TEORIA PROBABILITĂȚILOR

- Statistica prelucrează date **reale, concrete**.
- Metodele statisticii derivă însă din modele matematice **abstracte**.
- **Teoria probabilităților** este suportul abstract fundamental al statisticii inferențiale.
- Teoria probabilităților se ocupă de *fenomenele cu comportare aleatoare*.

EXPERIMENT ALEATOR

- Un experiment aleator este un act care satisface următoarele condiții:
  - toate situațiile finale distincte posibile sunt cunoscute *a priori*;
  - pentru oricare repetare particulară a experimentului, situația finală nu este cunoscută *a priori*;
  - experimentul poate fi repetat în condiții identice
- Nu se cunoaşte dinainte rezultatul, dar se cunosc toate rezultatele posibile.
- Exemplu. La aruncarea unui zar sunt 6 rezultate elementare posibile.

EVENIMENTE

Toate situațiile finale legate de un experiment aleator şi despre care, după efectuarea experimentului, putem spune cu certitudine că s-au produs sau nu.

- **Un eveniment** este o submulțime a mulțimii tuturor rezultatelor (situațiilor) finale elementare posibile.
  - Exemplu. Evenimentul ca un zar să arate, după aruncare, un număr par.
    - Nu se poate prevedea rezultatul unei singure repetări a unui eveniment aleator (factori!);
    - se poate prevedea însă **structura rezultatelor**.

EVENIMENT SIGUR – EVENIMENT IMPOSIBIL

- **Evenimentul sigur S**: un eveniment care se realizează cu certitudine la fiecare repetare a experimentului (*zar: "*$1 \leq x \leq 6$*"*).
- **Evenimentul imposibil Φ**: un eveniment care cu certitudine nu se produce la nici o repetare a experimentului (*zar: "*$x<1$ *sau* $7 \leq x$*"*).
  - Aceste două evenimente se ataşează oricărui experiment aleator.
  - Fiecare se realizează dacă nu se realizează celălalt.

EVENIMENTE CONTRARE

- Exemple:
  - {1,3} şi {2,4,5,6}
  - S şi Φ.

- Dat un eveniment E, lui îi corespunde **evenimentul contrar** $\overline{E}$, a cărui producere înseamnă *prin definiție* nerealizarea lui E.
- A contrar lui B ⇔ B contrar lui A.
- Proprietăți: $\overline{\overline{A}} = A$; $\overline{S} = \Phi$; $\overline{\Phi} = S$.

EVENIMENTE COMPATIBILE

- Evenimentele A şi B sunt **compatibile** dacă *se pot produce simultan*.
  - Există rezultate finale favorabile şi lui A, şi lui B.
  - Exemplu: La aruncarea unui zar, “par” şi “prim”.
- Evenimentele A şi B sunt **incompatibile** dacă *nu se pot produce simultan*.
  - Ca mulțimi de evenimente elementare
    - compatibile: nedisjuncte; incompatibile: disjuncte
- Generalizare pentru orice n≥2
    - compatibile global sau două câte două.

EVENIMENT IMPLICAT

- Evenimentul A **implică** evenimentul B (B **este implicat** de A) dacă *B se produce ori de câte ori se produce A*.
- Ca mulțimi de rezultate finale elementare, A este inclus în B.
- *Exemplu*. La aruncarea unui zar
  - A “putere nenulă a lui 2”; B “par” $\{2,4\} \subset \{2,4,6\}$
  - A “impar”; B “prim” $\{1,3,5\} \subset \{1,2,3,5\}$

OPERAȚII CU EVENIMENTE

- **Reuniunea**. Date evenimentele A şi B, evenimentul reuniune $A \cup B$ se produce atunci când *se produce cel puțin unul dintre evenimentele A , B*.
  - $\{1,2,5\} \cup \{3,4,5\} = \{1,2,3,4,5\}$.
- **Intersecția**. Date evenimentele A şi B, evenimentul intersecție $A \cap B$ se produce atunci când *se produc simultan şi A şi B*.
  - "par" $\cap$ "prim" = $\{2\}$ (compatibile)
  - "par" $\cap$ "impar" = Φ (incompatibile).

