Matematici speciale

(curs 2-S2)

28.02.2024

Recapitulare

Dată o experiența aleatoare, tripletul $(\Omega, \mathcal{K}, P)$ **spatiu de probabilitate**, unde Ω-spațiul tuturor realizărilor acestei exp, $\mathcal{K} \subset \mathcal{P}(\Omega)$ o familie (admisibila) de evenimente, iar $P : \mathcal{K} \to [0,1]$ funcția probabilitate

Ev. A	Multime A	$P(A) \in [0,1]$
Ev. sigur	Ω	1
Ev. imposibil	$\emptyset$	0
Ev. contrar lui A	$\complement_\Omega A$	$P(\complement_\Omega A) = 1 - P(A)$
Ev. reuniune	$A \cup B$	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$
Ev. intersectie	$A \cap B$	$P(A \cap B) = ?$
Ev. mutual exclusive	$A \cap B = \emptyset$	$P(A \cap B) = 0$
Ev. diferenta	$A \setminus B$	$P(A \setminus B) = P(A) - P(A \cap B)$

Definiția probabilităților condiționate

Fie E_1, E_2 evenimente din $\mathcal{K}$

$P(E_1) \neq 0$

Atunci $P(E_2) = P(E_2 | E_1) = \dfrac{P(E_1 \cap E_2)}{P(E_1)} \Rightarrow \boxed{P(E_1 \cap E_2) = P(E_1) \cdot P(E_2 | E_1)}$

$P_{E_1}(E_2)$

E_1 = eveniment condiționat = fixat

Propoziție: $0 \leq P(E_2 | E_1) \leq 1$

$P_{E_1}(\Omega) = 1;$

E, F - mutual exclusive $P_{E_1}(E \cup F) = P_{E_1}(E) + P_{E_1}(F)$

Aplicație

Alg. probabilist pt. verificare a egalității a 2 polinoame

(?) $F(x) = G(x)$

$a_m x^m + \ldots + x_0$ $\qquad$ $a(x - \alpha_1) \cdot \ldots (x - \alpha_d)$

Pas 1: $S = \{1, 2, \ldots, 100d\}$ alegem aleator

Pas 2: $F(\Omega) = G(\Omega) \xrightarrow{DA} F(x) = G(x)$

$\xrightarrow{NU} F(x) \neq G(x)$

Eroare: r - o răd. a polinomului $F - G$

$$\left.\begin{aligned} &\Rightarrow \underbrace{P(\text{alg. rasp. gresit})}_{P(E_1)} \leq \frac{d}{100d} = \frac{1}{100} \\ &P(E_2 | E_1) = \frac{d-1}{100d-1} \end{aligned}\right\} \Rightarrow P(E_1)\,P(E_2|E_1) \leq \frac{d(d-1)}{100d(100d-1)}$$

Eșantionare: extragerea unor bile din S

- cu înlocuire $\Rightarrow$ independența (extrag. nu se influențează reciproc)
- fără înlocuire $\Rightarrow$ extrag. se vor influența reciproc, pentru că se modifică numărul de elemente

În cazul independenței evenimentelor de tip E_i = la iterația i a algoritmului se extrage o rădăcină a lui $F - G$. Formula de condiționare iterată:

$$P(E_1 \cap E_2 \cap \ldots \cap E_k) =$$
$$P(E_1)P(E_2|E_1)P(E_3|E_1 \cap E_2)\ldots P(E_k|E_1 \cap E_2 \ldots \cap E_{k-1})$$

Eșantionarea cu înlocuire

$$P(E_1 \cap E_2 \ldots \cap E_k) = \prod_{i=1}^{k} P(E_i) \leq \left(\frac{d}{100d}\right)^k = \left(\frac{1}{100}\right)^k$$

Eșantionare fără înlocuire

$$P(E_1 \cap E_2) = P(E_1) \cdot P(E_2 | E_1) \leq \underbrace{\frac{d}{100d} \cdot \frac{d-1}{100d-1}}_{\text{fără înlocuire}} \leq \underbrace{\left(\frac{1}{100}\right)^2}_{\text{cu înlocuire}}$$

