n} \sum_{k=1}^n \mathbb{E}X_k \right| < \varepsilon \right) = 1, \forall \varepsilon > 0.$$

Teorema 1 (LNM în forma Cebyshev). Dacă $X_1, X_2, \dots, X_n, \dots$ este un șir de v.a. necorelate două câte două pentru care există dispersia fiecărei v.a. și o constantă c astfel încât dispersiile $\mathbb{D}X_k \leq c, k = 0, 1, 2, \dots$, atunci șirul acesta este guvernat de Legea Numerelor Mari.

Consecința 1. Dacă $X_1, X_2, \dots, X_n, \dots$ este un șir de v.a. independente pentru care există dispersia fiecărei v.a. și o constantă c astfel încât dispersiile $\mathbb{D}X_k \leq c, k = 0, 1, 2, \dots$, atunci șirul acesta este guvernat de Legea Numerelor Mari.

Consecința 2 (LNM în forma Bernoulli). Dacă $X_1, X_2, \dots, X_n, \dots$ este un șir de v.a. i.i.d. Bernoulli cu parametrul $p, 0 < p < 1$, atunci șirul acesta este guvernat de Legea Numerelor Mari, care în acest caz ia forma

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{1}{n} \sum_{k=1}^n X_k - p \right| < \varepsilon \right) = 1, \forall \varepsilon > 0.$$

Remarca 2. LNM în forma Bernoulli servește în calitate de fundament matematic pentru Principiul Regularității (Stabilității) Statistice. Într-adevăr, cum orice experiment aleator \mathcal{E} ce poate fi repetat independent unul de altul în aceleași condiții, pentru fiecare eveniment aleator A de probabilitate nenulă $p = P(A)$, poate fi transformat în probă Bernoulli cu probabilitatea "succesului" p , rezultă că v.a.

$$X_k = \begin{cases} 1, & \text{dacă în proba cu nr. } k \text{ se produce } A, \\ 0, & \text{în caz contrar,} \end{cases} \quad k = 1, 2, \dots$$

sunt guvernate de LNM în forma Bernoulli. Dar frecvența relativă

$$f_n(A) = \frac{1}{n} \sum_{k=1}^n X_k.$$

Prin urmare pentru astfel de experimente aleatoare LNM ia forma

$$\lim_{n \rightarrow \infty} P(|f_n(A) - P(A)| < \varepsilon) = 1, \forall \varepsilon > 0,$$

care devine, din punct de vedere matematic, o expresie riguroasă a Principiului Regularității Statistice.

De remarcat ca LNM expusă mai sus în diverse forme este, totuși, condiționată de existența valorii medii și a dispersiei, dar în anumite condiții existența dispersiei nu este obligatorie, așa cum se vede din

Teorema 2 (LNM în forma Hincin). *Dacă $X_1, X_2, \dots, X_n, \dots$ este un șir de v.a.i.i.d. și pentru care există valoarea medie $\mathbb{E}X_k = m, k=1, 2, \dots$, atunci șirul acesta este guvernat de Legea Numerelor Mari, care în acest caz ia forma*

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{1}{n} \sum_{k=1}^n X_k - m\right| < \varepsilon\right) = 1, \forall \varepsilon > 0.$$

Exemplul 3. În Statistică, atunci când avem de a face cu valorile (X_1, X_2, \dots, X_n) unei v.a. X , valori obținute în urma unor observații independente asupra acestei v.a., media aritmetică $\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$ servește în calitate de estimator pentru valoarea medie (teoretică) $\mathbb{E}X = m$, m fiind necunoscut. LNM în forma Hincin garantează că, odată cu creșterea lui n , această aproximare devine tot mai bună pentru m . Din păcate, însă, LNM nu ne spune nimic despre viteza cu care acest estimator se apropie de valoarea estimată m . Aceasta lacună vine să o înlăture

Teorema 3 (Teorema Limită Centrală pentru v.a.i.i.d.). *Dacă $X_1, X_2, \dots, X_n, \dots$ este un șir de v.a.i.i.d. și pentru care există valoarea medie $\mathbb{E}X_k = m$ și dispersia $\mathbb{D}X_k = \sigma^2, k=1, 2, \dots$, atunci*

$$\lim_{n \rightarrow \infty} P\left(\frac{\sum_{k=1}^n X_k - nm}{\sigma\sqrt{n}} \leq x\right) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du, \forall x \in \mathbb{R}.$$

Consecința 1. (Teorema Limită Centrală în forma Moivre-Laplace). *Dacă $X_1, X_2, \dots, X_n, \dots$ este un șir de v.a. i.i.d. Bernoulli cu parametrul*

$p, 0 < p < 1$, atunci

$$\lim_{n \rightarrow \infty} P \left(\frac{\sum_{k=1}^n X_k - np}{\sqrt{np(1-p)}} \leq x \right) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du, \forall x \in \mathbb{R}.$$

Următoarea afirmație vine să ilustreze modul în care $\frac{1}{n} \sum_{k=1}^n X_k$ converge către valoarea medie.

