Curs 4 - Probabilități condiționate, independență și formula lui Bayes

Cum se schimbă/modifică probabilitatea realizării unui eveniment atunci când avem informații suplimentare.

A - "zi ploioasă"

$\mathbb{P}(A)$ - probabilitatea inițială (prior)

B - "afară este înnorat"

$\mathbb{P}$(că va ploua știind că este înnorat)

Exp: Aruncăm cu banul (echilibrat) de 3 ori

a) Care este prob. să obținem HHH?

$A = \{$am obținut HHH$\} = \{HHH\}$

$\Omega = \{H, T\}^3$, $|\Omega| = 8$

$\mathbb{P}(A) = 1/8$

b) La prima aruncare am obținut cap. Cât este $\mathbb{P}$(A având informația suplimentară)?

$\Omega' = \{HHH, HHT, HTH, HTT\}$

Ω'

HHH	HHT	THH	THT
HTH	HTT	TTH	TTT

A

Ω

$\frac{1}{4} = \mathbb{P}(A \mid B)$

știind că

ev. care s-a realizat

Probabilitatea realizării ev. A știind că ev. B s-a realizat

Din persp. frecvenționistă: Pp că repetăm experimentul de N ori și de fiecare dată înregistrăm dacă A și/resp B s-au realizat

Ne uităm doar la experimentele în care B s-a realizat și vrem să calculăm frecvența de realizare a lui A:

$$\frac{N(A\cap B)}{N(B)} \quad \begin{array}{l} \longrightarrow \text{nr. de exp în care și } A \text{ și } B \text{ s-au realizat} \\ \longrightarrow \text{nr. de exp. în care } B \text{ s-a realizat} \end{array}$$

$$= \frac{\frac{N(A\cap B)}{N}}{\frac{N(B)}{N}} \sim \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}$$

<u>Def</u>: Dacă A și B sunt evenimente a.î. $\mathbb{P}(B) > 0$ atunci probabilitatea condiționată a lui A la B se notează cu $\mathbb{P}(A|B)$ și este definită prin

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

<u>Not</u>: $\mathbb{P}(A)$ - prior (prob. inițiale)

$\mathbb{P}(A|B)$ - posterior

"$A|B$" - <u>nu</u> este un eveniment

<u>Exp</u>: (cont.) B - la prima aruncare am obținut cap

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A)}{\mathbb{P}(B)} = \frac{1/8}{4/8} = \frac{1}{4}$$

pt că $A\cap B = A$

<u>Exp2</u>: Un pachet de cărți de joc și extragem (în mod aleator) două cărți de joc una după alta fără întoarcere.

A - "prima carte este de inimă roșie"

B - "a doua carte este de inimă roșie"

Vrem să calculăm $\mathbb{P}(B|A) = ?$

Sol:

Avem 13 cărți de inimă roșie

$$\mathbb{P}(B|A) = \frac{12}{51}$$

C - "a doua carte este de culoare roșie"

$$\mathbb{P}(C|A) = \frac{25}{51}$$

Să calculăm

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

$\mathbb{P}(A) = \frac{1}{4}$ → cazuri fav. să extragem prima carte

$$\mathbb{P}(A\cap B) = \frac{\cancel{13}\times 12}{\cancel{52}\times 51} = \frac{3}{51}$$ → a doua carte

$$\mathbb{P}(B|A) = \frac{3/51}{1/4} = \frac{12}{51}$$

Putem evalua și prob. $\mathbb{P}(A|B) = ?$

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)} = \frac{12}{51}$$

$\mathbb{P}(B) = 1/4$

În mod similar putem calcula $\mathbb{P}(A|C) = \frac{\mathbb{P}(A\cap C)}{\mathbb{P}(C)}$

$$\left.\begin{aligned}\mathbb{P}(A\cap C) &= \frac{\cancel{13}\times 25}{\cancel{52}\times 51} = \frac{25}{204}\\ \mathbb{P}(C) &= 1/2\end{aligned}\right\} \Rightarrow \mathbb{P}(A|C) = \frac{25}{102}$$

În general, $\mathbb{P}(A|B) \neq \mathbb{P}(B|A)$

Din def: $\mathbb{P}(A|B) = \dfrac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}$ pp că $\mathbb{P}(B)>0$

$$\Rightarrow \mathbb{P}(A\cap B) = \mathbb{P}(B)\cdot \mathbb{P}(A|B) = \mathbb{P}(A)\cdot \mathbb{P}(B|A)$$

(P) Pentru orice evenimente $A_1, A_2, \dots, A_n$ cu $\mathbb{P}(A_1\cap A_2\cap \dots A_{n-1})>$ avem

$$\mathbb{P}(A_1\cap A_2\cap \dots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\times \mathbb{P}(A_3|A_1\cap A_2)\times \dots \times \mathbb{P}(A_n|A_1\cap A_2\cap \dots \cap A_{n-1})$$

