Curs 1

Statistică şi prelucrarea datelor

Teoria probabilităţii este ştiinţa incertitudinii. Ea dezvoltă reguli matematice exacte pentru a înţelege şi a analiza propria noastră ignoranţă. Înţelegerea ne face să fim mai exacţi, să luăm decizii mai bune, să evităm riscul.

1.1 Câmp de evenimente

Teoria probabilităţilor studiază experimentele aleatoare, experimente care reproduse de mai multe ori pot să se desfăşoare de fiecare dată în mod diferit, rezultatul fiind imprevizibil. Exemple: aruncarea unui zar, tragerile la ţintă, durata de funcţionare a unei maşini etc.

Pentru a modela şi analiza un experiment aleator trebuie să introducem mulţimea rezultatelor posibile ale experimentului.

Definiţia 1.1.1 *Mulţimea rezultatelor posibile ale unui experiment se numeşte* ***spaţiu de selecţie****.*

Vom nota spaţiul de selecţie cu Ω.

Un rezultat posibil al experienţei se numeşte **probă**.

Definiţia 1.1.2 *O submulţime a spaţiului de selecţie se numeşte* ***eveniment****.*

Rezultă că evenimentul este format din una sau mai multe probe.

Un spaţiu de selecţie se numeşte **finit** dacă este o mulţime finită, şi spunem că este **discret** dacă este o mulţime cel mult numărabilă (finită sau numărabilă). Un spaţiu de selecţie este **continuu** dacă conţine un interval de numere reale.

Exemplul 1.1.3 Cel mai simplu experiment este acela în care sunt posibile două rezultate. Un asfel de experiment constă, de exemplu, în verificarea unui tranzistor pentru a vedea dacă este corespunzător sau nu. Spaţiu de selecţie este: $\Omega = \{C, D\}$ (corespunzător sau defect) şi este un spaţiu discret.

Noţiunile de spaţiu de selecţie şi de eveniment astfel introduse ne permit ca teoria mulţimilor să poată fi folosită în studiul evenimentelor. Traducem în limbaj de evenimente noţiuni şi simboluri caracteristice teoriei mulţimilor.

1. Drept submulţime a lui Ω se poate considera Ω. Cum indiferent de rezultatul ω al experienţei, $\omega \in \Omega$, rezultă că odată cu ω se realizează Ω. Evenimentul Ω se numeşte **eveniment cert** (sau **eveniment sigur**).

2. Drept submulţime a lui Ω putem considera mulţimea vidă $\emptyset$ care nu se realizează la nicio efectuare a experienţei, motiv pentru care se numeşte **eveniment imposibil**.

3. Fie evenimentul $A \subset \Omega$. Evenimentul complementar lui A în raport cu Ω, notat $\overline{A}$, se numeşte **eveniment contrar** evenimentului A. Acesta se realizează dacă şi numai dacă nu se realizează evenimentul A.

4. Fie evenimentele $A, B \subset \Omega$. Evenimentul A implică evenimentul B (scris $A \subset B$) dacă B se realizează prin toate probele lui A (şi prin alte probe), adică dacă $(\omega \in A) \Rightarrow (\omega \in B)$.

5. Fie $A, B \subset \Omega$ două evenimente. Evenimentul $A \cup B$ este evenimentul care se realizează prin probe ale evenimentelor A sau B.

6. Fie $A, B \subset \Omega$. Prin evenimentul $A \cap B$ înţelegem evenimentul care se realizeză prin probe ale evenimentulor A şi B.

7. Fie $A, B \subset \Omega$. Prin $A \setminus B$ înţelegem evenimentul care se realizează prin probe ale evenimentelor A şi $\overline{B}$.

Definiţia 1.1.4 *Fie $A \subset \Omega, A \neq \emptyset$. Evenimentul A se numeşte **eveniment elementar** dacă este implicat numai de el însuşi şi de evenimentul imposibil. Celelalte evenimente se numesc **evenimente compuse.***

Definiţia 1.1.5 *Fie $A, B \subset \Omega$. Evenimentele A, B se numesc **compatibile** dacă se pot realiza simultan, adică există probe care realizează atât pe A cât şi pe B ($A \cap B \neq \emptyset$). În caz contrar evenimentele se numesc **incompatibile** ($A \cap B = \emptyset$).*

Definiţia 1.1.6 *Perechea $\{\mathbf{F}, \Omega\}$, $\mathbf{F} \neq \emptyset$, $\mathbf{F} \subset \mathcal{P}(\Omega)$, se numeşte **câmp de evenimente** dacă:*

a) $\forall A \in \mathbf{F} \Rightarrow \overline{A} \in \mathbf{F}$;

b) $\forall A, B \in \mathbf{F} \Rightarrow A \cup B \in \mathbf{F}$.