Consecința 2. Dacă $X_1, X_2, \dots, X_n, \dots$ este un șir de v.a.i.i.d. și pentru care există valoarea medie $\mathbb{E}X_k = m$ și dispersia $\mathbb{D}X_k = \sigma^2, k=1, 2, \dots$, atunci:

a) pentru $\varepsilon > 0$ dat și n suficient de mare probabilitatea

$$P \left(\left| \frac{1}{n} \sum_{k=1}^n X_k - m \right| \leq \varepsilon \right) \simeq 2\Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) - 1.$$

b) pentru orice $\varepsilon > 0$ și $0 < \alpha < 1$ probabilitatea

$$P \left(\left| \frac{1}{n} \sum_{k=1}^n X_k - m \right| \leq \varepsilon \right) \geq 1 - \alpha$$

deîndată ce $n \geq n_0 = \left[\left(\frac{\sigma x_{1-\alpha/2}}{\varepsilon} \right)^2 \right] + 1$, unde $[a]$ este partea întreagă a numărului a , iar $x_{1-\alpha/2}$ este $1 - \alpha/2$ - quantilă a v.a. standard distribuită, adică

$$x_{1-\alpha/2} : \Phi(x_{1-\alpha/2}) = 1 - \alpha/2.$$

Remarca 2. Importanța Teoremi Limită Centrală rezidă, mai ales, în faptul că aceasta arată că, în anumite condiții, o sumă $S_n = \sum_{k=1}^n X_k$ de v.a.independente, luate într-un număr suficient de mare, indiferent care este distribuția lor, se comportă din punct de vedere probabilist ca și cum aceasta ar fi o v.a. normal distribuită. Mai exact, avem

Consecința 3. Dacă $X_1, X_2, \dots, X_n, \dots$ este un șir de v.a.i.i.d. și pentru care există valoarea medie $\mathbb{E}X_k = m$ și dispersia $\mathbb{D}X_k = \sigma^2, k=1, 2, \dots$, atunci pentru n suficient de mare

$$P \left(\sum_{k=1}^n X_k \leq x \right) \approx \Phi\left(\frac{x - nm}{\sigma\sqrt{n}}\right),$$

sau cu alte cuvinte $S_n = \sum_{k=1}^n X_k \simeq N(nm; \sigma\sqrt{n})$.

Exemplul 4. O Companie de asigurări are asigurate 10000 de automobile. Probabilitatea unei avarieri, ce intră sub incidența poliței de asigurare, fiind egală cu $p=0.006$. Fiecare proprietar de automobil asigurat plătește o poliță de asigurare anuală de 50 de euro, primind, în caz de avariere o sumă de 4000 de euro. Să se calculeze probabilitățile următoarelor evenimente: $A=\{\text{la sfârșit de an Compania fa fi în pierdere}\}$, $B_m=\{\text{la sfârșit de an Compania va avea un profit de, cel puțin, } m \text{ euro}\}$, $m=160000, 240000, 320000$.

Soluție. Considerăm v.a.

$$X_k = \begin{cases} 1, & \text{dacă automobilul cu nr. } k \text{ va fi avariat,} \\ 0, & \text{în caz contrar,} \end{cases}$$

$k = 1, 2, \dots, 10000$. Atunci evenimentul

$$A \equiv \left\{ 4000 \cdot \sum_{k=1}^{10000} X_k \geq 500000 \right\} = \left\{ \sum_{k=1}^{10000} X_k \geq 125 \right\}.$$

Deoarece v.a. $X_k \sim \text{Bernoulli}(0.006)$, $k=1, 2, \dots, 10000$, $n=10000$ fiind suficient de mare, putem aplica Teorema limită Centrală în forma Moivre-Laplace și Consecința 3:

$$\begin{aligned} P(A) &= P\left(\sum_{k=1}^{10000} X_k \geq 125\right) \approx 1 - \Phi\left(\frac{125 - 10000 \cdot 0.006}{\sqrt{10000 \cdot 0.006 \cdot (1 - 0.006)}}\right) = \\ &= 1 - \Phi(8.4168) = 1 - 1 = 0. \end{aligned}$$

Analogic,

$$B_{160000} \equiv \left\{ 4000 \cdot \sum_{k=1}^{10000} X_k \leq 500000 - 160000 \right\} = \left\{ \sum_{k=1}^{10000} X_k \leq 85 \right\},$$

prin urmare

$$\begin{aligned} P(B_{160000}) &= P\left\{\sum_{k=1}^{10000} X_k \leq 85\right\} \approx \Phi\left(\frac{85 - 10000 \cdot 0.006}{\sqrt{10000 \cdot 0.006 \cdot (1 - 0.006)}}\right) = \\ &= \Phi(3.2372) = 0.9987. \end{aligned}$$

Cititorul poate verifica desinestător că $P(B_{240000}) \approx 0.7413$ și $P(B_{320000}) \approx 0.1587$.