Obs!

$$\mathbb{P}(\underbrace{A_1\cap A_2}_{B}\cap A_3) = \mathbb{P}(A_3|\underbrace{A_1\cap A_2}_{B})\cdot \mathbb{P}(\underbrace{A_1\cap A_2}_{B}) = \mathbb{P}(A_3|A_1\cap A_2)\,\mathbb{P}(A_2|A_1)\,\mathbb{P}(A_1)$$

Exp! O familie are 2 copii.

a) Care este prob. ca cei 2 copii să fie de sex feminin știind că cel mai în vârstă este fată?

b) Care este prob. ——— știind că cel puțin / un copil este de sex feminin?

Sol:

Presupuneri: - nu putem avea decât sex feminin și masculin

$\mathbb{P}(B) = \mathbb{P}(F) = 1/2$

- genul unui copil nu influențează genul celuilalt

(nu luăm în calcul gemenii identici)

a) $\Omega = \{BB, BF, FB, FF\}$

$A = \{\text{ambii copii sunt } F\} = \{FF\}$

$B = \{\text{cel mai în vârstă este } F\} = \{FB, FF\}$

inf. suplimentară – un copil specific este F

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)} = \frac{1/4}{2/4} = 1/2$$

b) $C = \{\text{cel puțin unul dintre copii este } F\}$

$= \{BF, FB, FF\}$

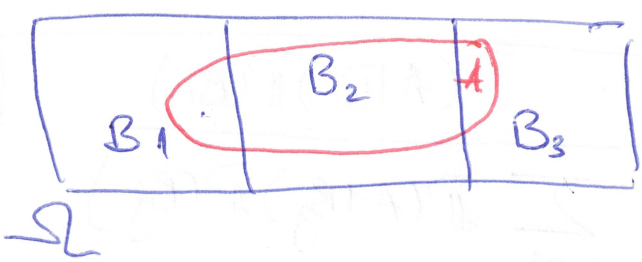
$$\mathbb{P}(A|C) = \frac{\mathbb{P}(A\cap C)}{\mathbb{P}(C)} = \frac{1/4}{3/4} = 1/3$$

c) Ne întâlnim în mod aleator cu unul dintre cei 2 și aflăm că este F. Care este prob. ca cei 2 copii să fie F?

Formula prob. totale:

Să presupunem că Ω este divizat în B_1, B_2, B_3 evenimente disjuncte 2 câte 2.

partiție



$$\begin{cases} B_1 \cup B_2 \cup B_3 = \Omega \\ B_1 \cap B_2 = \emptyset \\ B_2 \cap B_3 = \emptyset \\ B_1 \cap B_3 = \emptyset \end{cases}$$

$$\mathbb{P}(A) = \mathbb{P}(A\cap B_1) + \mathbb{P}(A\cap B_2) + \mathbb{P}(A\cap B_3)$$
$$= \mathbb{P}(A|B_1)\mathbb{P}(B_1) + \mathbb{P}(A|B_2)\mathbb{P}(B_2) + \mathbb{P}(A|B_3)\mathbb{P}(B_3)$$

ponderi

Fie A un eveniment cu $\mathbb{P}(A) > 0$, $\mathbb{P}(A) < 1$

$$\mathbb{P}(B) = \mathbb{P}(B|A)\mathbb{P}(A) + \mathbb{P}(B|A^c)\mathbb{P}(A^c)$$

(P) Dacă A și B sunt evenimente cu $0 < \mathbb{P}(B) < 1$ at.

$$\boxed{\mathbb{P}(A) = \mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|B^c)\mathbb{P}(B^c)}$$

Mai mult, dacă $B_1, B_2, \dots, B_n$ formează o partiție pe Ω cu $\mathbb{P}(B_i) > 0$ atunci

$$\boxed{\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A|B_i)\mathbb{P}(B_i)}$$

(T) (Formula lui Bayes)
Fie $(\Omega, \mathcal{F}, \mathbb{P})$ c.p, $A, B \in \mathcal{F}$ a.î. $\mathbb{P}(A) > 0$ și $\mathbb{P}(B) > 0$.

