

Statistic Descriptivă - continuare -



Parametrii în statistica descriptivă

Măsuri de centralitate

- ♦ Media
- ♦ Mediana
- ♦ Modulul

Măsuri de împrăștiere

- ♦ Amplitudine
- ♦ Variația
- ♦ Deviația standard
- ♦ Coeficientul de variație
- ♦ Eroarea standard

Măsuri de simetriei

- ♦ Asimetria
- ♦ Excesul

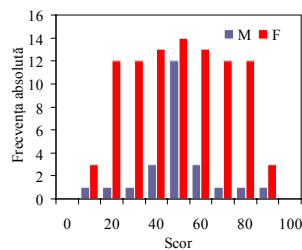
Măsuri de localizare

- ♦ Cvartiel (decile; percentile)



Distribuții de date

- Măsuri de centralitate: modulul, mediana, media aritmetică
- Împrăștierea
- Forma: simetria/asimetria, boltirea



Parametrii în statistica descriptivă

Măsuri de centralitate

- ♦ Media
- ♦ Mediana
- ♦ Modulul

Măsuri de împrăștiere

- ♦ Amplitudine
- ♦ Variația
- ♦ Deviația standard
- ♦ Coeficientul de variație
- ♦ Eroarea standard

Măsuri de simetriei

- ♦ Asimetria
- ♦ Excesul

Măsuri de localizare

- ♦ Cvartiel (decile; percentile)



Măsuri de împrăștiere

- Împrăștierea față de valoarea centrală
- Distribuția datelor unei variabile e cu atât mai mare cu cât valorile diferă mai mult unele față de celelalte

Parametrii:

1. Amplitudinea
2. Variația
3. Deviația standard
4. Coeficientul de variație
5. Eroarea standard



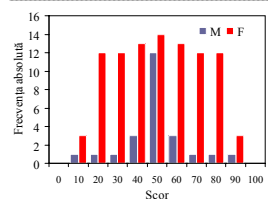
Măsuri de împrăștiere: Amplitudinea

$$A = X_{\max} - X_{\min}$$

- Nu se spune nimic despre modalitatea în care datele variază în jurul valori centrale
- Valorile extreme afectează semnificativ valoarea amplitudinii

Excel: RANGE (Descriptive Statistics)

- $A_M = 90 - 10 = 80$
- $A_F = 90 - 10 = 80$
 - Împrăștierea lor arată diferit



TRUNCHI COMUN, anul I (2008-2009) 7

Măsurile de împrăștiere: Media deviației

- De la medie

$$AD_{\bar{x}} = \frac{\sum_{i=1}^n |X_i - \bar{X}|}{n}$$

- De la mediană

$$AD_{Me} = \frac{\sum_{i=1}^n |X_i - Me|}{n}$$

StdID	Note	AD _{Medie}	AD _{Mediana}
34501	8	1,20	0,00
27896	3	-3,80	-5,00
32102	4	-2,80	-4,00
32654	8	1,20	0,00
32014	9	2,20	1,00
31023	9	2,20	1,00
30126	5	-1,80	-3,00
34021	9	2,20	1,00
33214	9	2,20	1,00
32016	4	-2,80	-4,00
Media	6,80		
Mediana	8,00		

Sorana D. BOLBOACA – INFORMATICA MEDICALA SI BIostatistica Curs 8

TRUNCHI COMUN, anul I (2008-2009) 8

Măsurile de împrăștiere: Media deviației

- Verificăm cât de diferite sunt notele a zece studenți față de medie prin folosirea distanțelor
- Cu cât nota e mai îndepărtată de medie cu atât deviația e mai mare
- Pentru a cuantifica cât de deviată e distribuția față de altă distribuție vom calcula sumele deviațiilor
- Diferența față de medie este foarte aproape de zero

