Universitatea din Bucureşti
Facultatea de Matematică şi Informatică

METODE MODERNE DE CALCUL ŞI SIMULARE

Cursul nr. 1 – Anul III – Tehnologia Informaţiei

Asist. dr. Bianca Mogoş

Bucureşti, Februarie 2014

Cuprins

Capitolul 1

Noţiuni de probabilităţi şi statistică

1.1 Experiment aleator. Spaţiu de selecţie.

Definiţia 1.1.1. *(Experiment aleator)*

Un experiment aleator este definit ca fiind o acţiune al cărei rezultat nu poate fi prezis cu certitudine şi care se poate modifica în urma repetării experimentului.

Observaţie 1.1.2. Variabilitatea rezultatelor ieşite în urma unui experiment aleator poate apărea ca urmare a unor erori de măsurare, a alegerii unor obiecte diferite pentru testare, etc.

Exemplul 1.1.3.

- Observarea pe un interval T de timp a funcţionării unui calculator.
- Înregistrarea consumului de energie electrică a unui combinat.

Definiţia 1.1.4. *(Spaţiu de selecţie)*

Spaţiul de selecţie al unui experiment aleator, notat prin S, reprezintă mulţimea tuturor rezultatelor posibile.

Exemplul 1.1.5.

- Prin aruncarea banului se pot obţine două rezultate. Astfel, spaţiul de selecţie este $\{0,1\}$.
- Prin rostogolirea unui zar cu şase feţe şi numărarea punctelor de pe o faţă se pot obţine şase rezultate posibile. Astfel, spaţiul de selecţie este $\{1,2,3,4,5,6\}$.

1.2 Evenimente

Definiţia 1.2.1. *(Eveniment)*

Un eveniment este orice submulţime de rezultate conţinute în spaţiul de selecţie. Un eveniment este elementar dacă el constă dintr-un singur rezultat şi compus dacă constă din mai multe rezultate.

Definiţia 1.2.2. *(Evenimentul sigur şi evenimentul imposibil)*

Evenimentul sigur este evenimentul care se realizează întotdeauna ca rezultat al experimentului; va fi notat cu Ω (se asociază mulţimii totale de rezultate).

Evenimentul imposibil nu se poate realiza ca rezultat al unui experiment; va fi notat cu $\emptyset$ (corespunde mulţimii vide).

Definiţia 1.2.3. *(Reuniunea şi intersecţia evenimentelor)*

Numim reuniunea evenimentelor $A_1, A_2, \ldots, A_k$, notată prin $\bigcup\limits_{j=1}^{k} A_j$, evenimentul care se realizează când cel puţin unul dintre evenimentele $A_1, A_2, \ldots, A_k$ se realizează.

Numim intersecţia evenimentelor $A_1, A_2, \ldots, A_k$, notată prin $\bigcap\limits_{j=1}^{k} A_j$, evenimentul care se realizează când se realizează toate evenimentele $A_1, A_2, \ldots, A_k$.

1.3 Axiomele probabilităţii. Câmp de probabilitate

Probabilitatea unui eveniment este o măsură prin care gestionăm/modelăm incertitudinea producerii fenomenelor şi a apariţiei datelor din lumea reală. Aceasta ne permite să apreciem gradul de încredere şi să cuantificăm lipsa de siguranţă inerentă în procesul care generează datele analizate.

Definiţia 1.3.1. *(Definiţia clasică a probabilităţii)*

Dacă într-un experiment cu "n" rezultate, "k"dintre ele favorizează realizarea evenimentului A, definim probabilitatea P(A) a evenimentului A prin

$$P(A) = \frac{k}{n} \tag{1.3.1}$$

Exerciţiul 1.3.2. *(Aruncarea cu zarul)*

1. Calculaţi probabilitatea ca în urma aruncării unui zar să iasă un număr par.
2. Care este probabilitatea ca în urma aruncării unui zar să iasă un număr impar, strict mai mare ca 3?

Teoria modernă a probabilităţilor a fost fondată la începutul secolului XX. O contribuţie esenţială a avut-o matematicianul rus A. Kolmogorov.

Fie Ω mulţimea evenimentelor elementare (rezultatelor posibile) ale unui experiment aleator.

Definiţia 1.3.3. *(σ – algebră)*

$\mathcal{B} \subset \mathcal{P}(\Omega)$ este o σ – algebră dacă

1. $\emptyset, \Omega \in \mathcal{B}$

2. Dacă $A \in \mathcal{B}$ atunci $A^C \in \mathcal{B}$

3. Dacă $(A_n)_{n\in N^*}, A_n \in \mathcal{B}$ atunci $\bigcup\limits_{n=1}^{\infty} A_n \in \mathcal{B}$

Perechea $(\Omega, \mathcal{B})$ se numeşte câmp de evenimente.