Observaţia 1.1.7 *Dacă Ω este o mulţime finită de evenimente, câmpul de evenimente se numeşte **finit**. Dacă Ω este cel mult numărabilă, atunci câmpul de evenimente se mai numeşte **câmp discret de evenimente**.*

Consecinţe ale definiţiei:

1. $\Omega \in \mathbf{F}$ deoarece $(\forall A \in \mathbf{F} \overset{a)}{\Rightarrow} \overline{A} \in \mathbf{F}) \overset{b)}{\Rightarrow} A \cup \overline{A} \in \mathbf{F} \Rightarrow \Omega \in \mathbf{F}$.

2. $\emptyset \in \mathbf{F}$ deoarece $\Omega \in \mathbf{F} \overset{a)}{\Rightarrow} \overline{\Omega} \in \mathbf{F}$, dar $\emptyset = \overline{\Omega} \Rightarrow \emptyset \in \mathbf{F}$.

3. Dacă $A, B \in \mathbf{F} \Rightarrow A \setminus B \in \mathbf{F}$ deoarece $A \setminus B = A \cap \overline{B} = \overline{\overline{A} \cup B} \in \mathbf{F}$.

4. Următoarele proprietăţi sunt echivalente:

$$(\forall A, B \in \mathbf{F} \Rightarrow A \cup B \in \mathbf{F}) \text{ şi } (\forall A, B \in \mathbf{F} \Rightarrow A \cap B \in \mathbf{F}),$$

   deoarece $A \cap B = \overline{\overline{A} \cup \overline{B}}$.

5. Din Definiţia 1.1.6 rezultă, folosind inducţia matematică, că pentru orice $n \geq 2$,

$$(A_j \in \mathbf{F}, 1 \leq j \leq n) \Rightarrow \bigcup_{j=1}^{n} A_j \in \mathbf{F}.$$

6. Din Consecinţa 4 rezultă că pentru orice $n \geq 2$,

$$(A_j \in \mathbf{F}, 1 \leq j \leq n) \Rightarrow \bigcap_{j=1}^{n} A_j \in \mathbf{F}.$$

Într-un câmp de evenimente au loc următoarele proprietăţi:

P1 *Două evenimente elementare distincte sunt incompatibile.*

Fie A_1 şi A_2 două evenimente elementare. Presupunem că nu ar fi incompatibile, adică $A_1 \cap A_2 = B \neq \emptyset$. Atunci $B \subset A_1, B \neq \emptyset$ deci A_1 nu este eveniment elementar, ceea ce este fals.

P2 *Într-un câmp finit de evenimente există evenimente elementare.*

Fie A_1 un eveniment. Dacă A_1 este elementar, afirmaţia este demonstrată. Dacă A_1 este eveniment compus există $A_2 \neq \emptyset$, $A_2 \neq A_1$ astfel încât $A_2 \subset A_1$. Dacă A_2 este eveniment elementar, afirmaţia este demonstrată. Dacă A_2 este eveniment compus se continuă raţionamentul anterior. Raţionamentul se opreşte deoarece câmpul este un câmp finit de evenimente.

P3 *Într-un câmp finit de evenimente orice eveniment al acestui câmp se poate scrie ca reuniune finită de evenimente elementare.*

Fie B un eveniment compus (dacă B este eveniment elementar atunci afirmaţia este demonstrată). Atunci există, conform proprietăţii P2, un eveniment elementar $A_1 \in \mathbf{F}$ şi un eveniment $B_1 \in \mathbf{F}$ astfel încâ $B = A_1 \cup B_1$, $B_1 = B \setminus A_1$ cu $A_1 \cap B_1 = \emptyset$. Dacă B_1 este eveniment elementar, afirmaţia este demonstrată. Dacă B_1 nu este eveniment elementar, există evenimentul elementar A_2 şi un eveniment $B_2 \in \mathbf{F}$ astfel încât $B_1 = A_2 \cup B_2$ şi deci $B = A_1 \cup A_2 \cup B_2$ şi raţionamentul se continuă. Numărul paşilor va fi finit deoarece cîmpul de evenimente este finit Deci $B = A_1 \cup A_2 \cup \ldots A_k$, unde $A_i, i = \overline{1,k}$ sunt evenimente elementare.

