# Curs 2/Partea 1: Probabilităţi condiţionate. Evenimente independente.

## 1.1 Probabilităţi condiţionate

Într-un experiment aleator producerea unui eveniment poate influenţa probabilitatea producerii altor evenimente. Pentru a cuantifica această influenţă se defineşte noţiunea de probabilitate a unui eveniment, condiţionată de producerea altui eveniment.

**Exemplul 1**. 9 studenţi, 4 băieţi şi 5 fete, participă la o competiţie pentru o bursă Erasmus+. Cristina este una dintre cele 5 concurente, şi are şansa de 1/9 să ia interviul pentru bursă (să fie admisă în programul Erasmus+). După concurs, înainte de a se anunţa oficial rezultatele, se află că a luat interviul o fată. Cum au participat 5 fete, şansa Andreei de a primi bursa Erasmus+ este acum de 1/5. Extra informaţia că concursul a fost câştigat de o fată a schimbat probabilitatea de succes a Andreei.

Dacă notăm cu $A$ evenimentul "Cristina ia interviul" şi cu $B$ evenimentul "o fată ia interviul", atunci probabilitatea evenimentului $A$ condiţionată de $B$ se notează $P_B(A)$ sau $P(A|B)$ şi reprezintă probabilitatea ca evenimentul $A$ să se producă, ştiind că $B$ s-a produs.

**Definiţia 1.1.1** *Probabilitatea unui eveniment $A$, condiţionată de evenimentul $B$, cu $P(B) \neq 0$, este prin definiţie*

$$P_B(A) = \frac{P(A \cap B)}{P(B)} \tag{1}$$

În exemplul precedent $P(A \cap B) = P(A) = 1/9$, iar $P(B) = 5/9$, deci $P_B(A) = (1/9)/(5/9) = 1/5$.

**Propoziţia 1.1.1** *Fie $(\Omega, \mathcal{K}, P)$ un câmp de probabilitate, $B \in \mathcal{K}$, un eveniment de probabilitate nenulă, $P(B) \neq 0$. Atunci $P_B$ este o probabilitate pe $\mathcal{K}$, adică:*
*1)* $0 \leq P_B(A) \leq 1$, $\forall A \in \mathcal{K}$;
*2)* $P_B(\Omega) = 1$;

*3) Pentru orice două evenimente, $A_1, A_2$, mutual exclusive, $A_1 \cap A_2 = \emptyset$, probabilitatea reuniunii condiţionată de $B$ este suma probabilităţilor condiţionate:*

$$P_B(A_1 \cup A_2) = P_B(A_1) + P_B(A_2) \tag{2}$$

**Demonstraţie**: 1) şi 2) sunt evidente.

3. Fie $A_1, A_2 \in \mathcal{K}$, şi $A_1 \cap A_2 = \emptyset$. Atunci şi evenimentele $A_1 \cap B, A_2 \cap B$ sunt mutual exclusive şi:

$$\begin{aligned} &P_B(A_1 \cup A_2) = \frac{P((A_1 \cup A_2) \cap B)}{P(B)} = \\ &\frac{P((A_1 \cap B) \cup (A_2 \cap B))}{P(B)} = \frac{P(A_1 \cap B) + P(A_2 \cap B)}{P(B)} = \\ &P_B(A_1) + P_B(A_2) \end{aligned} \tag{3}$$

□

### Formula de condiţionare iterată

Fie $A_1, A_2, \ldots, A_n$, $n$ evenimente într-un experiment. Exploatând formula de calcul a probabilităţii condiţionate, $P(A_2|A_1) = \dfrac{P(A_1 \cap A_2)}{P(A_1)}$ deducem recursiv modalitatea de calcul a probabilităţii intersecţiei a unui număr arbitrar de evenimente din cele $n$:

$$P(A_1 \cap A_2) = P(A_1)P(A_2|A_1)$$

$$P(A_1 \cap A_2 \cap A_3) = P((A_1 \cap A_2) \cap A_3) = P(A_1 \cap A_2)P(A_3|A_1 \cap A_2) =$$

$$=P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2)$$

$$\vdots$$

$$\begin{aligned} &P(A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_k) = P(A_1)\, P(A_2|A_1)\, P(A_3|A_1 \cap A_2) \cdots \\ &\cdots P(A_k|A_1 \cap A_2 \cap \cdots \cap A_{k-1}) \end{aligned} \tag{4}$$

**Exemplul 2**. O cutie conţine 20 chip-uri de memorie din care 5 sunt defecte. Se aleg 3 la întâmplare. Să se calculeze a) probabilitatea ca toate trei să fie defecte; b) exact un chip să fie defect.

