Curs 1/Partea 1: Introducere in Teoria Probabilităţilor

1.1 Experiment aleator, evenimente şi probabilităţi

Cursul de teoria probabilitaţilor cu aplicaţii în CS[1] are ca scop dezvoltarea bazelor teoretice necesare pentru dezvoltarea şi analiza unor modele probabiliste frecvent utilizate în ingineria şi ştiinţa calculatoarelor, precum şi fundamentarea algorimilor de generare de numere aleatoare, de simulare a sistemelor ce includ componente aleatore: simularea cozilor, a serviciului CPU, a navigării aleatoare în WWW (algorimul `PageRank-Google`) etc. Bazele teoretice sunt prezentate la curs, dezvoltarea abilităţii de modelare probabilistă şi analiza modelului, se realizează la seminar şi prin exersare individuală (teme), iar însuşirea tehnicilor de simulare se realizează prin conceperea în echipă (6 studenţi) a codului `C` ce implementează algoritmi de simulare, la proiect.

Datele cu care operează teoria probabilităţilor şi statistica sunt obţinute fie prin observaţii asupra evenimentelor necontrolabile din natură, societate, fie ca rezultat al unui experiment controlat, într-un laborator sau experiment pe calculator. Pentru a avea o terminologie unică, definim un *experiment* ca fiind procesul prin care efectuăm o observaţie sau o măsurătoare.

Experimentele care pot avea rezultate diferite în funcţie de o serie de circumstanţe şi rezultatele nu pot fi cunoscute înaintea realizării experimentului se numesc experimente aleatoare.

Exemple de experimente aleatoare:

• Înregistrarea numărului de cereri de acces la un server WEB, într-un interval de timp $(0, t]$ (experimentul constă în observarea fluxului sosirii cererilor de acces);

• observarea numărului de comparaţii într-un algoritm de sortare;

• observarea timpului în care CPU răspunde la o comandă de la un terminal interactiv;

• observarea timpului de viaţă (de bună funcţionare până la prima cădere) a unei componente electronice;

Definiţia 1.1.1 *Rezultatul atomic (nedecompozabil) al unui experiment aleator se numeşte*

[1] notiţele de curs sunt adaptate conform carţii/curs prof. dr. Emilia Petrişor 2016

realizare. Colecţia tuturor realizărilor acoperă orice posibilitate (adică este exhaustivă) şi nici o realizare nu se suprapune peste o alta (realizările sunt mutual exclusive).

O colecţie de realizări se numeşte **eveniment**, *iar mulţimea tuturor realizărilor – evenimentul sigur sau spaţiul observabilelor.*

Numele de eveniment sigur este folosit doar în limba română. În limba engleză colecţia tuturor realizărilor posibile se numeşte *sample space*.

• Evenimentul sigur se produce cu certitudine la fiecare efectuare a experimentului.

• Evenimentul care nu se produce, ori de câte ori repetăm experimentul în condiţii identice, se numeşte **eveniment imposibil**.

Evenimentul sigur se notează Ω, evenimentul imposibil cu $\emptyset$, iar evenimentele particulare ce sunt părţi ale evenimentului sigur se notează cu $A, B, C, \ldots$.

Experimentul clasic, ce ilustrează aceste noţiuni, este aruncarea zarului. Realizările posibile ale experimentului sunt apariţia feţei cu numărul $1, 2, 3, 4, 5, 6$. Evenimentele de apariţie a feţei cu numărul k, $k = \overline{1,6}$, se numesc evenimente elementare. Evenimentul de apariţie a unei feţe cu număr impar este reprezentat simbolic de mulţimea $A = \{1, 2, 5\}$, în timp ce evenimentul apariţiei unei feţe cu numărul mai mare sau egal ca 3 este $B = \{3, 4, 5, 6\}$.

În orice experiment aleator, unui eveniment A îi corespunde *evenimentul contrar sau opus* notat $\complement_\Omega A$ ($\complement$ notează complementara mulţimii A faţa de Ω). Pentru simplitate vom nota evenimentul opus lui A, prin $\overline{A}$. Producerea evenimentului A înseamnă nerealizarea evenimentului contrar şi reciproc ($\complement_\Omega(\complement_\Omega(A)) = A$). Evenimentul sigur şi evenimentul imposibil sunt contrare unul altuia: $\complement_\Omega\emptyset = \Omega$.