P4 *Într-un câmp finit de evenimente reuniunea tuturor evenimentelor elementare ale lui* $\mathbf{F}$ *este* Ω.

Într-adevăr, cum $\Omega \in \mathbf{F}$, rezultă din P3 că Ω se poate scrie sub forma $\Omega = A_1 \cup A_2 \cup \ldots \cup A_s$. Presupunem că în câmpul finit de evenimente mai există un eveniment elementar $A_n \neq A_j, j = \overline{1,s}$. Atunci

$$A_n \cap \Omega = A_n = A_n \cap (A_1 \cup \ldots \cup A_s) = (A_n \cap A_1) \cup \ldots \cup (A_n \cap A_s) = \emptyset$$

conform P1.

Nu de puţine ori de un real folos ne va fi descompunerea unui eveniment într-o reuniune de evenimente incompatibile două câte două.

Definiţia 1.1.8 *Fie* $\{\mathbf{F}, \Omega\}$ *un câmp de evenimente şi* $A_1, A_2, \ldots, A_n \in \mathbf{F}$. *Spunem că familia de evenimente* $A_1, A_2 \ldots A_n$ *formează un* ***sistem complet de evenimente*** *dacă:*

a) $A_i \neq \emptyset,\ i = \overline{1,n}$;

b) $A_i \cap A_j = \emptyset,\ \forall i \neq j,\ i, j = \overline{1,n}$;

c) $\bigcup\limits_{i=1}^{n} A_i = \Omega$.

Observaţia 1.1.9 *Într-un câmp finit de evenimente, mulţimea tuturor evenimentelor elementare ataşate unui experiment formează un sistem complet de evenimente.*

1.2 Definiţia probabilităţii

Introducem noţiunea de probabilitate ataşată unui eveniment aparţinând unui **câmp finit de evenimente**. Probabilitatea este folosită pentru a cuantifica şansa de a se produce un eveniment. "Probabilitatea ca să plouă este de 30%" este o afirmaţie care cuantifică posibilitatea de a ploua.

Definiţia 1.2.1 *Se numeşte* ***probabilitate (măsură de probabilitate)*** *o funcţie definită pe un câmp de evenimente* $\{\mathbf{F},\Omega\}$ *cu valori reale,* $P : \mathbf{F} \to [0, 1]$, *care satisface următoarele axiome:*

a) $P(\Omega) = 1$;

b) $P(A \cup B) = P(A) + P(B), \forall A, B \in \mathbf{F}, A \cap B = \emptyset$.

Observaţia 1.2.2 *Axioma c) din definiţie se extinde prin inducţie la orice număr finit de evenimente incompatibile două câte două, deci dacă* $A_i \cap A_j = \emptyset,\ i \neq j,\ i, j = \overline{1,n}$, *atunci*

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i).$$

Proprietăţi:

P1. $P(\emptyset) = 0$.

$(\Omega \cup \emptyset = \Omega$ şi $\Omega \cap \emptyset = \emptyset) \Rightarrow P(\Omega) = P(\Omega \cup \emptyset) = P(\Omega) + P(\emptyset) \Rightarrow P(\emptyset) = 0$.

P2. Pentru orice eveniment $A \in \mathbf{F}$ are loc $P(\overline{A}) = 1 - P(A)$.

Relaţia rezultă din $\Omega = \overline{A} \cup A, \overline{A} \cap A = \emptyset \overset{c)}{\Rightarrow} P(\Omega) = P(\overline{A}) + P(A) \overset{b)}{\Rightarrow} 1 = P(\overline{A}) + P(A) \Rightarrow P(\overline{A}) = 1 - P(A)$.

P3. $P(A \setminus B) = P(A) - P(A \cap B), \forall A, B \in \mathbf{F}$.

Din $(A \setminus B) \cup (A \cap B) = A$ şi $(A \setminus B) \cap (A \cap B) = \emptyset \overset{c)}{\Rightarrow} P(A \setminus B) + P(A \cap B) = P(A)$.

P4. Pentru orice $A, B \in \mathbf{F}$ cu proprietatea $A \subset B$ are loc relaţia $P(A) \leq P(B)$.

Într-adevăr, ţinând seama de P2 şi de faptul că $A \subset B$ avem $0 \leq P(B \setminus A) = P(B) - P(B \cap A) = P(B) - P(A)$. Deci $P(B) - P(A) \geq 0$ sau $P(B) \geq P(A)$.