Rezolvare: Notăm cu $E_i$ evenimentul ”al i-lea chip ales este defect”, $i = \overline{1,3}$. La a) trebuie calculată probabilitatea:

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

Dar $P(E_1) = 5/20 = 1/4$. $P(E_2|E_1)$ este probabilitatea ca al doilea chip extras să fie defect ştiind că şi primul a fost defect. După prima alegere au rămas 4 chipuri defecte din 19, deci $P(E_2|E_1) = 4/19$. Analog $P(E_3|E_1 \cap E_2) = 3/18$ şi deci:

$$P(E_1 \cap E_2 \cap E_3) = \frac{5}{20}\frac{4}{19}\frac{3}{18}$$

b) Fie $D_i$ evenimentul "al i-lea chip ales este defect" şi $B_i$ evenimentul "al i-lea chip ales este bun". Prin urmare trebuie să calculăm:

$$P((D_1 \cap B_2 \cap B_3) \cup (B_1 \cap D_2 \cap B_3) \cup (B_1 \cap B_2 \cap D_3)) = P(D_1 \cap B_2 \cap B_3) + \\ + P(B_1 \cap D_2 \cap B_3) + P(B_1 \cap B_2 \cap D_3)$$

Fiecare din probabilităţile din membrul drept al ultimei relaţii se calculează aplicând formula de condiţionare iterată.

**Exemplul 3**. Google Translate a folosit până de curând un algoritm numit *Statistical Machine Translation* (prescurtat SMT), care presupune crearea pentru fiecare limbă a unui model probabilist al limbii respective. Crearea acestui model constituie, în limbaj de machine learning, etapa "de învăţare a limbii respective" şi anume: dintr-o bază de texte în limba respectivă, stocate în format electronic, se estimează probabilităţile de a întâlni într-o (viitoare) traducere o anumită frază *f*. Pentru a estima probabilităţile $P(f)$, frazele *f*, care sunt stringuri de cuvinte, se divid în substringuri de $n$ cuvinte, numite $n$-grame. Particularizând, avem unigram ($n = 1$), bigram ($n = 2$), trigram ($n = 3$). O frază se notează astfel:

$$f = w_1 w_2 \ldots w_I.$$

În contextul maşinii de tradus, fraza $f$ este privită ca evenimentul ca într-un text din limba respectivă să se întâlnească succesiunea de cuvinte $w_1 w_2 \ldots w_I$.

Pentru a estima probabilitatea $P(f)$ din textele din corpus (denumirea oficială a bazei de texte dintr-o limbă), teoretic se procedează astfel: dacă $w$ este un unigram (adică un cuvânt), atunci probabilitatea de a întâlni cuvântul $w$ este estimată de:

$$p(w) = \frac{\text{de câte ori apare cuvântul } w \text{ în baza de texte în limba analizată}}{\text{numărul total de cuvinte din baza de texte în limba respectivă}}.$$

Probabilitatea ca, într-un text, cuvântul $w_2$ să urmeze după cuvântul $w_1$ este:

$$p(w_2|w_1) = \frac{p(w_1 w_2)}{p(w_1)} = \frac{\text{de câte ori apare în baza de texte succesiunea de cuvinte } w_1 w_2}{p(w_1)}.$$

Apoi

$$p(w_3|w_1 w_2) = \frac{\text{de câte ori apare în baza de texte stringul "} w_1 w_2 w_3 \text{"}}{\text{de câte ori apare stringul "} w_1 w_2 \text{"}} \quad \text{etc.}$$

Mai sus s-a precizat că *teoretic* probabilităţile de apariţie a $n$-gramelor se estimează ca frecvenţe de apariţie. În practică, se aplică o serie de ajustări pentru a elimina din calcule $n$-gramele de probabilităţi nule etc.

Aplicând formula condiţionării iterate, rezultă că probabilitatea (teoretică) ca fraza $f$ să fie întâlnită într-o cerere de traducere este

$$p(f) = p(w_1)p(w_2|w_1) \cdots p(w_{I-1}|w_1 w_2 \ldots w_{I-2}) p(w_I|w_1 w_2 \ldots w_{I-1}),$$

unde $\mathtt{p(w_I|w_1w_2\ldots w_{I-1})}$ este probabilitatea să întâlnim cuvântul $w_I$ după succesiunea de cuvinte $w_1w_2\ldots w_{I-1}$. Înlocuind probabilităţile din membrul drept cu estimatorii deduşi din baza de texte, după ce au fost ajustaţi conform unor reguli, se obţine probabilitatea estimată, adică învăţată de maşina de tradus din baza de texte.