StdID	Note	AD _{Medie}	AD _{Mediana}
34501	8	1,20	0,00
27896	3	-3,80	-5,00
32102	4	-2,80	-4,00
32654	8	1,20	0,00
32014	9	2,20	1,00
31023	9	2,20	1,00
30126	5	-1,80	-3,00
34021	9	2,20	1,00
33214	9	2,20	1,00
32016	4	-2,80	-4,00
Sum		0,00	-12,00

Sorana D. BOLBOACA – INFORMATICA MEDICALA SI BIostatistica Curs 8

TRUNCHI COMUN, anul I (2008-2009) 9

Măsurile de împrăștiere: SS

- Astfel utilizăm pătratul deviației față de medie
- Obținem astfel suma pătratelor abaterilor de la medie

$$SS = \sum_{i=1}^n (X_i - \bar{X})^2$$

StdID	Note	AD _{Medie}	AD _{Medie} ²
34501	8	1,20	1.39
27896	3	-3,80	14.59
32102	4	-2,80	7.95
32654	8	1,20	1.39
32014	9	2,20	4.75
31023	9	2,20	4.75
30126	5	-1,80	3.31
34021	9	2,20	4.75
33214	9	2,20	4.75
32016	4	-2,80	7.95
Sum		0,00	55,60

Sorana D. BOLBOACA – INFORMATICA MEDICALA SI BIostatistica Curs 8

TRUNCHI COMUN, anul I (2008-2009) 10

Măsurile de împrăștiere: Variația

- Media sumei pătratelor abaterilor de la medie se numește VARIAȚIA (se exprimă în pătratul unităților de măsură al valorilor observate)
- Variația populației:

$$\sigma^2 = \frac{SS}{n} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}$$

- Variația eșantionului (pentru a corecta faptul că variația eșantionului tinde să subestimeze variația populației):

$$s^2 = \frac{SS}{n-1} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$$

Sorana D. BOLBOACA – INFORMATICA MEDICALA SI BIostatistica Curs 8

TRUNCHI COMUN, anul I (2008-2009) 11

Măsurile de împrăștiere: Variația

Pentru a calcula variația:

- Calculează media.
- Află diferența dintre valoarea datei și medie pentru fiecare subiect.
- Calculează pătratul deviației față de medie.
- Calculează suma pătratelor diferențelor.
- Împarte suma pătratelor diferențelor la n dacă lucrezi cu toată populație sau la $(n-1)$ dacă lucrezi cu un eșantion al populației.

- $s^2 = 55,60/9 = 6,18$

StdID	Note	AD _{Medie}	AD _{Medie} ²
34501	8	1,20	1.39
27896	3	-3,80	14.59
32102	4	-2,80	7.95
32654	8	1,20	1.39
32014	9	2,20	4.75
31023	9	2,20	4.75
30126	5	-1,80	3.31
34021	9	2,20	4.75
33214	9	2,20	4.75
32016	4	-2,80	7.95
Sum		0,00	55,60

Sorana D. BOLBOACA – INFORMATICA MEDICALA SI BIostatistica Curs 8

TRUNCHI COMUN, anul I (2008-2009) 12

Măsurile de împrăștiere: Deviația Standard

Deviația standard = abaterea standard = ecartul tip

- Are aceeași unitate de măsură ca și media și datele seriei
- Variația se folosește în statistica inferențială
- Deviația standard se folosește în statistica descriptivă

$$s = \sqrt{s^2} = \sqrt{\frac{SS}{n-1}} = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}}$$

Sorana D. BOLBOACA – INFORMATICA MEDICALA SI BIostatistica Curs 8

Măsuri de împrăștiere: Deviația Standard

Interval	Procent observații conținute
$\bar{X} \pm 1 \cdot s$	68.3
$\bar{X} \pm 2 \cdot s$	95.5
$\bar{X} \pm 3 \cdot s$	99.7



Măsuri de împrăștiere: Coeficientul de variație

- Măsură relativă a dispersiei datelor
- Formula de calcul:

$$CV = \frac{s}{\bar{X}}$$

- Evaluare a abaterii standard în raport cu valoarea medie
- Are avantajul de a fi un indicator independent de unitățile de măsură