- Avem
$$P(E_j|E_1 \cap E_2 \cap \ldots E_{j-1}) \leq \frac{d-(j-1)}{100d-(j-1)}$$
-
$$P(E_1 \cap E_2 \cap \ldots E_k) \leq \prod_{j=1}^{k} \frac{d-(j-1)}{100d-(j-1)} \leq \left(\frac{1}{100}\right)^k$$

Evenimente independente

Def: realizarea evenimentului E_1 nu influențează evenimentul E_2

$$P(E_1 \cap E_2) = P(E_1)\cdot \underline{P(E_2|E_1)} = P(E_1)\cdot \underline{P(E_2)}$$

$$P(E_2|E_1) = P(E_2)$$

Mai general, evenimentele $E_1, E_2, \ldots, E_n$ cu proprietatea:

$$P(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_k}) = P(E_{i_1})P(E_{i_2})\cdots P(E_{i_k}),$$

pentru orice $k \in \{2, \ldots n\}$ și indicii $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ se numesc evenimente independente.

Observații legate de evenimente independente

- Pentru a testa independența a n- evenimente $E_1, E_2, \ldots, E_n$, avem de verificat $C_n^2 + C_n^3 + \ldots C_n^n = 2^n - (1+n)$ relații.

1. $\{E_1, E_2, \ldots, E_n\}$ → evenimente independente 2 câte 2

$$P(E_i \cap E_j) = P(E_i)\cdot P(E_j) \ ; \ i, j = \overline{1,n}$$

2. $\{E_1, \ldots, E_n\}$ → evenimente independente

$$P(E_{i_1} \cap \ldots E_{i_k}) = \prod_{j=1}^{k} P(E_{i_j})$$

ex: Dacă E_1, E_2, E_3 sunt trei evenimente, două câte două independente, atunci E_1, E_2, E_3 nu sunt neapărat independente, adică relațiile: $P(E_1 \cap E_2) = P(E_1)P(E_2), P(E_1 \cap E_3) = P(E_1)P(E_3), P(E_2 \cap E_3) = P(E_2)PE_1C)$, dar nu implică obligatoriu:

$P(E_1 \cap E_2 \cap E_3) = P(E_1)P(E_2)P(E_3)$. $= \prod_{i=1}^{3} P(E_i)$

$n = 3 \Rightarrow 2^3 - (3+1) = 4$ relații

Remărci legate de conceptul de probabilitate condiționată

- dat evenimentul E_1, avem $P(E_1)$;
- date două evenimente E_1, E_2, avem $P(E_1 \cap E_2) = P(E_1, E_2)$ probabilitatea ca evenimentele să se realizeze simultan (joint probability)
- **probabilitate condiționată**: $P(E_2|E_1) := \frac{P(E_1 \cap E_2)}{P(E_1)}$
- regula produsului: $P(E_1 \cap E_2) = P(E_1)P(E_2|E_1)$
- simetria probabilitații ev. intersecție: $P(E_1 \cap E_2) = P(E_2 \cap E_1)$
- avem: $P(E_2|E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)} = \frac{P(E_1|E_2)P(E_2)}{P(E_1)}$
- **prob cond. nu este simetrica**: $P(E_1|E_2) \neq P(E_2|E_1)$
- avem: $P(E_2|\bar{E}_1) = 1 - P(\bar{E}_2|\bar{E}_1)$
- ev. independente: $P(E_2|E_1) = P(E_2) \Leftrightarrow P(E_1 \cap E_2) = P(E_1)P(E_2)$

$$P((E_1 \cap E_2) \cap E_3) = P(E_1 \cap E_2) \cdot P(E_3 | E_1 \cap E_2)$$

$$P(E_1 \cap E_2 \cap E_3) = P(E_1) \cdot P(E_2 | E_1) \cdot P(E_3 | E_1 \cap E_2)$$