Probleme propuse.

1. Prețul de vânzare al unui anumit produs agricol la Piața Centrală variază de la zi, fiind o v.a. X cu valoare medie egală cu 2 lei/kg. Un cumpărător afirmă că probabilitatea că în ziua următoare prețul la acest produs va fi mai mare sau egal cu 10 lei/kg este mai mică sau egală cu 0.2. Este oare această afirmație fondată?

2. Producatorul de automobile Mitsubishi Colt garantează, conform cărții tehnice a acestui autoturism, un consum mediu de benzina în afara orașului egal cu 5 l/100km cu o abatere standart egală cu 0.25 l/100km. Să se estimeze valoarea probabilității că proprietarul unui atare autoturism va înregistra un consum real de benzina în afara orașului cuprins între 4.5 și 5.5 l/100km.

3. Presupunem ca durata vieții măsurată în sute de ore a memoriei hard produs pentru un anumit tip de laptop este o v.a. măsurată în sute de ore și definită de d.d. $f(x) = \lambda e^{-\lambda x} I_{[0,+\infty)}(x)$. Considerăm că v.a. X_1, X_2, \dots, X_n sunt v.a.i.i.d. cu d.d. $f(x)$. Pentru v.a. $\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$, luând $n=100$ și $\lambda=1/10$, calculați cu aproximare probabilitatea $P(\bar{X} \geq 15)$.

5. Elemente de Teoria Informației

5.1. Obiectul de studiu al Teoriei Informației

Teoria Informației (TI) studiază modele matematice (preponderant statistico-probabiliste) ale fenomenelor legate de recepționarea, stocarea și transmiterea informației. Părintele TI este Claude Shannon, anul 1948 fiind considerat anul apariției acestei teorii.

Noțiunea de informație este o noțiune primară, adică imposibil de încadrat într-o definiție strictă, la fel ca noțiunea de punct în geometrie sau noțiunea de mulțime din Teoria Mulțimilor. Încercând, totuși, să o explicăm, am putea spune că informația pentru un sistem oarecare, biologic sau tehnic, este un mesaj despre evenimente care au sau au avut sau vor avea loc atât în exteriorul sistemului cât și în interiorul lui.

Menționăm că între informație și nedeterminare există o legătură strânsă. O informație este informație, în sensul adevărat al cuvântului, dacă și numai dacă ea înlătură o anumită nedeterminare.

Într-un experiment se obține o informație numai atunci când nu cunoaștem rezultatul său înainte de efectuarea experimentului, adică numai atunci când rezultatul înlătură nedeterminarea care exista la început. Mai mult, cu cât nedeterminarea de la începutul experimentului este mai mare, cu atât mai mare va fi cantitatea de informație pe care o obținem atunci când aflăm rezultatul experimentului.

5.2. Entropia ca măsură a nedeterminării/cantității de informație

În cazul unui eveniment aleator A asociat experimentului aleator \mathcal{E} și pentru care probabilitatea sa $P(A) > 0$, Claude Shannon definește cantitatea de informație ce se conține în mesajul “s-a produs evenimentul A ” ca fiind numărul $I(A) = -\log_2 P(A)$. Cea mai mică unitate a cantității de informație se numește bit, unde $I(A) = 1 \text{ bit}$ atunci când $P(A) = 1/2$.

Evident, $I(A)$ poate fi interpretată și ca, măsura nedeterminării legate de producerea evenimentului A , această nedeterminare fiind cu atât mai mare cu cât probabilitatea $P(A)$ va fi mai mică. În extremis, gradul de nedeterminare este egal cu 0 dacă $P(A) = 1$.

Ne interesează măsura nedeterminării înlăturate sau cantitatea de informație furnizată de realizarea unui experiment aleator \mathcal{E} . Pentru aceasta presupunem că în acest experiment observatorul este interesat (înregistrează) producerea unuia din evenimentele A_1, A_2, \dots, A_n care alcătuiesc un grup complet de evenimente, adică acestea sunt disjuncte două câte două și suma lor coincide cu evenimentul sigur. Realizarea efectivă a experimentului \mathcal{E} , înregistrarea unui anumit eveniment dintre evenimentele A_1, A_2, \dots, A_n , ca rezultat al experimentului, înlătură nedeterminarea pe care am avut-o la început.

Exemplu. Fie 2 experimente aleatoare:

\mathcal{E}_1 -aruncarea unei monede “perfecte” și grupul complet de evenimente $A_1 = \{\text{aparitia stemei}\} = \{S\}$, $A_2 = \{\text{aparitia banului}\} = \{B\}$

\mathcal{E}_2 -aruncarea unei monede deformate, astfel încât $P\{S\} = 0.9$, $P\{B\} = 0.1$.