Măsuri de împrăștiere: Coeficientul de variație

- Interpretarea omogenității:

Coeficient de variație (CV)	Interpretare: populația poate fi considerată
$CV < 10\%$	<u>omogenă</u>
$10\% \leq CV < 20\%$	<u>relativ omogenă</u>
$20\% \leq CV < 30\%$	<u>relativ eterogenă/relativ heterogenă</u>
$> 30\%$	<u>eterogenă/heterogenă</u>



Măsuri de împrăștiere: Eroarea standard

$$ES = \frac{s}{\sqrt{n}}$$

- intervine în estimarea statistică în determinarea intervalelor de încredere



Parametrii în statistica descriptivă

Măsuri de centralitate	Măsuri de împrăștiere
	Măsuri de localizare
	♦ Cvartile (decile; percentile)



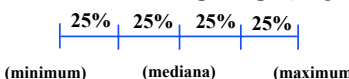
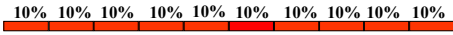
Măsuri de localizare

- Cvartile
- Percentile
- Decile
- Funcția Excel pentru cvartile:
- **QUARTILE**



TRUNCHI COMUN, anul I (2008-2009) 19

Măsurile de localizare: Cvartile – Decile

- Cvartile:**
 - Împarte seria de date în patru părți egale:
- Decile:**
 - Împarte seria de date în 10 părți egale:

Sorana D. BOLBOACA – INFORMATICĂ MEDICALĂ ȘI BIOSTATISTICĂ Curs 8

TRUNCHI COMUN, anul I (2008-2009) 20

Măsurile de localizare

- Percentile:**
 - Împarte seria de date în 100 părți egale
- Simetria unei distribuții analizată cu ajutorul cvartilelor:**
 - Fie Q_1, Q_2, Q_3 prima (1/3), a doua (1/2) și a treia (3/4) cvartilă:
 - $Q_2 - Q_1 \approx Q_3 - Q_2$ (\approx înseamnă aproximativ egal cu) \rightarrow distribuția este aproximativ simetrică
 - $Q_2 - Q_1$ e diferită de $Q_3 - Q_2 \rightarrow$ distribuția este asimetrică (spre stânga sau spre dreapta)

Sorana D. BOLBOACA – INFORMATICĂ MEDICALĂ ȘI BIOSTATISTICĂ Curs 8

TRUNCHI COMUN, anul I (2008-2009) 21

Măsurile de localizare: cvartile

- n = numărul de observații
- se ordonează datele crescător
- proporția observațiilor care au valori mai mici decât a k observație din serie $Q_k = k/(n+1)$
- $k = Q_k(n+1)$ ($Q_1 = 0,25; Q_2 = 0,50; Q_3 = 0,75$)
- Dacă k e număr întreg a k -a observație este valoarea estimată a cvartilei
- Dacă k nu este întreg, considerăm j valoarea întreagă a lui k :

valoarea cvartilei = $X_j + (X_{j+1} - X_j) \cdot (k - j)$

Sorana D. BOLBOACA – INFORMATICĂ MEDICALĂ ȘI BIOSTATISTICĂ Curs 8

TRUNCHI COMUN, anul I (2008-2009) 22

Măsurile de localizare: cvartile

- Greutatea la naștere a 10 copii exprimată în kg:
3,25; 4,10; 3,45; 3,40; 2,80; 2,97; 3,05; 4,40; 4,30; 3,80
- Ordonarea ascendentă a datelor:

2,80	2,97	3,05	3,25	3,40	3,45	3,80	4,10	4,30	4,40
X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}

Sorana D. BOLBOACA – INFORMATICĂ MEDICALĂ ȘI BIOSTATISTICĂ Curs 8

TRUNCHI COMUN, anul I (2008-2009) 23

Măsurile de localizare: cvartile

2,80	2,97	3,05	3,25	3,40	3,45	3,80	4,10	4,30	4,40
X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}