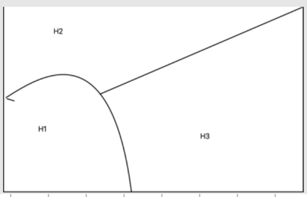
Formula probabilității condiționate

Formula lui Bayes

Notații

- $(\Omega, \mathcal{K}, P)$ un câmp de probabilitate
- Fie evenimentele $H_1, H_2, \dots H_n \in \mathcal{K}$, ce formează o partiție a evenimentului sigur Ω:
  - $\cup_{i=1}^{n} H_i = \Omega$
  - $H_i \cap H_j = \emptyset$, $i \neq j$
- Evenimentele H_i, $i \in \{1, n\}$ se numesc **ipoteze**.
  Ipotezele sunt acceptate cu o anumită probabilitate, $P(H_i)$.
- Avem:

$$1 = P(\Omega) = P(H_1) + P(H_2) + \dots + P(H_n)$$



- Dacă apare o nouă informație A, atunci nivelul de veridicitate (verosimilitate) al acestei informații, se reprezintă prin probabilitățile condiționate $P(A|H_k)$, $k \in \{1, n\}$.
- $P(A|H_k)$ este nivelul de verosimilitate al informației A în condițiile acceptării ipotezei H_k.
- $P(H_k)$- probabilităților ipotezelor se mai numesc **probabilități apriorice**

Cum se calculează probabilitatea $P(A)$, adică nivelul de veridicitate al informației A?

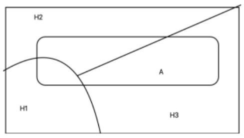


Figure: Evenimentele implicate în formula lui Prob Totale

Formula probabilității totale

Dacă $A \in \mathcal{K}$ este un eveniment informație, atunci gradul/nivelul de veridicitate al acestei informații este:

$$P(A) = \sum_{i=1}^{n} P(H_i)P(A|H_i)$$

Demonstrație:

- $H_1, H_2, \ldots, H_n$ fiind mutual exclusive două câte două, deci și $(A \cap H_i), (A \cap H_j)$, $i \neq j$, mutual exclusive două câte două;
- scriem $A = (A \cap H_1) \cup (A \cap H_2) \cup \cdots \cup (A \cap H_n)$;
- $P(A) = \sum_{i=1}^{n} P(A \cap H_i)$;
- Din formula probabilității condiționate avem: $P(A \cap H_i) = P(H_i)P(A|H_i)$,

Formula lui Bayes

pe baza unei noi informații (A), reevaluând prob. ev. ipoteză (Hi)

$$P(H_k|A) = \frac{P(H_k)P(A|H_k)}{\sum_{i=1}^{n} P(H_i)P(A|H_i)} \quad k \in \{1, n\}$$

Demonstrație Exprimăm $P(H_k \cap A)$ în două moduri:$P(H_k \cap A) = P(H_k)P(A|H_k) = P(A)P(H_k|A)$.
Din ultima egalitate avem:

$$P(H_k|A) = \frac{P(H_k)P(A|H_k)}{P(A)} = \frac{P(H_k)P(A|H_k)}{\sum_{i=1}^{n} P(H_i)P(A|H_i)}, \quad k \in \{1, n\}$$

$P(H_k|A)$−se numesc probabilități posterioare

Aplicații

1. Definim ev:

H_1 - persoana selectată are boala $\qquad P(H_1) = 0,02$

H_2 - persoana selectată nu are boala $\qquad P(H_2) = P_C(H_1) = 1 - 0,02 = 0,98$

C.E. $\begin{cases} H_1 \cup H_2 = \Omega \\ H_1 \cap H_2 = \emptyset \end{cases}$

A_+ - test pozitiv

A_- - test negativ

$\underbrace{P(A_- \mid H_1)}_{\text{rata fals negativă}} = 0,001 \qquad P(A_+ \mid H_1) = 0,999$

$\underbrace{P(A_+ \mid H_2)}_{\text{rata fals pozitivă}} = 0,005 \qquad P(A_- \mid H_2) = 0,995$

$$P(H_1 \mid A_+) = \frac{P(H_1) \cdot P(A_+ \mid H_1)}{P(A_+)} = 0,8$$

$$P(A_+) = P(H_1) \cdot \underbrace{P(A_+ \mid H_1)}_{?} + P(H_2) \cdot P(A_+ \mid H_2)$$

$$P(A_+ \mid H_1) = 1 - P(A_- \mid H_1) = 0,999$$