SPAȚIUL DE SELECȚIE AL UNUI EXPERIMENT

- Spațiul de selecție (evenimentelor) unui experiment este *o mulțime* cu proprietatea că
  - orice *eveniment elementar* rezultat în urma experimentului
  - corespunde *unui singur element* al acestei mulțimi.
- Spațiul de selecție se schimbă în funcție de punctul de vedere din care este privit experimentul
- Exemplu. Aruncarea a două monezi (o repetare):
  - {(BB),(BS),(SB),(SS)} – ce fețe apar
  - {(2,0),(1,1),(0,2)} – de câte ori apare fiecare față
  - {ID, DIF} – sunt sau nu identice fețele

FRECVENȚĂ

- Fie un experiment şi A un eveniment ataşat.
- Se repetă experimentul de **n** ori, de **a** ori producându-se evenimentul A
  - iar de **n-a** ori, evenimentul contrar lui A.
- Numărul $f_n = \frac{a}{n}$ este frecvența relativă a evenimentului A în instanțierea respectivă a experimentului.
- De la o instanțiere la alta a experimentului, frecvența variază: $0 \leq \mathbf{a} \leq \mathbf{n}$; $0 \leq f_n \leq 1$.
  - Pentru multe fenomene, când **n** creşte, f_n se apropie de o constantă.

DEFINIȚIA CLASICĂ A PROBABILITĂȚII (1)

– Laplace, sec. al-XIX-ea.

<u>Evenimente egal probabile</u>

Dacă **n** evenimente posibile într-un experiment

– ce nu pot fi descompuse în evenimente mai simple

- au *aceeaşi şansă de a se produce*,
- atunci ele sunt *egal probabile* (*equally likely*)
- iar probabilitatea fiecăruia este 1/**n**.

DEFINIȚIA CLASICĂ A PROBABILITĂȚII (2)

Probabilitatea evenimentului A

– este egală cu *raportul* dintre

– numărul de evenimente egal probabile ce definesc A

– şi numărul total de evenimente egal probabile.

Exemple.

– A **-** evenimentul imposibil: P(A) = 0.

– A – evenimentul sigur: P(A) = 1.

• Monotonie: $0 \leq P(A) \leq 1$.

DEFINIȚIA CLASICĂ A PROBABILITĂȚII

Dacă în urma efectuării unui experiment

- pot rezulta **n** evenimente elementare egal probabile
- şi dacă **m** dintre ele definesc evenimentul A
- atunci probabilitatea evenimentului A este: P(A) = **m** / **n** ("numărul de cazuri favorabile împărțit la numărul total de cazuri").

$$P(\overline{A}) = \frac{n - m}{n} = 1 - P(A)$$

- **Exemplu**. "Cel puțin un 1" în două aruncări ale unui zar: P(A) = (36-25) / 36 = 11 / 36 .

CRITICA DEFINIȚIEI CLASICE A PROBABILITĂȚII

- Egal-probabilitatea nu există decât în experimente artificiale (extrageri de bile de culori diferite)
  - zar sau monedă perfecte nu există.
- Dacă există o infinitate de cazuri elementare, atunci toate probabilitățile sunt practic 0.
- La fenomene sociale definiția nu este aplicabilă, când nu se cunoaşte numărul de cazuri.
  - Exemplu. Frecvența în populație anul viitor depinde de numărul *necunoscut* de persoane de peste un an.

DEFINIȚIA CLASICĂ A PROBABILITĂȚII (5)

- **Teorema 1**. Dacă $A_1, \ldots, A_k$ sunt evenimente global incompatibile, atunci:

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

- **Demonstrație**. A_1 definit de n_1 evenimente elementare, …, A_k de n_k evenimente elementare; $A = \cup A_i$ este definit (disjuncte!) de $\sum n_i$ evenimente elementare.