- $k_1 = Q_1(10+1) = 0,25 \cdot 11 = 2,75$ (valoarea e cuprinsă între a doua și a treia observație)

 $Q_1 = 2,97 + (3,05 - 2,97) \cdot (2,75 - 2) = 3,03$
- $k_2 = Q_2(10+1) = 0,50 \cdot 11 = 5,5$ (valoarea e cuprinsă între a cincea și a șasea observație)

 $Q_2 = 3,40 + (3,45 - 3,40) \cdot (5,50 - 5) = 3,43$
- $k_3 = Q_3(10+1) = 0,75 \cdot 11 = 8,25$ (valoarea e cuprinsă între a opta și a noua observație)

 $Q_3 = 4,10 + (4,30 - 4,10) \cdot (8,25 - 8) = 4,15$

Sorana D. BOLBOACA – INFORMATICĂ MEDICALĂ ȘI BIOSTATISTICĂ Curs 8

TRUNCHI COMUN, anul I (2008-2009) 24

Măsurile de localizare: cvartile

2,80	2,97	3,05	3,25	3,40	3,45	3,80	4,10	4,30	4,40
X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}

- $Q_1 = 3,03$
- $Q_2 = 3,43$
- $Q_3 = 4,15$

$$Q_2 - Q_1 = 3,43 - 3,03 = 0,40$$

$$Q_3 - Q_2 = 4,15 - 3,43 = 0,72$$

Interpretare???

Sorana D. BOLBOACA – INFORMATICĂ MEDICALĂ ȘI BIOSTATISTICĂ Curs 8

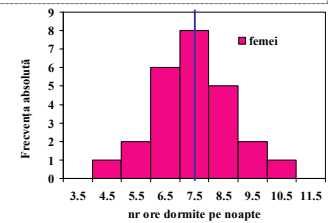
Parametrii în statistica descriptivă

Măsuri de centralitate	Măsuri de împrăștiere
Măsuri de simetrie ♦ Asimetria ♦ Excesul	Măsuri de localizare



Măsuri de simetrie

- Într-o distribuție simetrică
media aritmetică = mediana = valoarea modală



Măsuri de simetrie: Asimetria

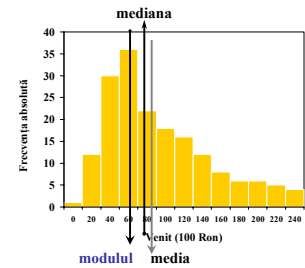
- este destinată să indice pentru o serie sau o distribuție de date:
 - extinderea asimetriei (abaterea de la aspectul simetric)
 - direcția asimetriei (pozitivă sau negativă)
- Formula de calcul:

$$M_3 = \frac{\sum_{i=1}^n (X_i - \bar{X})^3}{n}$$



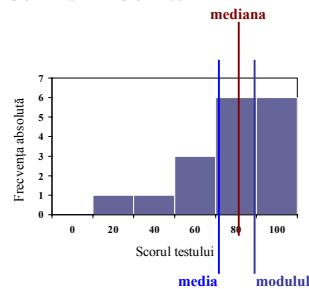
Măsuri de simetrie: Asimetria

- Asimetrie la stânga / pozitivă:
 - Modulul** = 7000 Ron
 - Mediana** = 8870 Ron
 - Media** = 9360 Ron
- Modulul < Mediana < Media aritmetică**



Măsuri de simetrie: Asimetria

- Asimetrie la dreapta / negativă:
 - Modulul > Mediana > Media aritmetică**
- Excel:
 - = SKEW(număr1, ..., numărn)



Măsuri de simetrie: Boltirea

- o măsură a formei unei serii sau distribuții de date, care măsoară înălțimea aplatizării/boltirii unei distribuții în comparație cu o distribuție normală

$$\alpha_4 = \frac{1}{n} \cdot \frac{\sum_{i=1}^n (X_i - \bar{X})^4}{S^4} - 3$$

- Excel:
 - = KURT(număr1, ..., numărn)