Avem, astfel, repartițiile corespunzătoare:

$$\mathcal{E}_1 : \begin{pmatrix} S & B \\ 05 & 08 \end{pmatrix}, \mathcal{E}_2 : \begin{pmatrix} S & B \\ 09 & 01 \end{pmatrix}.$$

Evident, \mathcal{E}_1 conține mai multă nedeterminare decât \mathcal{E}_2 , fapt care poate fi confirmat folosind noțiunea de entropie definite în continuare.

Fie un experiment aleator \mathcal{E} în care putem observa, în urma efectuării lui, unul din evenimentele A_1, A_2, \dots, A_n care alcătuiesc un grup complet de evenimente. Presupunem că $p_i = P(A_i) > 0, i = 1, \dots, n$.

Cantitatea de informație I furnizată în urma efectuării experimentului \mathcal{E} depinde de evenimentul A_i care se va produce, prin urmare aceasta este o v.a. data de repartiția ,

$$I : \begin{pmatrix} I(A_1) & I(A_2) & \dots & I(A_n) \\ p_1 & p_2 & \dots & p_n \end{pmatrix}, p_i > 0, i = 1, \dots, n, \quad (1)$$

Anume valoarea medie $\mathbb{E}I$ a acestei v.a. poate fi luată în calitate de măsurător al cantității de informație / gradului de nedeterminare furnizat de experimentul \mathcal{E} . Mai exact, putem da următoarea

Definiție. Vom numi *entropie* sau *măsură a nedeterminării* sau *cantitate de informație a experimentului \mathcal{E}* numărul

$$H(p_1, p_2, \dots, p_n) = - \sum_{i=1}^n p_i \cdot \log_2 p_i,$$

considerând că $p \cdot \log_2 p = 0$ dacă $p = 0$.

Revenind la exemplul de mai sus, constatăm ca entropiile lor sunt egale cu $H(\mathcal{E}_1) = 1 \text{ bit}$, $H(\mathcal{E}_2) = -(0.9 \log_2 0.9 + 0.1 \log_2 0.1) < H(\mathcal{E}_1)$ deoarece funcția $H(p) = -(p \log_2 p + (1-p) \log_2 (1-p))$ are valoare maximală ,egala cu 1, pentru $p = 0.5$. Cantitatea de informație I furnizată în urma efectuării experimentului \mathcal{E} depinde de evenimentul A_i care se va produce, prin urmare aceasta este o v.a. data de repartiția , $p_i > 0, i = 1, \dots, n$,

5.3. Proprietățile entropiei

Entropia posedă un șir de proprietăți, cele mai importante fiind prezentate în următoarele teoreme.

Teorema 1. *Entropia $H(p_1, p_2, \dots, p_n)$ ce corespunde experimentului \mathcal{E} cu repartiția (1) posedă următoarele proprietăți :*

- a) $H(p_1, p_2, \dots, p_n) \geq 0$;
- b) $H(p_1, p_2, \dots, p_n) = 0$ dacă și numai dacă există un indice i_0 astfel încât $p_{i_0} = 1$;
- c) $H(p_1, p_2, \dots, p_n) < H(1/n, 1/n, \dots, 1/n)$ pentru orice

$$(p_1, p_2, \dots, p_n), p_i \geq 0, i = 1, \dots, n, ;$$

$$d) \quad H(p_1, p_2, \dots, p_n, 0) = H(p_1, p_2, \dots, p_n).$$

Având două experimente \mathcal{E}_1 și \mathcal{E}_2 independente putem define entropia experimentului $\mathcal{E} = (\mathcal{E}_1, \mathcal{E}_2)$ ce constă în efectuarea lor simultană. Notăm entropia experimentului cu $H(\mathcal{E})$. Are loc

Teorema 2. $H(\mathcal{E}) = H(\mathcal{E}_1, \mathcal{E}_2) = H(\mathcal{E}_1) + H(\mathcal{E}_2)$.