Definiţia 1.3.4. *(Probabilitate)*

Se numeşte probabilitate pe $\mathcal{B}$ o funcţie nenegativă $P : \mathcal{B} \to [0,1]$ cu proprietăţile:

1. $P(\Omega) = 1$

2. Dacă $A, B \in \mathcal{B}$ şi $A \cap B = \emptyset$ atunci $P(A \cup B) = P(A) + P(B)$.

Proprietăţi 1.3.5.

1. $P(A^C) = 1 - P(A)$, *deoarece* $A \cup A^C = \Omega, A \cap A^C = \emptyset$.

2. $P(\emptyset) = P(\Omega^C) = 1 - P(\Omega) = 0$.

3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, *dacă* $A \cap B \neq \emptyset$.
   Relaţia 3) se demonstrează astfel:

   $A = A \cap (B \cup B^C) = (A \cap B) \cup (A \cap B^C)$ cu $A \cap B \cap B^C = \emptyset$.

   $B = B \cap (A \cup A^C) = (B \cap A) \cup (B \cap A^C)$ cu $B \cap A \cap A^C = \emptyset$.
   Dar, $A \cup B = (A \cap B) \cup (A \cup B^C) \cup (B \cap A^C)$.
   Aplicăm axioma 2) din definiţia probabilităţii şi obţinem

   $P(A) = P(A \cap B) + P(A \cap B^C)$

   $P(B) = P(A \cap B) + P(B \cap A^C)$

   $P(A \cup B) = P(A \cap B) + P(A \cap B^C) + P(B \cap A^C)$
   Din cele trei relaţii deduse mai sus rezultă proprietatea 3.

Definiţia 1.3.6. *(Câmp de probabilitate)*

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

Exerciţiul 1.3.7. *(Problema controlului alimentelor)*
Se ştie că într-o cutie sunt 550 de mere. Se verifică aleator 25 dintre ele dacă sunt stricate. Dacă 2% dintre merele din cutie sunt stricate, care este probabilitatea ca printre cele 25 testate săgăsim 2 mere stricate?

Exerciţiul 1.3.8. *(Problema zilei de naştere)*
Care este probabilitatea ca cel puţin 2 persoane dintr-un grup de n indivizi să aibă aceeaşi zi de naştere?

1.4 Probabilităţi condiţionate

Definiţia 1.4.1. *Evenimente independente*

Două evenimente A şi B din $\mathcal{B}$ sunt independente dacă nu se influenţează, adică realizarea evenimentului A nu depinde de realizarea lui B şi reciproc.

Definiţia 1.4.2. *(Probabilitate condiţionată)*

Se numeşte probabilitate condiţionată a evenimentului A de către evenimentul B şi se notează $P(A|B)$ sau $P_B A$ probabilitatea de realizare a evenimentului A calculată în condiţia că evenimentul B s-a realizat $(P(B) > 0)$ şi se defineşte ca:

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

Exemplul 1.4.3.

Un tratament s-a aplicat la 700 de pacienţi care sufereau de piatră la rinichi. S-a constatat că doar 562 s-au vindecat. De asemenea, se ştie că 357 au pietre mici şi 315 dintre aceştia s-au vindecat, iar 343 au pietre mari şi 247 s-au vindecat. Să se calculeze probabilitatea ca un pacient să se vindece ştiind că acesta are piatră mică. Dar, dacă are piatră mare?

Soluţie:

Fie

$\Omega = \{1, 2, \ldots, 700\}$ mulţimea tuturor pacienţilor,

$S = \{1, 2, \ldots, 562\}$ mulţimea pacienţilor vindecaţi şi

$E = \{563, \ldots, 700\}$ mulţimea pacienţilor nevindecaţi.

Obţinem următoarele probabilităţi:

$P(\omega) = \frac{1}{700}, \forall$ evenimentul elementar $\omega \in \Omega$

$P(S) = \frac{562}{700} \approx 0,8$

$P(E) = \frac{138}{700} = \approx 0,2$

Notăm cu

$\Omega_1 =$ pacienţii care au pietre mici

$\Omega_2 =$ pacienţii care au pietre mari

Rezultă $P(\Omega_1) = \frac{357}{700} = 0,51$ şi $P(\Omega_2) = \frac{343}{700} = 0,49$. Astfel

$$P(S|\Omega_1) = \frac{P(S \cap \Omega_1)}{P(\Omega_1)} = \frac{315}{700} \cdot \frac{1}{0,51} = 0,88$$

Analog, se arată că

$P(S|\Omega_2) = 0,72.$

Exerciţiul 1.4.4.