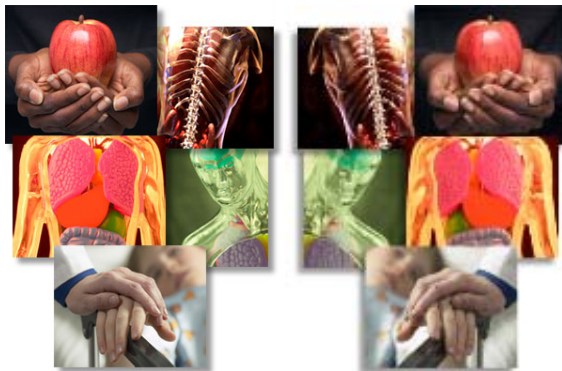
Importanța unităților de măsură

- Dacă la fiecare din datele seriei adunăm sau scădem o constantă atunci:
 - Media va crește respectiv scădea cu valoarea constantei adăugate
 - Deviația standard nu se va modifica
- Dacă înmulțim sau împărțim fiecare din datele seriei cu o constantă:
 - Media se va înmulți sau se va împărți cu valoarea constantei
 - Deviația standard se va înmulți sau împărți cu valoarea constantei



De ținut minte!

- Unitățile de măsură influențează statisticile descriptive.
- Statisticile descriptive trebuie aplicate diferențiat în funcție de scala de măsură a variabilei.
- Este utilă cunoașterea parametrilor descriptivi.



Probabilitatea



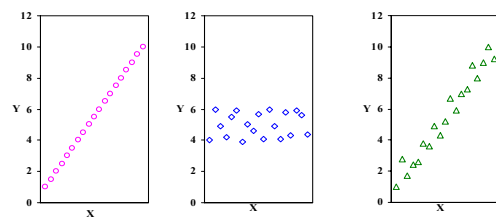
Probabilitatea

- Evenimentul aleatoriu
- Probabilități condiționate
- Proprietățile probabilităților
- Reguli
- Exemple: Procesul aleatoriu binomial



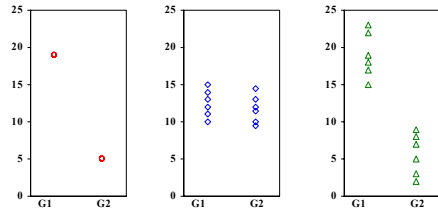
Probabilități

- Datele pot fi generate de:
 - Un proces pur sistematic
 - Un proces aleatoriu
 - Combinația dintre un proces sistematic și un proces aleator



Probabilități

- Datele pot fi generate de:
 - Un proces pur sistematic
 - Un proces aleatoriu
 - Combinația dintre un proces sistematic și un proces aleator



Probabilități & Ipoteze

- Dorim să știm care din cele 3 explicații este mai corectă
- Procesul pur sistematic este ușor de confirmat sau respins dacă ne uităm la distribuția grafică a datelor (rar se întâlnește în științele medicale)
- Ne-au mai rămas doar două întrebări:
 - procesul e aleator
 - combinația dintre un proces aleator și un proces pur sistematic



Evenimentul aleator

Definiții:

- **Probă:** aplicarea experimentului asupra unui element al colectivității
- **Eveniment:** rezultatul unei probe
- **Eveniment aleator:** evenimentul ce apare ca rezultat al unei singure probe (sau încercări)



Evenimentul aleatoriu binomial

- La aruncarea cu moneda avem două rezultate posibile (capul sau pajura) asociate cu o probabilitate specifică (ex. 0,05)
- Nu putem spune prea multe despre frecvența absolută a unuia din cele două evenimente posibile pentru un număr mic de aruncări dar putem spune multe despre frecvența relativă obținută în urma aruncării monedei de mai multe ori.