După numeroase teste s-a ajuns la concluzia că nu e nevoie să se estimeze probabilităţi de tipul $p(w_k|w_1w_2\ldots w_{k-1})$ pentru $k > 4$, deoarece în baza de texte există puţine `k`-grame, $\mathtt{k} > \mathtt{4}$, care să se repete în diferite texte şi cel mai adesea se repetă bigrame, trigrame.

În noiembrie 2016, Google a anunţat că Google Translate va deveni un sistem de traducere neuronal, numit Google Neural Machine Translation (GNMT), bazat pe reţele neuronale (artificiale): o frază ce se doreşte a fi tradusă va fi considerată ca un întreg şi nu bucată cu bucată, cum s-a întâmplat în cazul sistemului SMT. Pentru a vedea cum funcţionează sistemul de tradus GNMT vezi https://arxiv.org/pdf/1609.08144.pdf.

## 1.2 Evenimente independente

După ce am introdus noţiunea de probabilitate condiţionată, ne întrebăm dacă producerea unui eveniment influenţează neapărat şansa de producere a altui eveniment. Intuitiv, ne aşteptăm ca unele evenimente să fie şi independente, adică să nu fie influenţată şansa de producere a unuia de producerea celuilalt. În cele ce urmează caracterizăm matematic independenţa evenimentelor.

Fie $A$ şi $B$ două evenimente într-un experiment, de probabilităţi nenule, $P(A), P(B) \neq 0$. Din formula probabilităţii condiţionate $P_B(A) = \dfrac{P(A\cap B)}{P(B)}$ sau $P_A(B) = \dfrac{P(A\cap B)}{P(A)}$, rezultă că:

$$P(A\cap B) = P_B(A)P(B) = P_A(B)P(A). \tag{5}$$

Dacă $P_B(A)$=P(A) şi $P_A(B) = P(B)$, atunci realizarea evenimentului $B$ nu influenţează probabilitatea evenimentului $A$ şi reciproc. În acest caz $P(A\cap B) = P(A)P(B)$.

Pe de altă parte $P(A\cap B) = P(A)P(B)$ implică ţinând seama de (5) că $P_B(A) = P(A)$ şi $P_A(B) = P(B)$. Acest fapt justifică definiţia:

**Definiţia 1.2.1** *Fie $(\Omega, \mathcal{K}, P)$ un câmp de probabilitate. Evenimentele $A$, $B \in \mathcal{K}$ de probabilităţi nenule, cu proprietatea că $P(A\cap B) = P(A)P(B)$ se numesc evenimente independente. Mai general evenimentele $A_1, A_2, \ldots, A_n$ cu proprietatea că:*

$$P(A_{i_1}\cap A_{i_2}\cap\cdots\cap A_{i_k}) = P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k}), \tag{6}$$

*pentru orice $k = \overline{2,n}$ şi indicii $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ se numesc evenimente independente.*

Remarcăm că independenţa a mai mult de două evenimente necesită verificarea mai multor relaţii. De exemplu evenimentele $A_1, A_2, A_3$ sunt independente dacă sunt satisfăcute relaţiile:

$$P(A_i\cap A_j) = P(A_i)P(A_j),\ \forall\ 1\leq i<j\leq 3,\ \text{(adică } C_3^2 = 3 \text{ relaţii)}$$

şi

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2) \cdot P(A_3)$$

În cazul general a $n \geq 2$ evenimente, $A_1, A_2, \ldots, A_n$, avem de verificat:

$$C_n^2 + C_n^3 + \cdots + C_n^n = (C_n^0 + C_n^1 + C_n^2 + C_n^3 + \cdots + C_n^n) - C_n^0 - C_n^1 = 2^n - (1+n)$$

relaţii de forma (6) pentru a testa independenţa evenimentelor.

A nu se confunda evenimentele mutual exclusive cu evenimentele independente!!!!

• Două evenimente $A$, $B$ sunt mutual exclusive dacă $A \cap B = \emptyset$, adică producerea lor simultană este imposibilă, şi deci $P(A \cap B) = 0$.

• Evenimentele $A, B$ sunt independente dacă $P(A \cap B) = P(A)P(B)$.