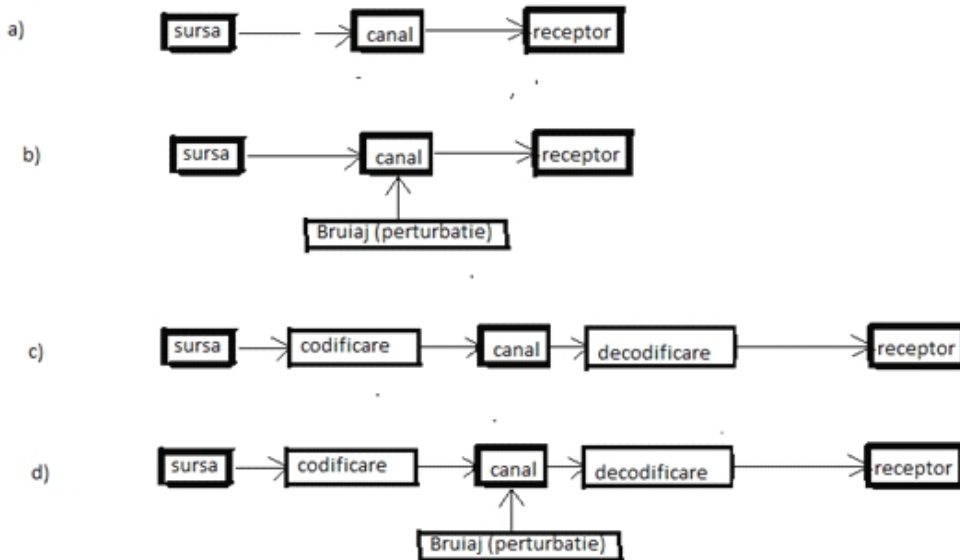
Exemplu. Considerăm în calitate de experiment aleator $\mathcal{E} = (\mathcal{E}_1, \mathcal{E}_2)$ aruncarea simultană a două monede perfecte, \mathcal{E}_1 fiind aruncarea primei monede, iar \mathcal{E}_2 aruncarea celei de a doua monede. Evident, $H(\mathcal{E}_1) = H(\mathcal{E}_2) = 1 \text{ bit}$, dar cantitatea de informație $I(\mathcal{E})$ furnizată de rezultatul efectuării experimentului \mathcal{E} este o v.a. cu repartiția

$$I : \begin{pmatrix} I(SS) & I(SB) & I(BS) & I(BB) \\ 14 & 14 & 14 & 14 \end{pmatrix}.$$

Prin urmare $H(\mathcal{E}) = H(\mathcal{E}_1, \mathcal{E}_2) = H(\mathcal{E}_1) + H(\mathcal{E}_2)$.

5.4. Transmiterea informației. Codificarea. Teoreme de codificare

Orice sistem de transmitere a informației se încadrează în una din următoarele scheme generale:



Pentru a înțelege rolul noțiunii de entropie în Teoria Informației ne vom referi în continuare, drept exemplu, la schema c) de transmitere a informației cu codificare/decodificare, dar fără bruiaj. Dar pentru început vom vedea ce reprezintă codificarea/decodificarea în sistemele de comunicație de orice natură, pornind de la următorul

Exemplu. Considerăm trimiterea unui mesaj prin sistemul SMS (system de mesaje scurte) al unui telefon mobil. Pentru aceasta folosim, initial, în calitate de semnale caracterele tastaturii QWERTY, numărul lor total fiind notat cu N . Cu alte cuvinte aceste caractere formează o mulțime $X = \{x_1, x_2, \dots, x_N\}$ de semnale. În momentul trimiterii lui (apăsării butonului “send”) fiecare semnal/character x din X este codificat digital, adică cu ajutorul unei mulțimi de semnale simple $Y = \{y_1, y_2, \dots, y_Q\} = \{0, 1\}$.

Definiție. Vom numi codificare (codare) asocierea la un anumit system de semnale care poartă o informație (mesaj) a unor succesiuni/șiruri de semnale a unui alt system de semnale. Primele se numesc semnale initiale, cele de a doua se numesc semnale simple.

Este firesc ca mulțimea Y de semnale simple să conțină mai puține semnale decât mulțimea X cea ce se și întâmplă în exemplul nostru, $Q < N$.

În Teoria Informației se folosesc, în funcție de rezultatele scontate, diferiți algoritmi de codare/decodare.

Motivele care ne conduc la necesitatea codării informației sunt multiple. Iată câteva dintre ele:

- a) Natura sistemului concret de transmitere a informației (telegraf, telefon mobil, sistemul Morse, etc.) implică un anumit mod de codare;
- b) Prezența bruiajelor (perturbațiilor) canalelor de transmitere a informației implică utilizarea codării pentru a diminua efectul nociv asupra corectitudinii informației recepționate;
- c) La fel, caracterul intim, secret al informației implică codarea ei.

Observație. Motive de tipul c) pot avea drept consecință faptul că, uneori, $\text{card}Y \geq \text{card}X$.

Fie $X = \{x_1, x_2, \dots, x_N\}$ mulțimea de semnale inițiale, iar $Y = \{y_1, y_2, \dots, y_Q\}$ mulțimea de semnale simple. În procesul alcătuirii/trimiterii unui mesaj (unei informații), folosind semnalele din X , fiecare x_i din X are o anumită probabilitate (frecvența relativă) $p(x_i)$ cu care se întâlnește în acest mesaj,