Calculul probabilităților

- Studiul probabilităților producerii unui fenomen sau a unui eveniment dintr-un complex de fenomene sau de evenimente de același fel



Probabilitatea

- o măsură a șansei de realizare a unui eveniment
- $\Pr(A) \in [0, 1]$
- Fie A un eveniment:
 - $\Pr(A)$ = probabilitatea evenimentului A
 - Dacă evenimentul este o certitudine: $\Pr(A) = 1$
 - Dacă evenimentul este imposibil de realizat: $\Pr(A) = 0$



Probabilitatea subiectivă vs obiectivă

Subiectivă:

- stabilită subiectiv sau empiric pe baza experienței sau prin studii populaționale foarte largi
- implică evenimente elementare care **nu sunt** echiprobabile (echiprobabil = are tot atâtea șanse de a se produce ca și alte evenimente)

Formula de calcul:

- Dacă un eveniment A se poate realiza în S probe dintr-o serie de n încercări echiprobabile, atunci probabilitatea evenimentului A este dată de numărul de cazuri favorabile raportat la numărul de cazuri posibile
- $Pr(A) = (\text{nr cazuri favorabile}) / (\text{nr cazuri posibile})$



Șanse și rații

- Șansele sunt probabilități exprimate procentual
- Iau valori între 0% și 100%
 - Exemplu: o probabilitate de 0,75 este egală cu o șansă de 75%
- Rația unui eveniment este probabilitatea ca un eveniment să se întâmple împărțit la probabilitatea ca acel eveniment să nu se întâmple
 - Poate lua orice valoare pozitivă
 - Fie A evenimentul de interes. Rația de probabilitate = $Pr(A) / [1 - Pr(A)]$ (unde $1 - Pr(A) = Pr(\text{non}A)$)
 - Exemplu: dacă $Pr(A) = 0,75$ atunci rația de probabilitate este de 3 la 1 ($0,75 / (1 - 0,75) = 0,75 / 0,25 = 3/1$)



Spațiul unui eveniment

- este determinat de mulțimea tuturor rezultatelor posibile ale unui proces aleatoriu
 - La aruncarea cu zarul spațiul de evenimente este format din {1, 2, 3, 4, 5, 6}
 - La aruncarea unei monede spațiul de evenimente este {cap, pajură}.
- un eveniment este un membru al spațiului evenimentului
 - “cap” este un eveniment posibil la aruncarea unei monede
 - “un număr mai mic sau egal cu 3” este un eveniment posibil la aruncarea unui zar
- **Evenimentele au asociate probabilități de producere!**



Probabilități: proprietăți

- Iau valori între 0 și 1:

$$0 \leq Pr(A) \leq 1$$
- $Pr(\text{spațiu al unui eveniment}) = 1$
- Probabilitatea de a se întâmpla un eveniment este 1 minus probabilitatea de a nu se întâmpla acel eveniment

$$Pr(A) = 1 - P(\text{not}A)$$



Concepte de bază

- Evenimente compatibile: evenimente care se pot realiza simultan:
 - $A = \{TAS < 140 \text{ mmHg}\}$
 - $B = \{TAD < 90 \text{ mmHg}\}$
- Evenimente incompatibile: evenimente care nu se pot realiza simultan:
 - $A = \{TAS < 140 \text{ mmHg}\}$
 - $B = \{140 \leq TAS < 200 \text{ mmHg}\}$



Concepte de bază

- Evenimentul A implică evenimentul B dacă evenimentul B se produce ori de câte ori se produce evenimentul A:
 - Simbol $A \subset B$
 - $A = \{TBC\}$
 - $B = \{\text{testul la tuberculină pozitiv}\}$



Probabilități condiționate

- Probabilități condiționate:
 - Fie A și B două evenimente
 - Prin probabilitatea condiționată a lui A de către B (simbol: $\Pr(A|B)$) se înțelege probabilitatea de a se realiza evenimentul A dacă în prealabil s-a realizat evenimentul B
 - Exemplu: $\Pr(\text{Test pozitiv tuberculină}|TBC)$ este probabilitatea de a obține un test pozitiv la tuberculină la un pacient care are TBC.
 - $P(B|A)$ nu este același lucru cu $P(A|B)$**