O altă observaţie este că dacă $A, B, C$ sunt trei evenimente, două câte două independente, atunci $A, B, C$ nu sunt neapărat independente, adică relaţiile:

$$P(A \cap B) = P(A)P(B), \quad P(A \cap C) = P(A)P(C), \quad P(B \cap C) = P(B)P(C)$$

nu implică obligatoriu:

$$P(A \cap B \cap C) = P(A)P(B)P(C).$$

**Propoziţia 1.2.1** *Dacă $A$ şi $B$ sunt evenimente independente atunci a) $\complement A$ şi $B$; b) $A$ şi $\complement B$; c) $\complement A$ şi $\complement B$ sunt independente.*

**Demonstraţie**: Arătăm doar că a) are loc, celelalte cazuri demonstrându-se similar. În demonstraţie ţinem seama că $A$ şi $B$ sunt independente, deci $P(A \cap B) = P(A)P(B)$ şi de faptul că pentru orice eveniment $B$, de probabilitate nenulă, $P_B$ este funcţie de probabilitate.

$$\begin{aligned}
P(\complement A \cap B) &= P_B(\complement A)P(B) = (1 - P_B(A))P(B) = \left(1 - \frac{P(A \cap B)}{P(B)}\right)P(B) \\
&= \left(1 - \frac{P(A)P(B)}{P(B)}\right)P(B) = (1 - P(A))P(B) = P(\complement A)P(B).
\end{aligned} \tag{7}$$

□

**Observaţia 1.2.1** *Propoziţia de mai sus este adevărată pentru un număr arbitrar de evenimente independente $A_1, A_2, \ldots, A_n$, în sensul că dacă $A_1$, $A_2, \ldots, A_n$ sunt evenimente independente, atunci şi evenimentele $B_1, B_2, \ldots$, $B_n$, sunt independente, unde:*

$$B_{i_1} = A_{i_1}, B_{i_2} = A_{i_2}, \ldots, B_{i_k} = A_{i_k}, B_j = \overline{A}_j,$$

*unde $i_1, i_2, \ldots, i_k$ sunt elemente distincte din $\{1, 2, \ldots, n\}$ şi $j \in \{1, 2, \ldots, n\} \setminus \{i_1, i_2, \ldots, i_k\}, 1 \leq k \leq n$.*

**Exemplul 4**. a) Fie $A$ şi $B$ două evenimente independente într-un experiment. Ştiind că $P(A) = 0.25, P(B) = 0.3$ să se calculeze $P(A \cup B)$, $P_B(\complement A)$ şi $P(A \cap \complement B)$.

$$\begin{aligned} &P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - P(A)P(B) = \\ &= 0.25 + 0.3 - 0.075 = 0.475 \end{aligned}$$

$P_B(\complement A) = 1 - P_B(A)$. Evenimentele $A$ şi $B$ fiind independente $P_B(A) = P(A) = 0.25$. Deci $P_B(\complement A) = 1 - 0.25 = 0.75$.

Cum şi evenimentele $A$ şi $\complement B$ sunt independente, rezultă că: $P(A \cap \complement B) = P(A) \cdot P(\complement B) = P(A)(1 - P(B)) = 0.25 \cdot 0.7 = 0.175$.

**Exemplul 5**. Un proiect constă din trei sarcini independente şi probabilităţile ca aceste sarcini să fie îndeplinite la timp sunt, respectiv: 0.9, 0.8, 0.85. Să se calculeze probabilitatea:
a) ca toate cele trei sarcini să fie îndeplinite la timp.
b) ca primele două să fie îndeplinite la timp, iar a treia nu.
c) cel puţin una din sarcini să fie îndeplinită la timp.

**Rezolvare:** Fie $A_i$, $i = 1, 2, 3$, evenimentul "sarcina i este îndeplinită la timp".
a) $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3) = (0.9)(0.8)(0.85)$;
b) Evenimentul a cărui probabilitate se cere la b) este $A_1 \cap A_2 \cap \complement A_3$. Evenimentele $A_1, A_2, A_3$ fiind independente, sunt independente şi evenimentele: $A_1, A_2, \complement A_3$. Deci

$$P(A_1 \cap A_2 \cap \complement A_3) = P(A_1)P(A_2)P(\complement A_3) = 0.9, 0.8(1 - 0.85)$$

c) $P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3) = 0.9 + 0.8 + 0.85 - (0.9)(0.8) - (0.9)(0.85) - (0.8)(0.85) + (0.9)(0.8)(0.85)$.

Mai simplu,

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) &= 1 - P(\complement(A_1 \cup A_2 \cup A_3)) = 1 - P(\overline{A_1} \cap \overline{A}_2 \cap \overline{A}_3) = \\ &= 1 - P(\overline{A}_1)P(\overline{A}_2)P(\overline{A}_3) \\ &= 1 - (1 - P(A_1))(1 - P(A_2))(1 - P(A_3)) \end{aligned}$$