Probabilitatea unui eveniment A este un număr notat $P(A) \in [0, 1]$, ce reprezintă şansa pe care o are evenimentul de a se produce.

În concluzie noţiunile primare în teoria probabilităţilor sunt cele de *eveniment* într-un experiment aleator şi de *probabilitate* a evenimentului.

1.2 Spaţiu discret de probabilitate

În prima parte a cursului considerăm experimente aleatoare în care numărul realizărilor este finit sau infinit numărabil, adică mulţimea realizărilor este în corespondenţa bijectivă cu mulţimea numerelor naturale.

O bijecţie $f : \mathbb{N} \to \Omega$, $f(n) = \omega_n$, indexează elementele lui Ω, adică:

$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_n, \ldots\}$$

Notăm cu $\mathcal{P}(\Omega)$ mulţimea părţilor lui Ω, adică mulţimea tuturor submulţimilor sale.

Dacă mulţimea Ω are n elemente, atunci $\mathcal{P}(\Omega)$ are 2^n elemente.

Dacă mulţimea Ω este numărabilă, atunci mulţimea părţilor sale $\mathcal{P}(\Omega)$ este infinită, dar ne-numărabilă.

Având precizat evenimentul sigur asociat unui experiment aleator, sarcina cea mai dificilă este să atribuim o probabilitate de producere fiecărui eveniment posibil în acel experiment.

Avem două cazuri în care probabilitatea unui eveniment se atribuie relativ simplu şi intuitiv:

a) Mulţimea observabilelor este finită şi toate realizările experimentului sunt egal probabile (adică nu există motiv ca o realizare să se producă mai frecvent ca alta). În acest caz *se defineşte probabilitatea unui eveniment A:*

$$P(A) = \frac{|A|}{|\Omega|}$$

ca raportul dintre numărul cazurilor favorabile şi numărul cazurilor posibile. Numărul cazurilor favorabile este numărul realizărilor a căror colecţie constituie evenimentul A. De exemplu în aruncarea zarului (presupus ca un cub "perfect") probabilitatea de apariţie a unei feţe cu număr impar este $\frac{3}{6} = \frac{1}{2}$. 6 este numărul cazurilor posibile, iar 3 numărul cazurilor favorabile.

Ori de câte ori într-un enunţ sau problemă se afirmă că se alege, se selectează la întâmplare o variantă din n, se presupune (prin convenţie) că cele n variante sunt echiprobabile, adică au aceeaşi probabilitate.

Exemplul 1. Să se calculeze probabilitatea ca alegând la întâmplare un număr de patru cifre în baza 10 acesta să fie un număr valid şi în baza 8.

Rezolvare: Mulţimea tuturor posibilităţilor (evenimentul sigur al experimentului) este mulţimea numerelor de forma:

$$\Omega = \{x = (c_3c_2c_1c_0)_{10} \,|\, c_i \in \{0, 1, 2, 3, \ldots, 9\}, c_3 \neq 0\}$$

adică mulţimea tuturor 4–listelor de elemente din $\{0, 1, 2, 3, \ldots, 9\}$, minus mulţimea 4–listelor ce încep cu 0, care este în corespondenţa bijectivă cu mulţimea 3–listelor de elemente din $\{0, 1, 2, 3, \ldots, 9\}$. Evenimentul de interes (mulţimea cazurilor favorabile) este

$$E = \{(o_3o_2o_1o_0)_8 \,|\, o_i \in \{0, 1, 2, 3, 4, 5, 6, 7\}, o_3 \neq 0\}$$

Cardinalul lui Ω este $10^4 - 10^3$, iar al lui E este $8^4 - 8^3$. Deci probabilitatea lui E este:

$$P(E) = \frac{|E|}{|\Omega|} = \frac{8^4 - 8^3}{10^4 - 10^3} = \left(\frac{4}{5}\right)^3 \frac{7}{9}$$