. Să luăm, drept exemplu, mesajul “*măine_va_ploua*”. Acesta are lungimea 14, incluzând simbolul “*pauză*”- “_”, iar semnalele “*m*”, “*a*” și “*b*”, luate ca exemplu, au frecvența relativă egală, respectiv, cu $p(m) = 1/14$, $p(a) = 1/7$, $p(b) = 0$. Prin urmare, trimiterii unui mesaj \mathcal{E} i se poate

asocia cantitatea de informație $H(\mathcal{E})$. Cum în procesul codării (criptării) fiecărui semnal x_i se asociază un șir de semnale simple din Y , șirul având lungimea n_i , $i = 1, 2, \dots, N$, atunci putem calcula lungimea medie L a unui șir de semnale simple folosite în procesul de codare:

Acum putem formula, drept exemplu, două rezultate teoretice ce vizează procesul de codificare a semnalelor transmise prin intermediul unui sistem/canal de transmitere a informației fără bruiaj, rezultate care arată care sunt limitele și posibilitățile codificării.

Propoziție. Pentru ca să fie posibilă codificarea prin intermediul șirurilor de semnale simple de lungimea n_i , $i = 1, 2, \dots, N$, care se fie atașate semnalelor inițiale x_1, x_2, \dots, x_N , este necesar și suficient să fie îndeplinită inegalitatea

$$\sum_{i=1}^N Q^{-n_i} < 1.$$

Teorema codificării. Ca să putem efectua o codificare, folosind Q semnale simple, pentru a transmite o cantitate de informație $H(\mathcal{E})$, este necesar ca lungimea medie L a șirurilor de codificare, atașate semnalelor inițiale din mulțimea X , purtătoare de informație, să nu fie inferioară numărului $H(\mathcal{E})\log_2 Q$, adică $L \geq H(\mathcal{E})\log_2 Q$.

Exemplu (continuare). În directă legătură cu trimiterea mesajului “*măine va ploua*”, respectiv, experimentul \mathcal{E} , cantitatea de informație I va fi o v.a. cu repartiția

$$I: \begin{array}{cccccccccccc} I(\hat{a}) & I(a) & I(e) & I(0) & I(u) & I(i) & I(m) & I(n) & \dots & I(v) & I(_) \\ 1/14 & 2/14 & 1/14 & 1/14 & 1/14 & 1/14 & 1/14 & 1/14 & \dots & 1/14 & 2/14 \end{array}$$

Prin urmare $H(\mathcal{E}) = 2[-(2/14) \cdot \log_2(2/14)] + 10[-(1/14) \cdot \log_2(1/14)] \approx 3.5216$.

În cazul codificării digitale $Y = \{0, 1\}$, $Q = 2$. Prin urmare ca acest mesaj să poată fi codificat digital este necesar ca lungimea medie a șirurilor de codificare să fie mai mare sau egală cu $H(\mathcal{E})\log_2 Q = 3.5216$.

BIBLIOGRAFIE

- [1]. Kolmogorov A.N., *Noțiunile de bază ale teoriei probabilităților*, M., Ed. Nauka, 1974 (l. rusă).
- [2]. Feller W., *An Introduction to Probability Theory and its Applications*, Volume I, 3rd edition, N.-Y., London, John Wiley & Sons Inc., 1968.
- [3]. Feller W., *An Introduction to Probability Theory and its Applications*, Volume II, 2nd edition, N.-Y., London, John Wiley & Sons Inc., 1971.
- [4]. Iosifescu M., Mihoc Gh., Theodorescu R., *Teoria probabilităților și statistica matematică*, Ed. Tehnică, București, 1966.
- [5]. Jay L. Devore, *Probability & Statistics (for Engineering and the Sciences)*, 8th edition, BROOKS / COLE, Chicago, 2009
- [6]. Mittelhammer R. C., *Mathematical statistics for economics and business*, Ed. Springer-Verlag, N.-Y. Inc., 1996.
- [7]. Silviu Guiașu, Radu Theodorescu, *Matematica și Informatica*, Edit. Stiintifică, București, 1967.

ANEXA 1

Tabelul de valori a distribuției normale standard:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du$$