DEFINIȚIA CLASICĂ A PROBABILITĂȚII (6)

- **Teorema 2**. Dacă A şi B sunt evenimente într-un spațiu de selecție *finit* S, atunci:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- **Demonstrație**. P(A) + P(B) conține suma probabilităților evenimentelor din $A \cap B$ *de două ori*.
- **Teorema 3**. Dacă $A_1, \ldots, A_k$ sunt evenimente global incompatibile şi dacă $\cup A_i = S$, atunci: $\sum P(A_i) = 1$.
- **Demonstrație**. Se aplică teorema 1, ținând seama de faptul că $P(\cup A_i) = P(S) = 1$.

DEFINIȚIA PROBABILITĂȚII BAZATĂ PE FRECVENȚE

- Fie un experiment aleator şi fie un eveniment A care se produce de m(A) ori la m repetări ale experimentului.
- **Definiție**. Probabilitatea evenimentului A se defineşte prin: $P(A)=\lim_{m\to\infty}\frac{m(A)}{m}$
- Există limita ? (legea numerelor mari)
- Definiția este consistentă numai în situația (cvasi-imposibilă) - când condițiile rămân identice la toate repetările experimentului aleator.

DEFINIȚIA AXIOMATICĂ A PROBABILITĂȚII

- **Spațiul de selecție** al unui experiment aleator se consideră a fi o mulțime de elemente astfel încât oricărui rezultat elementar al experimentului să-i corespundă un singur element (**punct**) al mulțimii.

– Considerăm în continuare doar spații *finite*.

Definiții.

- 1.- Orice mulțime de puncte este un **eveniment**.
- 2.- Un eveniment se nume**şte elementar dacă** are cardinalitate 1.
- 3.- Evenimentul A **s-a produs** dacă rezultatul experimentului este un punct din mulțimea ce defineşte A.
  - Evenimentul imposibil nu conține nici un punct.
  - Orice alt eveniment este o reuniune de evenimente elementare.

PROBABILITATE – AXIOME (1)

- Fie spațiul de selecție $S=\{e_1,\ldots,e_n\}$.
- Asociem fiecărui eveniment un număr numit probabilitate: $P : 2^S \to \mathbf{R}$

Axiome.

I. Probabilitatea asociată oricărui eveniment este mai mare sau egală cu 0.

II. Probabilitatea asociată întregului spațiu de selecție (evenimentul sigur) este 1.

III. Dacă evenimentele A şi B sunt incompatibile ($A\cap B=\Phi$), atunci:

$$\mathbf{P(A\cup B) = P(A)+P(B).}$$

PROBABILITATE – AXIOME (2)

- Axiomele probabilității sunt verificate de:
  - definiția clasică
  - definiția cu frecvențe
  - definiția subiectivă
    - unde P(A) este gradul individual de încredere că A se va produce.

CONSECINȚE ALE AXIOMELOR

1.- $\mathbf{P(\Phi) = 0}$ (Φ - evenimentul imposibil)

Demonstrație.

$$1 = P(S) = P(S \cup \Phi) = P(S) + P(\Phi) = 1 + P(\Phi)$$

2.- $\forall A$, $A \neq \Phi$, $A = \cup_i \{e_i\}$, e_i evenimente elementare: $\mathbf{P(A) = \cup_i P(e_i)}$

Demonstrație. Prin inducție.

3.- $\mathbf{P(E) + P(\overline{E}) = 1}$

4.- $$\mathbf{P\left(\bigcup_{i=1}^{k} E_i\right) = \sum_{i=1}^{k} P(E_i) - \sum_{i<j} P(E_i \cap E_j) + \sum_{i<j<k} P(E_i \cap E_j \cap E_k) - \ldots}$$

$$\mathbf{+ (-1)^{k-1} P(E_1 \cap E_2 \cap \ldots \cap E_k)}$$

PROBABILITĂȚI CONDIȚIONATE (1)

- Dacă avem informații despre producerea evenimentului B, le putem folosi pentru a calcula probabilitatea producerii, în aceste condiții, a evenimentului A.

- **Definiție**. Probabilitatea evenimentului A, dat faptul că B are loc ($P(B) \neq 0$), se numeşte **probabilitate condiționată a lui A, dat B**:

$$P(A/B) = P(A \cap B) / P(B)$$

  – Intuitiv: B devine noul spațiu de selecție, A mai putându-se realiza numai prin evenimente din $A \cap B$.