Definiţia 1.2.3 *Se numeşte* ***câmp de probabilitate*** *un câmp de evenimente* $\{\mathbf{F},\Omega\}$ *pe care am definit o probabilitate* P. *Se notează* $\{\mathbf{F},\Omega, P\}$.

Observaţia 1.2.4 *Dacă Ω este o mulţime finită, fie $\Omega = \{A_1, A_2, \ldots, A_n\}$, atunci orice eveniment $A \in F, A \neq \emptyset$ pote fi scris ca o reuniune finită de evenimente elementare, $A = A_{i_1} \cup A_{i_2} \cup \ldots \cup A_{i_k}$. Atunci conform Observaţiei 1.2.2 obţinem*

$$P(A) = P(A_{i_1}) + \ldots + P(A_{i_k}).$$

Deci pentru a cunoaşte probabilitatea unui eveniment oarecare din F este suficient să cunoaştem probabilitatea tuturor evenimentelor elementare care-l compun. Probabilitatea unui eveniment A este suma probabilităţilor evenimentelor elementare ce-l compun. Evident, probabilităţile evenimentelor elementare satisfac condiţiile

$$P(A_i) \geq 0,\ i = \overline{1, n}, \tag{1.1}$$

$$P(A_1) + P(A_2) + \ldots + P(A_n) = P(\Omega) = 1. \tag{1.2}$$

Deci, fiind date toate evenimentele elementare care compun Ω, familia F este perfect determinată şi deci câmpul de probabilitate mai depinde de alegerea a n numere, probabilităţile evenimentelor elementare care satisfac condiţiile (1.1) şi (1.2). În cazul particular când evenimentele elementare sunt echiprobabile $P(A_1) = P(A_2) = \ldots = P(A_n) = \dfrac{1}{n}$, şi dacă $A = A_{i_1} \cup A_{i_2} \cup \ldots \cup A_{i_k}$, obţinem $P(A) = \dfrac{k}{n}$, deci ajungem astfel la definiţia probabilităţii dintr-un câmp de probabilitate în care spaţiul de selecţie este o mulţime cel mult numărabilă: probabilitatea unui eveniment este raportul dintre numărul evenimentelor elementare favorabile producerii evenimentului dat şi numărul total de evenimente elementare ale câmpului. De aici rezultă necesitatea de a studia metodele de numărare şi modurile de selectare ale evenimentelor.

Definiţia 1.2.5 *Numim* ***probabilitatea evenimentului*** *A* ***condiţionată de evenimentul*** *B, $P(B) \neq 0$, (notat $P_B(A)$ sau $P(A|B)$) probabilitatea de realizare a evenimentului A în ipoteza că evenimentul B s-a realizat, probabilitate definită prin*

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{1.3}$$

Observaţia 1.2.6 *Dacă presupunem că $P(A) \neq 0$, atunci*

$$P(B|A) = \frac{P(A \cap B)}{P(A)}. \tag{1.4}$$

Observaţia 1.2.7 *Din relaţiile (1.3) şi (1.4) reţinem $P(A \cap B) = P(B) \cdot P(A|B) = P(A) \cdot P(B|A)$.*

Fie $\{\mathbf{F}, \Omega, P\}$ un câmp de probabilitate şi $A, B \in \mathbf{F}$.

Definiţia 1.2.8 *Evenimentele A şi B sunt* ***independente (în probabilitate)*** *dacă probabilitatea ca unul să se realizeze* ***nu*** *depinde de faptul că celălalt s-a realizat sau nu, altfel spus*

$$P(A \cap B) = P(A) \cdot P(B). \tag{1.5}$$

Teorema 1.2.9 **a)** *Evenimentele A şi B cu $P(A) \cdot P(B) \neq 0$ sunt independente dacă şi numai dacă $P(B|A) = P(B)$.*

b) *Dacă evenimentele A şi B sunt independente atunci şi evenimentele*

$$\overline{A} \text{ şi } B; \quad A \text{ şi } \overline{B}; \quad \overline{A} \text{ şi } \overline{B}$$

sunt independente.