Atunci

$$\boxed{\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B)\mathbb{P}(B)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B)\cdot\mathbb{P}(B)}{\mathbb{P}(A|B)\cdot\mathbb{P}(B) + \mathbb{P}(A|B^c)\mathbb{P}(B^c)}}$$

Mai mult, dacă $B_1, B_2, \dots, B_n$ formează o partiție pe Ω cu $\mathbb{P}(B_i) > 0$ atunci

$$\boxed{\mathbb{P}(B_i|A) = \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\sum_{j=1}^{n} \mathbb{P}(A|B_j)\mathbb{P}(B_j)}}$$

<u>Exp3:</u> Avem o urnă în care sunt 5 bile roșii și 2 albastre. Extragem 2 bile una după cealaltă și vrem să calculăm prob. ca a 2-a bilă să fie roșie?

Sol: $\Omega = \{rr, ra, ar, aa\}$

R_1 – "prima bilă extrasă este de culoare roșie"

R_2 – "a doua bilă ——"

A_1 – "prima bilă extrasă albastră"

A_2 – "a doua bilă ——"

$\mathbb{P}(R_2) = ?$ $(5/7)$

$$\mathbb{P}(R_2) = \mathbb{P}(R_2 \mid R_1)\mathbb{P}(R_1) + \mathbb{P}(R_2 \mid A_1)\mathbb{P}(A_1)$$

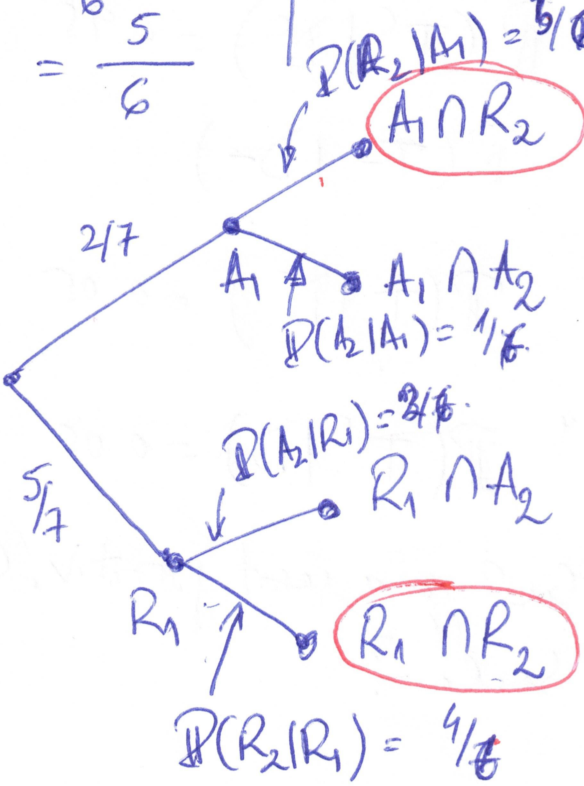
$\mathbb{P}(R_1) = \frac{5}{7}$

$\mathbb{P}(A_1) = 2/7.$

$\mathbb{P}(R_2 \mid R_1) = \frac{4}{6}$

$\mathbb{P}(R_2 \mid A_1) = \frac{5}{6}$

$$\Rightarrow \mathbb{P}(R_2) = \frac{4}{6} \times \frac{5}{7} + \frac{5}{6} \times \frac{2}{7} = \frac{30}{42} = \frac{5}{7}$$



$\mathbb{P}(R_2) = \mathbb{P}(A_1 \cap R_2) + \mathbb{P}(R_1 \cap R_2)$

$\mathbb{P}(A_1 \cap R_2) = 5/6 \times 2/7$

$\mathbb{P}(R_1 \cap R_2) = \frac{4}{6} \cdot \frac{5}{7}$

Exp: (Testăm pt o boală / afecțiune rară)

Să presupunem că frecvența afecțiunii / bolii în populație este de 1%.

Presupunem că efectuăm un test a cărui acuratețe este ridicată. (rată de "false positive" de 5%, "false negative" de 5%) 95% → sensitivitatea și specificitatea

D – "persoana are afecțiunea / boala" $\boxed{D+}$

T – "testul a ieșit pozitiv" $\boxed{T+}$

sensitivitatea (rata true pozitiv) $\mathbb{P}(T \mid D) = 0.95$

$\mathbb{P}(T+ \mid D+)$

specificitatea (rata true negative) $\mathbb{P}(T^c \mid D^c) = 0.95$

$\mathbb{P}(T- \mid D-)$

rata "false positive" $\mathbb{P}(T \mid D^c) = 0.05$

rata "false negative" $\mathbb{P}(T^c \mid D) = 0.05$

Pp că am efectuat testul și a ieșit pozitiv. Care este probabilitatea să avem boală?