PROBABILITĂȚI CONDIȚIONATE (2)

- Se aruncă două zaruri. Primul dintre ele arată 4. Care este probabilitatea ca suma celor două zaruri să fie 6?
- E: “suma este 6”; F: “primul zar arată 4”

  ($P(F) \neq 0$).

I. Fără informația suplimentară:

  Suma 6: (1,5), (2,4), (3,3), (4,2), (5,1)

  $P(E) = 5/36 < 1/6$.

II. Cu informația “primul zar arată 4”:

  $P(E/F) = P(E \cap F) / P(F) = (1/36) / (6/36) = 1/6$

  $(>5/36)$

PROBABILITĂȚI CONDIȚIONATE (3)

- Trei bărbați îşi aruncă pălăriile într-o cameră întunecoasă, apoi aleg la întâmplare câte o pălărie din cameră. Care este probabilitatea ca nici unul dintre cei trei să nu-şi nimerească propria pălărie?
  - E: “nici unul nu alege propria pălărie”.
  - E^c: “cel puțin unul alege propria pălărie”.
  - E_i: “persoana i alege pălăria proprie”.
  - $P(E_i)=1/3$ – evident;
  - $P(E_i \cap E_j)=P(E_i)\cdot P(E_j/E_i)=(1/3)\cdot(1/2)=1/6,\ i\neq j$;
  - $P(E_1 \cap E_2 \cap E_3)=P(E_1 \cap E_2)\cdot P(E_3/E_1 \cap E_2)=(1/6)\cdot 1$
  - $P(E^c)=P(E_1 \cup E_2 \cup E_3)=3\cdot(1/3)-3\cdot(1/6)+1/6=2/3$
  - $P(E)=1/3$.

EVENIMENTE INDEPENDENTE (1)

- **Definiție**. Evenimentele A şi B se numesc **independente** dacă apariția unuia nu influențează *probabilitatea* de apariție a celuilalt: $P(A/B) = P(A)$ şi $P(B/A) = P(B)$.
- **Teoremă**. Dacă A şi B sunt evenimente independente, atunci $P(A \cap B)=P(A) \cdot P(B)$.
- **Definiție**. Evenimentele $A_1, \ldots, A_k$ sunt **global independente** dacă oricum s-ar alege p evenimente dintre cele k ($p \leq k$), oricare p-1 dintre acestea nu influențează probabilitatea celuilalt.

EVENIMENTE INDEPENDENTE (2)

- E: “suma celor două zaruri este 6”
- F: “primul zar arată 4”
- G: “suma zarurilor este 7”
- **E şi F nu sunt independente!**
  - $P(E)\cdot P(F) = (5/36)\cdot(1/6) = 5/216 < 1/36 = P(E\cap F)$
  - Intuitiv, valoarea primului zar poate face ca E să fie fals – dacă ea este 6.
- **G şi F sunt independente.**
  - $P(G \cap F) = 1/36 = (6/36)\cdot(6/36) = P(G)\cdot P(F)$
  - Indiferent de valoarea primului zar, $P(G) \neq 0$.

FORMULA LUI BAYES (1)

- **Teoremă (înmulțirea probabilităților dependente).**

Dacă $A_1, \ldots, A_k$ sunt evenimente pentru care $P(A_1 \cap \ldots \cap A_k) \neq 0$, atunci: $\mathbf{P(A_1 \cap \ldots \cap A_k) = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1 \cap A_2) \cdot \ldots \cdot P(A_k/A_1 \cap \ldots \cap A_{k-1})}$

  - **Demonstrație**: se aplică definiția.
  - Pentru evenimente independente, $P(A_1 \cap \ldots \cap A_k) = P(A_1) \cdot \ldots \cdot P(A_k)$

- **Teoremă (formula probabilității totale).** Dacă $A_1, \ldots, A_k$ realizează o partiție a spațiului de selecție S şi dacă X este un eveniment din S ($X \subseteq S$), atunci:

$$P(X) = P(A_1) \cdot P(X/A_1) + \ldots + P(A_k) \cdot P(X/A_k)$$

(interpretare grafică)