Presupunem că 1% din populaţie este bolnavă. Un test de analiză a sângelui dă următoarele rezultate: în 95% dintre cazuri testul iese pozitiv pentru persoanele bolnave şi în 2% dintre cazuri iese pozitiv pentru persoanele sănătoase. Care este probabilitatea ca o persoană căreia i-a ieşit testul pozitiv să fie bolnavă?

Definiţia 1.4.5. *(Evenimente independente în totalitate)*

Spunem că evenimentele $A_1, A_2, \ldots, A_k$ sunt independente în totalitate dacă sunt independente două câte două, sunt independente trei câte trei,..., sunt independete toate "k".

Formal, avem relaţiile:

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_p}) = P(A_{i_1})P(A_{i_2}) \ldots P(A_{i_p}),$$
$$\forall i_1, \ldots, i_p \in \overline{1,k} \text{ distincte }, p = \overline{1,k}$$

Exemplul 1.4.6. *(Evenimentele nu sunt independente în totalitate)*

Fie spaţiul de selecţie $\Omega = \{1,2,3,4\}$ şi evenimentele.

$A = \{1,2\}$

$B = \{1,3\}$

$C = \{1,4\}$

Cum $A \cap B = \{1\}$ rezultă $P(A \cap B) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(A)P(B)$. Rezultă că evenimentele A şi B sunt independente.

Iar $P(A \cap B \cap C) = \frac{1}{4} \neq P(A)P(B)P(C) = \frac{1}{8}$. Deducem că evenimentele A, B şi C nu sunt independente.

1.5 Formula lui Bayes

Fie $(A_i)_{i=\overline{1,k}}$ o partiţie a mulţimii Ω, ceea ce înseamnă că

$$A_i \cap A_j = \emptyset, \forall i,j \in \overline{1,k}, i \neq j, P(A_j) > 0, \forall j = \overline{1,k} \text{ şi } \bigcup_{j=1}^{k} A_j = \Omega.$$

Fie $X \in \mathcal{B}$ un eveniment oarecare. Atunci

$$P(X) = P(X \cap \Omega) = P(\bigcup_{j=1}^{k} X \cap A_j) = \sum_{j=1}^{k} P(X \cap A_j) = \sum_{j=1}^{k} P(A_j) \cdot P(X|A_j)$$

Obţinem relaţia

$$P(X) = \sum_{j=1}^{k} P(A_j) \cdot P(X|A_j)$$

Atunci probabilitatea condiţionată a evenimentului A_j de către evenimentul X se poate calcula cu formula

$$P(A_j|X) = \frac{P(A_j) \cdot P(X|A_j)}{\sum_{j=1}^{k} P(A_j) \cdot P(X|A_j)}. \tag{1.5.1}$$

Relaţia obţinută în ecuaţia (1.5.1) se numeşte formula lui Bayes.

Exemplul 1.5.1. *(Exemplu bazat pe formula lui Bayes)*

Presupunem că trei maşini A, B şi C lucrează cu următoarele performanţe: probabilitatea ca o piesă să fie defectă este 0,01; 0,02 şi 0,04 pentru maşinile A, B şi respectiv C. În aceeaşi zi maşina A a produs 3900 de piese, maşina B 4200 de piese, iar maşina C 3600 de piese. Se alege la întâmplare o piesă din producţia acelei zile şi se constată că este defectă. Care este probabilitatea să fi fost produsă de maşina C?

Soluţie: Notăm cu

Ω = mulţimea tuturor pieselor produse într-o zi

A = evenimentul: piesa este produsă de maşina A

B = evenimentul: piesa este produsă de maşina B

C = evenimentul: piesa este produsă de maşina C

Atunci

$P(A) = \dfrac{3900}{11700} = 0,33$

$P(B) = \dfrac{4200}{11700} = 0,36$

$P(A) = \dfrac{3600}{11700} = 0,31$

De asemenea, notăm cu D = evenimentul: piesa este defectă. Avem din enunţ probabilităţile condiţionate: $P(D|A) = 0,01; P(D|B) = 0,02$ şi $P(D|C) = 0,04$

Atunci aplicând formula Bayes obţinem

$$P(C|D) = \frac{P(C) \cdot P(D|C)}{\sum_{A \in \{A,B,C\}} P(A) \cdot P(D|A)} = \frac{0,31 \cdot 0,04}{0,0229} = 0,54.$$

Bibliografie

[Craiu (1998)] M. Craiu (1998), *Statistică matematică: teorie şi probleme*, Editura Matrix Rom, Bucureşti

[Martinez, Martinez (2002)] W. L. Martinez, A. R. Martinez (2002), *Computational Statistics Handbook with MATLAB*, Chapman & Hall/CRC, Boca Raton London New York Washington, D.C.