$\mathbb{P}(D \mid T) = ?$

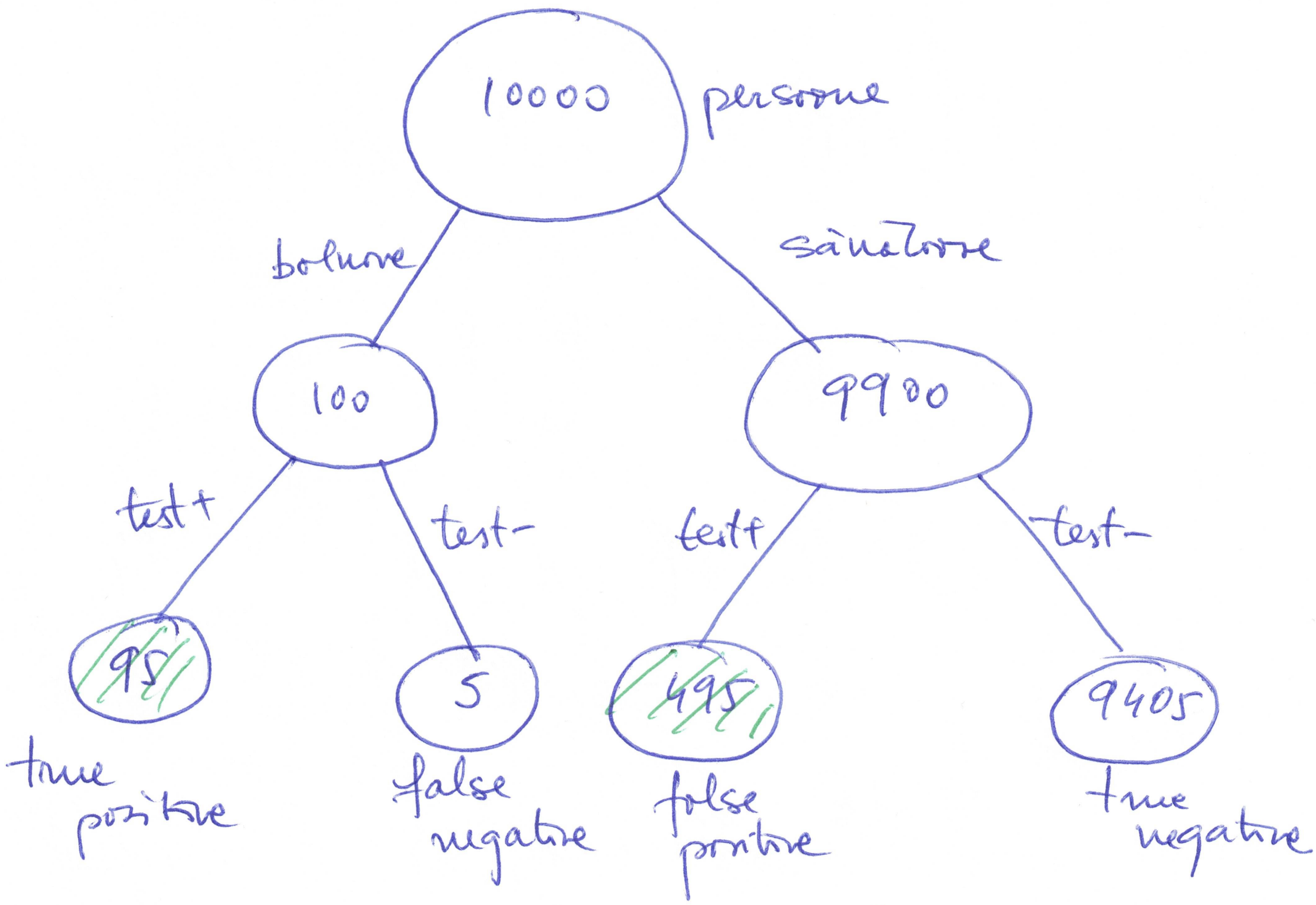
$$\mathbb{P}(D \mid T) = \frac{\mathbb{P}(T \mid D)\,\mathbb{P}(D)}{\mathbb{P}(T)} = \frac{\mathbb{P}(T \mid D)\,\mathbb{P}(D)}{\mathbb{P}(T \mid D)\,\mathbb{P}(D) + \mathbb{P}(T \mid D^c)\,\mathbb{P}(D^c)}$$

Din ipoteză $\mathbb{P}(D) = 0.01$

$\mathbb{P}(T|D) = 0.95$

$\mathbb{P}(T|D^c) = 1 - \mathbb{P}(T^c|D^c) = 0.05$

$$\mathbb{P}(D|T) = \frac{0.95 \times 0.01}{0.95 \times 0.01 + 0.05 \times 0.99} \simeq 0.16$$



$$\mathbb{P}(D|T) = \boxed{\frac{95}{95+495} \simeq 0.16}$$

b) Efectuăm testul a doua oară și avem tot un rsp +

Cât este acum șansa de îmbolnăvire?