Demonstraţie. a) Evenimentele A şi B sunt independente $\Rightarrow P(A \cap B) = P(A) \cdot P(B)$. Dar $P(A \cap B) = P(A) \cdot P(B|A) \Rightarrow P(B|A) = P(B)$. Reciproc, deoarece $P(B) = P(B|A) = \dfrac{P(A \cap B)}{P(A)}$ rezultă $P(A \cap B) = P(A) \cdot P(B)$, adică (1.5).

b) Demonstrăm că (1.5) implică $\overline{A}, B$ independente:

$$P(\overline{A} \cap B) = P(\overline{A}) \cdot P(B) \tag{1.6}$$

Într-adevăr, dacă evenimentele A şi B sunt independente, $P(A \cap B) = P(A) \cdot P(B)$ şi deoarece $\overline{A} \cap B = B \setminus A$ avem $P(B \setminus A) = P(B) - P(A \cap B) = (1 - P(A)) \cdot P(B) = P(\overline{A})P(B)$. Analog se demonstrează că A şi $\overline{B}$, respectiv $\overline{A}$ şi $\overline{B}$ sunt independente. $\square$

Definiţia 1.2.10 *Date evenimentele $A_1, A_2, \ldots, A_n$, vom spune că sunt* ***global independente*** *dacă probabilitatea oricărei subintersecţii este egală cu produsul probabilităţilor evenimentelor intersectate, adică dacă*

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

oricare ar fi $I \subset \{1, 2, ..., n\}$.

Observaţia 1.2.11 *Din Definiţia 1.2.10 rezultă că dacă trei evenimente sunt independente două câte două nu rezultă că sunt independente în totalitatea lor.*

Exemplul lui ***S. N. Bernstein*** *ne va ilustra acest lucru.*
Considerăm un tetraedru omogen cu feţele colorate astfel:

- *una în alb,*
- *una în negru,*
- *una în roşu şi*
- *a patra în toate cele trei culori.*

Aruncând tetraedrul pe o masă el se aşează pe una din feţe; ne interesează probabilitatea apariţiei fiecărei culori şi independenţa evenimentelor corespunzătoare. Notăm cu

- *A_1 evenimentul care constă în apariţia culorii albe,*
- *A_2 evenimentul care constă în apariţia culorii negre şi*
- *A_3 evenimentul care constă în apariţia culorii roşii.*

Avem:

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{2}$$

deoarece pentru fiecare culoare sunt patru cazuri posibile şi două favorabile (faţa cu culorea respectivă şi faţa cu cele trei culori). Se constată că

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

dar

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

Bibliografie

[1] Martha L. Abell, James P. Braselton, John A. Rafter, *Statistics with Mathematica*, Academic Press, San Diego, 1998.

[2] Petru Blaga, *Statistică...prin MATLAB*, Presa Universitară Clujeană, Cluj-Napoca, 2002.

[3] Teresa Bradley - *Essential Statistics for Economics*, Business and Management, John Willey&Sons, Inc., 2007

[4] Nicoleta Breaz, Lucia Căbulea, Ariana Pitea, Gheorghiţă Zbăganu, Rodica Tudorache, Ioan Rasa, *Probabilităţi şi statistică*, Editura SudIS, Iaşi, 2013.

[5] Mariana Craiu, *Statistică matematică, teorie şi probleme*, MATRIX ROM, Bucureşti, 1998.

[6] Kai Lai Chung, *Elementary Probability Theory with Stochastic Process*, Springer Verlag, Berlin, 1982.

[7] Michael J. Evans, Jeffrey S. Rosenthal, *Probability and Statistics*, W. H. Freeman and Company, New York, 2004.

[8] Teohari Ganciu, Cristina Ţugurlan, *Elemente de statistică şi fiabilitate*, Editura Universităţii "Gheorghe Asachi", Iaşi, 2002.

[9] John J. Kinney, *Probability*, Willey&Sons, Inc., 1997.

[10] Douglas C. Montgomery, George C. Runger, *Applied Statistics and Probability for Engineers*, Fourth Edition, John Willey&Sons, Inc., 2006.

[11] Ariadna Lucia Pletea, Mircea Lupan, *Statistică prin... MATLAB şi MATHEMATICA*, Editura Universităţii Tehnice "Gheorghe Asachi", Iaşi, 2009.

[12] Ariadna Lucia Pletea, Liliana Popa, *Teoria Probabilităţilor*, Editura Universităţii Tehnice "Gheorghe Asachi", Iaşi, 1999.

[13] Tatiana Stănăşilă, *Metode statistice pentru ingineri, teorie, exerciţii, aplicaţii*, MATRIX ROM, Bucureşti, 1998.