FORMULA LUI BAYES (2)

- Fie evenimentele $A_1,\ldots, A_k$ care realizează o partiție a spațiului de selecție S;
- fie X este un eveniment din S ($X \subseteq S$);
- se cunosc probabilitățile "*a priori*": $P(A_1),\ldots,P(A_k)$, probabilitățile condiționate $P(X/A_1),\ldots,P(X/A_k)$.
- Se efectuează experimentul şi se produce X.
- Să se determine $P(A_1/X),\ldots,P(A_k/X)$ (probabilitățile "*a posteriori*").

FORMULA LUI BAYES (3)

- **Teoremă (formula lui Bayes)**.

$$\mathbf{P(A_i/X) = \frac{P(A_i)\cdot P(X/A_i)}{\sum_{j=1}^{k} P(A_j)\cdot P(X/A_j)}}$$

- **Demonstrație**.

  $P(A_i)\cdot P(X/A_i) = P(X)\cdot P(A_i/X) = P(A_i \cap X)$ şi se ține seama de formula probabilității totale.

FORMULA LUI BAYES – EXEMPLE (1)

- **Instanțierea.** Din trei prizonieri, unul, ales la întâmplare fără ştirea lor, va fi executat, iar ceilalți doi vor fi eliberați. Unul dintre prizonieri cere gardianului să-i indice pe unul dintre ceilalți doi, care va fi eliberat (oricum, cel puțin unul dintre cei doi va fi eliberat). Dacă gardianul îi răspunde, capătă prizonierul mai multă informație decât avea deja?
- E_i: “prizonierul i este ales pentru execuție”. $P(E_1) = P(E_2) = P(E_3) = \mathbf{1/3}$.
- F_i: “prizonierul i este eliberat”. $P(F_1) = P(F_2) = P(F_3)$

$$P(E_1/F_2) = \frac{P(F_2/E_1)\cdot P(E_1)}{\sum_{i=1}^{3} P(F_2/E_i)\cdot P(E_i)} = \frac{1\cdot(1/3)}{1/3+0+1/3} = \frac{1}{2} > \frac{1}{3}$$

FORMULA LUI BAYES – EXEMPLE (2)

- **E-mail**. Fie o cutie e-mail cu trei foldere şi fie α_i probabilitatea de a găsi un mesaj dacă acesta se află în folderul *i* şi folderul este examinat superficial. Presupunând că mesajul a fost căutat în folderul *1* şi nu a fost găsit, care este probabilitatea ca mesajul să se afle totuşi în folderul *1* ?
- F_i : "mesajul se află în folderul *i*".
- E: "mesajul nu a fost găsit în folderul *1*".

$$P(F_1/E) = \frac{P(E/F_1)\cdot P(F_1)}{\sum_{i=1}^{3} P(E/F_i)\cdot P(F_i)} = \frac{(1-\alpha_1)\cdot(1/3)}{(1-\alpha_1)\cdot(1/3)+1/3+1/3} = \frac{1-\alpha_1}{3-\alpha_1}$$

FORMULA LUI BAYES – EXEMPLE (3)

- **Test-grilă**. Se dă o singură întrebare, cu m variante de răspuns. Fie p probabilitatea ca studentul S să ştie să răspundă corect la întrebare; probabilitatea de a ghici răspunsul este 1/m. Notând cu K evenimentul ca studentul să cunoască răspunsul şi cu C evenimentul ca S să răspundă corect, să se calculeze P(K/C).

$$P(K/C) = \frac{P(C/K)\cdot P(K)}{P(C/K)\cdot P(K) + P(C/\overline{K})\cdot P(\overline{K})} =$$

$$= \frac{1\cdot p}{p + (1/m)\cdot(1-p)} = \frac{m\cdot p}{1+(m-1)\cdot p}$$

- $m \rightarrow \infty \Rightarrow P(K/C) = 1.$ $(m=5,\ p=1/2) \Rightarrow P(K/C)=5/6.$