x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.9	.00005	.00005	.00004	.00004	.00004	.00004	.00004	.00004	.00003	.00003
-3.8	.00007	.00007	.00007	.00006	.00006	.00006	.00006	.00005	.00005	.00005
-3.7	.00011	.00010	.00010	.00010	.00009	.00009	.00008	.00008	.00008	.00008
-3.6	.00016	.00015	.00015	.00014	.00014	.00013	.00013	.00012	.00012	.00011
-3.5	.00023	.00022	.00022	.00021	.00020	.00019	.00019	.00018	.00017	.00017
-3.4	.00034	.00032	.00031	.00030	.00029	.00028	.00027	.00026	.00025	.00024
-3.3	.00048	.00047	.00045	.00043	.00042	.00040	.00039	.00038	.00036	.00035
-3.2	.00069	.00066	.00064	.00062	.00060	.00058	.00056	.00054	.00052	.00050
-3.1	.00097	.00094	.00090	.00087	.00084	.00082	.00079	.00076	.00074	.00071
-3.0	.00135	.00131	.00126	.00122	.00118	.00114	.00111	.00107	.00104	.00100
-2.9	.00187	.00181	.00175	.00169	.00164	.00159	.00154	.00149	.00144	.00139
-2.8	.00256	.00248	.00240	.00233	.00226	.00219	.00212	.00205	.00199	.00193
-2.7	.00347	.00336	.00326	.00317	.00307	.00298	.00289	.00280	.00272	.00264
-2.6	.00466	.00453	.00440	.00427	.00415	.00402	.00391	.00379	.00368	.00357
-2.5	.00621	.00604	.00587	.00570	.00554	.00539	.00523	.00508	.00494	.00480
-2.4	.00820	.00798	.00776	.00755	.00734	.00714	.00695	.00676	.00657	.00639
-2.3	.01072	.01044	.01017	.00990	.00964	.00939	.00914	.00889	.00866	.00842
-2.2	.01390	.01355	.01321	.01287	.01255	.01222	.01191	.01160	.01130	.01101
-2.1	.01786	.01743	.01700	.01659	.01618	.01578	.01539	.01500	.01463	.01426
-2.0	.02275	.02222	.02169	.02118	.02068	.02018	.01970	.01923	.01876	.01831
-1.9	.02872	.02807	.02743	.02680	.02619	.02559	.02500	.02442	.02385	.02330
-1.8	.03593	.03515	.03438	.03362	.03288	.03216	.03144	.03074	.03005	.02938
-1.7	.04457	.04363	.04272	.04182	.04093	.04006	.03920	.03836	.03754	.03673
-1.6	.05480	.05370	.05262	.05155	.05050	.04947	.04846	.04746	.04648	.04551
-1.5	.06681	.06552	.06426	.06301	.06178	.06057	.05938	.05821	.05705	.05592
-1.4	.08076	.07927	.07780	.07636	.07493	.07353	.07215	.07078	.06944	.06811
-1.3	.09680	.09510	.09342	.09176	.09012	.08851	.08691	.08534	.08379	.08226
-1.2	.11507	.11314	.11123	.10935	.10749	.10565	.10383	.10204	.10027	.09853
-1.1	.13567	.13350	.13136	.12924	.12714	.12507	.12302	.12100	.11900	.11702
-1.0	.15866	.15625	.15386	.15151	.14917	.14686	.14457	.14231	.14007	.13786
-0.9	.18406	.18141	.17879	.17619	.17361	.17106	.16853	.16602	.16354	.16109
-0.8	.21186	.20897	.20611	.20327	.20045	.19766	.19489	.19215	.18943	.18673
-0.7	.24196	.23885	.23576	.23270	.22965	.22663	.22363	.22065	.21770	.21476

-0.6 .27425 .27093 .26763 .26435 .26109 .25785 .25463 .25143 .24825 .24510
-0.5 .30854 .30503 .30153 .29806 .29460 .29116 .28774 .28434 .28096 .27760
-0.4 .34458 .34090 .33724 .33360 .32997 .32636 .32276 .31918 .31561 .31207
-0.3 .38209 .37828 .37448 .37070 .36693 .36317 .35942 .35569 .35197 .34827
-0.2 .42074 .41683 .41294 .40905 .40517 .40129 .39743 .39358 .38974 .38591
-0.1 .46017 .45620 .45224 .44828 .44433 .44038 .43644 .43251 .42858 .42465
-0.0 .50000 .49601 .49202 .48803 .48405 .48006 .47608 .47210 .46812 .46414
0.0 .50000 .50399 .50798 .51197 .51595 .51994 .52392 .52790 .53188 .53586
0.1 .53983 .54380 .54776 .55172 .55567 .55962 .56356 .56749 .57142 .57535
0.2 .57926 .58317 .58706 .59095 .59483 .59871 .60257 .60642 .61026 .61409
0.3 .61791 .62172 .62552 .62930 .63307 .63683 .64058 .64431 .64803 .65173
0.4 .65542 .65910 .66276 .66640 .67003 .67364 .67724 .68082 .68439 .68793
0.5 .69146 .69497 .69847 .70194 .70540 .70884 .71226 .71566 .71904 .72240
0.6 .72575 .72907 .73237 .73565 .73891 .74215 .74537 .74857 .75175 .75490
0.7 .75804 .76115 .76424 .76730 .77035 .77337 .77637 .77935 .78230 .78524
0.8 .78814 .79103 .79389 .79673 .79955 .80234 .80511 .80785 .81057 .81327
0.9 .81594 .81859 .82121 .82381 .82639 .82894 .83147 .83398 .83646 .83891
1.0 .84134 .84375 .84614 .84849 .85083 .85314 .85543 .85769 .85993 .86214
1.1 .86433 .86650 .86864 .87076 .87286 .87493 .87698 .87900 .88100 .88298
1.2 .88493 .88686 .88877 .89065 .89251 .89435 .89617 .89796 .89973 .90147
1.3 .90320 .90490 .90658 .90824 .90988 .91149 .91309 .91466 .91621 .91774
1.4 .91924 .92073 .92220 .92364 .92507 .92647 .92785 .92922 .93056 .93189
1.5 .93319 .93448 .93574 .93699 .93822 .93943 .94062 .94179 .94295 .94408
1.6 .94520 .94630 .94738 .94845 .94950 .95053 .95154 .95254 .95352 .95449
1.7 .95543 .95637 .95728 .95818 .95907 .95994 .96080 .96164 .96246 .96327
1.8 .96407 .96485 .96562 .96638 .96712 .96784 .96856 .96926 .96995 .97062
1.9 .97128 .97193 .97257 .97320 .97381 .97441 .97500 .97558 .97615 .97670
2.0 .97725 .97778 .97831 .97882 .97932 .97982 .98030 .98077 .98124 .98169
2.1 .98214 .98257 .98300 .98341 .98382 .98422 .98461 .98500 .98537 .98574
2.2 .98610 .98645 .98679 .98713 .98745 .98778 .98809 .98840 .98870 .98899
2.3 .98928 .98956 .98983 .99010 .99036 .99061 .99086 .99111 .99134 .99158
2.4 .99180 .99202 .99224 .99245 .99266 .99286 .99305 .99324 .99343 .99361
2.5 .99379 .99396 .99413 .99430 .99446 .99461 .99477 .99492 .99506 .99520
2.6 .99534 .99547 .99560 .99573 .99585 .99598 .99609 .99621 .99632 .99643
2.7 .99653 .99664 .99674 .99683 .99693 .99702 .99711 .99720 .99728 .99736
2.8 .99744 .99752 .99760 .99767 .99774 .99781 .99788 .99795 .99801 .99807
2.9 .99813 .99819 .99825 .99831 .99836 .99841 .99846 .99851 .99856 .99861
3.0 .99865 .99869 .99874 .99878 .99882 .99886 .99889 .99893 .99896 .99900