În majoritatea experimentelor realizările nu sunt egal probabile.

b) În cazul în care un experiment aleator are un număr finit de realizări ce nu sunt egal probabile, nu există o modalitate teoretică care să permită calculul probabilităţii cu acurateţe absolută. În inginerie, de obicei, se analizează datele existente din repetarea experimentului de n ori, în aceleaşi condiţii, şi apoi *se aproximează probabilitatea unui*

eveniment ca raportul dintre numărul cazurilor $k, 0 \leq k \leq n$, *în care evenimentul de interes s-a produs şi numărul experimentelor,* $P(A) \approx \frac{k}{n}$. Această aproximare se bazează pe ipoteza că probabilitatea exactă a evenimentului de interes este

$$p = \lim_{n\to\infty} \frac{k(n)}{n},$$

numită limita frecvenţelor experimentale de producere a evenimentului.

Exemplul 2. *Caching* înseamnă stocarea datelor într-un sistem de stocare ce poate fi rapid accesat sau într-un sistem de stocare mai apropiat (ca distanţă) de locul de unde se accesează datele. Browserele au *cache*, router-ele, etc. Browserul memorează în cache cópii ale paginilor WEB, accesate de utilizator mai des. Router-ul stochează în cache adrese şi informaţii de rutare, bazat pe experimentare virtuală (virtuală, în sensul că nu stă cineva să numere de câte ori se cere forwardarea către o adresă, ci există o aplicaţie inteligentă care monitorizează numărul de cereri de forwardare către acea adresă) într-o perioadă de timp fixată.

În cazul cache-lui unui browser, experimentul aleator în urma căruia se aleg paginile ce se stochează este următorul: dacă din n accesări de pagini WEB într-o perioadă dată de timp, utilizatorul accesează o pagină de k ori, atunci raportul k/n numit şi *hit ratio* reprezintă probabilitatea ca utilizatorul să acceseze acea pagină. Evident ca diferite pagini au probabilităţi diferite de a fi accesate (deci realizările "se accesează pagina Pag_i, $i = \overline{1,n}$ nu sunt egal probabile). Sunt salvate în cache cópii al paginilor cu *hit ratio* ridicat.

În CS se simulează (imită) un fenomen aleator printr-un algoritm ce poate genera toate evenimentele tipice pentru acel fenomen. Dacă algoritmul rulat de n ori generează de $k = k(n)$ ori acelaşi eveniment, A, atunci probabilitatea de producere a evenimentului A este aproximată de $p \approx \frac{k(n)}{n}$. Evident că rulând din nou programul de simulare pentru un n mai mare, numărul $k(n)$ va fi diferit de rularea precedentă, dar admiţând că probabilitatea exactă este

$$p = \lim_{n\to\infty} \frac{k(n)}{n},$$

atunci cu cât n este mai mare cu atât $\frac{k(n)}{n}$ aproximează mai bine probabilitatea exactă, p.

De exemplu dacă avem un algoritm ce generează biţi aleatori (deci simulează două evenimente posibile: bitul 0 si bitul 1), atunci generând $n = 1000$ de biţi şi numărând câţi biţi 1 a generat (de exemplu $k(1000) = 335$), se estimează probabiliatea de a genera bitul 1 prin $335/1000 \approx 1/3$.

Chiar dacă atribuirea probabilităţii se face în mod intuitiv sau bazat pe experienţă sau experimentare, ea trebuie să respecte nişte axiome care fundamentează ştiinţific predicţiile.

În teoria probabilităţilor familia tuturor evenimentelor asociate unui experiment aleator

este reprezentată de o familie $\mathcal{K}$ de părţi ale evenimentului sigur Ω, $\mathcal{K} \subseteq \mathcal{P}(\Omega)$. Familia $\mathcal{K}$ nu se alege în mod arbitrar, ci în aşa fel încât să verifice următoarele condiţii:

1. $\Omega \in \mathcal{K}$;
2. $A \in \mathcal{K} \Rightarrow \complement_\Omega A \in \mathcal{K}$
3. $A, B \in \mathcal{K} \Rightarrow A \cup B \in \mathcal{K}$; (**Reuniunea a două evenimente A şi B este evenimentul care se produce dacă cel puţin unul dintre cele două se produce.**

O astfel de familie o numim în continuare, familie admisibilă de evenimente.)

Propoziţia 1.2.1 *Dacă $\mathcal{K}$ este o familie admisibilă de evenimente, atunci:*
a) $A, B \in \mathcal{K} \Rightarrow A \cap B \in \mathcal{K}$; (**Intersecţia a două evenimente A şi B este evenimentul care se produce dacă ambele se produc.**)
b) $A, B \in \mathcal{K} \Rightarrow A \setminus B \in \mathcal{K}$.

Demonstraţie: a) Din definiţia familiei admisibile de evenimente avem că $A, B \in \mathcal{K} \Rightarrow \complement_\Omega A, \complement_\Omega B \in \mathcal{K}$ şi deci şi $\complement_\Omega A \cup \complement_\Omega B \in \mathcal{K}$. Complementul acestui eveniment este $\complement_\Omega(\complement_\Omega A \cup \complement_\Omega B) = A \cap B \in \mathcal{K}$.
b) Fie $A, B \in \mathcal{K}$. $A \setminus B = \{\omega \in A \text{ şi } \omega \notin B\} = A \cap \complement_\Omega B \in \mathcal{K}$. □

Evident că $A_i \in \mathcal{K}$, $i = \overline{1,n}$, implică $\cup_{i=1}^n A_i \in \mathcal{K}$ şi $\cap_{i=1}^n A_i \in \mathcal{K}$.

Două evenimente A, B cu proprietatea că producerea lor simultană este imposibilă, $A \cap B = \emptyset$, se numesc *evenimente incompatibile* sau *mutual exclusive* (se exclud unul pe altul). De exemplu, la aruncarea unei monede evenimentul A: "apariţia capului" şi evenimentul B: "apariţia stemei", sunt mutual exclusive.

Două evenimente A,B cu proprietatea că $A \cap B \neq \emptyset$ se numesc *evenimente compatibile*, adică există cel puţin o realizare a experimentului ce favorizează producerea ambelor evenimente.

Definiţia 1.2.1 *Fie $\mathcal{K}$ o familie admisibilă de evenimente asociate mulţimii Ω. O probabilitate pe $\mathcal{K}$ este o funcţie $P : \mathcal{K} \to [0,1]$ ce verifică condiţiile:*
1) $P(\Omega) = 1$.
2) Dacă $A, B \in \mathcal{K}$ sunt două evenimente mutual exclusive, adică $A \cap B = \emptyset$, atunci $P(A \cup B) = P(A) + P(B)$.

Tripletul $(\Omega, \mathcal{K}, P)$ se numeşte spaţiu de probabilitate.

Proprietăţi: a) Dacă $A \in \mathcal{K}$ atunci probabilitatea evenimentului opus lui A este $P(\complement_\Omega A) = 1 - P(A)$. În particular probabilitatea evenimentului imposibil este 0: $P(\emptyset) = 0$.
b) Dacă $A, B \in \mathcal{K}$ şi $A \cap B \neq \emptyset$, atunci $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
c) Dacă $A, B \in \mathcal{K}$ şi $A \subseteq B$, atunci $P(A) \leq P(B)$.

Demonstraţie: a) Deoarece $\Omega = A \cup \complement_\Omega A$ şi evenimentele A, $\complement_\Omega A$ sunt mutual exclusive, rezultă că $1 = P(A) + P(\complement_\Omega A)$, adică $P(\complement_\Omega A) = 1 - P(A)$. $\emptyset = \complement_\Omega \Omega$ şi deci $P(\emptyset) = 1 - P(\Omega) = 1 - 1 = 0$.
b) Fie $A, B \in \mathcal{K}$. $A \cup B$=$A \cup (B \setminus A)$. Cum A şi $B \setminus A$ sunt mutual exclusive avem:

$$P(A \cup B) = P(A) + P(B \setminus A) \tag{1}$$

Pe de altă parte $B = (A \cap B) \cup (B \setminus A)$ şi cum aceasta este o reuniune de evenimente disjuncte (mutual exclusive) rezultă că:

$$P(B) = P(A \cap B) + P(B \setminus A) \tag{2}$$

Scăzând relaţiile anterioare obţinem:

$$P(A \cup B) - P(B) = P(A) - P(A \cap B), \tag{3}$$

adică $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
c) Deoarece $A \subseteq B$, avem că $B = A \cup (B \setminus A)$. Astfel $P(B) = P(A) + P(B \setminus A)$. Cum $P(B \setminus A) \geq 0$, rezultă că $P(B) \geq P(A)$. □