3.1 .99903 .99906 .99910 .99913 .99916 .99918 .99921 .99924 .99926 .99929
3.2 .99931 .99934 .99936 .99938 .99940 .99942 .99944 .99946 .99948 .99950
3.3 .99952 .99953 .99955 .99957 .99958 .99960 .99961 .99962 .99964 .99965
3.4 .99966 .99968 .99969 .99970 .99971 .99972 .99973 .99974 .99975 .99976
3.5 .99977 .99978 .99978 .99979 .99980 .99981 .99981 .99982 .99983 .99983
3.6 .99984 .99985 .99985 .99986 .99986 .99987 .99987 .99988 .99988 .99989
3.7 .99989 .99990 .99990 .99990 .99991 .99991 .99992 .99992 .99992 .99992
3.8 .99993 .99993 .99993 .99994 .99994 .99994 .99994 .99995 .99995 .99995
3.9 .99995 .99995 .99996 .99996 .99996 .99996 .99996 .99996 .99997 .99997

ANEXA 2

Tabel pentru α -cuantilele uzuale $Q(\alpha)$, $\alpha \in (0, 1)$ ale v.a. $\chi^2(\nu)$ distribuite
 Hi-pătrat cu ν grade de libertate:
 $P(\chi^2(\nu) \leq Q(\alpha, \nu)) = \alpha$, unde pentru $\nu \geq 40$ cuantilele $Q(\alpha, \nu) \approx Q(\alpha, 40)$

ν	$Q(.005)$	$Q(.01)$	$Q(.025)$	$Q(.05)$	$Q(.1)$	$Q(.9)$	$Q(.95)$	$Q(.975)$	$Q(.99)$	$Q(.995)$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.143	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.290	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.653	40.647	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.994
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
31	14.458	15.655	17.539	19.281	21.434	41.422	44.985	48.232	52.192	55.003
32	15.134	16.362	18.291	20.072	22.271	42.585	46.194	49.480	53.486	56.328
33	15.815	17.074	19.047	20.867	23.110	43.745	47.400	50.725	54.775	57.648
34	16.501	17.789	19.806	21.664	23.952	44.903	48.602	51.966	56.061	58.964
35	17.192	18.509	20.569	22.465	24.797	46.059	49.802	53.204	57.342	60.275
36	17.887	19.233	21.336	23.269	25.643	47.212	50.998	54.437	58.619	61.581
37	18.586	19.960	22.106	24.075	26.492	48.364	52.192	55.668	59.893	62.885
38	19.289	20.691	22.878	24.884	27.343	49.513	53.384	56.896	61.163	64.183
39	19.996	21.426	23.654	25.695	28.196	50.660	54.572	58.120	62.429	65.477
40	20.707	22.164	24.433	26.509	29.051	51.805	55.759	59.342	63.691	66.767