Din proprietatea b) rezultă că $P(A \cup B) \leq P(A) + P(B)$. Această inegalitate se numeşte *inegalitatea lui Boole*.

Propoziţia 1.2.2 *a) Dacă evenimentele $A_1, A_2, \ldots, A_n$ sunt mutual exclusive două câte două, atunci $P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$;*
b) Pentru n evenimente arbitrare $A_1, A_2, \ldots, A_n$ probabilitatea reuniunii lor este:

$$\begin{aligned} P(\cup_{i=1}^n A_i) \;=\; & \textstyle\sum_{i=1}^n P(A_i) - \sum_{1 \leq i_1 < i_2 \leq n} P(A_{i_1} \cap A_{i_2}) + \\ & \textstyle\sum_{1 \leq i_1 < i_2 < i_3 \leq n} P(A_{i_1} \cap A_{i_2} \cap A_{i_3}) + \cdots \\ & (-1)^{k+1} \textstyle\sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) + \cdots + \\ & (-1)^{n+1} P(A_1 \cap A_2 \cap \cdots \cap A_n) \end{aligned} \tag{4}$$

În cazul particular $n = 3$ egalitatea b) din propoziţia de mai sus este:

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) \;=\; & P(A_1) + P(A_2) + P(A_3) \\ - \; & P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) \\ + \; & P(A_1 \cap A_2 \cap A_3) \end{aligned} \tag{5}$$

Exemplul 3. În criptografie (ştiinţa criptării (codificării) informaţiei pentru a fi transmisă în siguranţă către un destinatar) se lucrează cu spaţiul de probabilitate $(\Omega, \mathcal{P}(\Omega), P)$, unde Ω este mulţimea finită a stringurilor de n biţi

$$\{(s_1 s_2 \ldots s_n) \,|\, s_i \in \{0, 1\}\},$$

identificată cu mulţimea aplicaţiilor $s : \{1, 2, \ldots, n\} \to \{0, 1\}$, adică

$$\Omega = \{0,1\}^{\{1,2,\ldots,n\}} \stackrel{\text{notaţie}}{=} \{0,1\}^n,$$

Cardinalul mulţimii este $|\Omega| = 2^n$. Un element al lui $\{0,1\}^n$ se numeşte mesaj şi se notează cu $\mathfrak{m}$. O probabilitate pe $\{0,1\}^n$ este definită de o funcţie pozitivă $p : \{0,1\}^n \to \mathbb{R}^+$, ce asociază fiecărui mesaj un număr din intervalul $[0,1]$ şi

$$\sum_{\mathfrak{m} \in \{0,1\}^n} p(\mathfrak{m}) = 1$$

Probabilitatea ca un mesaj să fie generat, selectat, decriptat (depinde de context) se defineşte atunci prin $P(\{\mathfrak{m}\}) = p(\mathfrak{m})$.

Cea mai simplă probabilitate este definită de funcţia $p(\mathfrak{m}) = \dfrac{1}{|\Omega|} = \dfrac{1}{2^n}$, adică fiecare mesaj de n biţi are aceeaşi probabilitate de a fi generat, selectat sau decriptat (mesajele sunt echiprobabile). De obicei în criptografie se folosesc probabilităţi P' ce sunt ϵ–apropiate de această probabilitate uniformă. Mai precis, distanţa dintre două probabilităţi P, P' definite de funcţiile p, respectiv p', se defineşte prin:

$$\text{dist}(P, P') = \frac{1}{2} \sum_{\mathfrak{m} \in \{0,1\}^n} |p(\mathfrak{m}) - p'(\mathfrak{m})|$$