Probabilități condiționate

	TBC+	TBC-
Test+	15	12
Test-	25	18

- Fie:
 - $A = \{TBC+\}$
 - $B = \{\text{Test}+\}$
- $\Pr(A) = (15+25)/(15+12+25+18) = 0,57$ (prevalența bolii)
- $\Pr(\text{non}A) = (12+18)/(15+12+25+18) = 0,73$
- $\Pr(B|A) = \text{probabilitatea unui test pozitiv la un pacient cu TBC} = 15/(15+25) = 0,38 = \text{SENSIBILITATE (Se)}$



Probabilități condiționate

	TBC+	TBC-
Test+	15	12
Test-	25	18

- Fie:
 - $A = \{TBC+\}$
 - $B = \{\text{Test}+\}$
- $\Pr(\text{non}B|\text{non}A) = \text{probabilitatea de a obține un test negativ știind că testul se aplică unui pacient indemn de TBC} = 18/(18+12) = 0,60 = \text{SPECIFICITATE (Sp)}$
- $\Pr(A|B) = \text{probabilitatea ca o persoană cu TBC să prezinte un test pozitiv} = 15/(15+12) = 0,56 = \text{VALOAREA PREDICTIVĂ POZITIVĂ (VPP)}$



Probabilități condiționate

	TBC+	TBC-
Test+	15	12
Test-	25	18

- Fie:
 - $A = \{TBC+\}$
 - $B = \{\text{Test}+\}$
- $\Pr(\text{non}A|\text{non}B) = \text{probabilitatea ca o persoană indemnă TBC să prezinte un test negativ} = 18/(18+25) = 0,42 = \text{VALOAREA PREDICTIVĂ NEGATIVĂ (VPN)}$



Probabilități condiționate

	TBC+	TBC-
Test+	15	12
Test-	25	18

- Fie:
 - $A = \{TBC+\}$
 - $B = \{\text{Test}+\}$
- Rata falșilor pozitivi: $\text{RFP} = \Pr(B|\text{non}A)$
- Rata falșilor negativi: $\text{RFN} = \Pr(\text{non}A|B)$



Evenimente independente: probabilități condiționate

- Două evenimente A și B se numesc *independente* dacă și numai dacă

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B).$$
- În acest caz:
 - $\Pr(B|A) = \Pr(B|\text{non}A) = \Pr(B)$
 - $\Pr(A|B) = \Pr(A|\text{non}B) = \Pr(A)$
- ceea ce exprimă bine independența celor două evenimente prin faptul că probabilitatea evenimentului B (respectiv A) nu depinde de realizarea evenimentului A (respectiv B).
 - Exemplu: dacă aruncăm moneda de două ori probabilitatea ca la a doua aruncare să obținem cap este întotdeauna 0,5 indiferent ce am obținut la prima aruncare



Operații cu evenimente

- **REUNIUNEA (SAU):**
 - $A \cup B$ - se produce cel puțin unul dintre evenimentele A sau B
- **INTERSECȚIA (ȘI):**
 - $A \cap B$ - evenimentele A și B se produc simultan
- **NEGAREA:**
 - nonA



Reguli de probabilitate

- Probabilitatea de apariție a evenimentului A sau B: **ADUNARE**

$$\Pr(A \text{ sau } B) = \Pr(A) + \Pr(B)$$

- evenimentele A și B sunt mutual exclusive

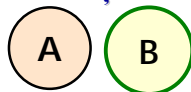
- Probabilitate de A și B: **ÎNMULȚIRE**

$$P(A \text{ și } B) = P(A) \cdot P(B)$$

- evenimentele A și B sunt independente



Reguli de adunare a probabilităților



- Fie A și B două evenimente:
 $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$
- Evenimente mutual exclusive: $\Pr(A \cap B) = 0$



Reguli de adunare a probabilităților

- $A = \{\text{TAS mamă} > 140 \text{ mmHg}\}$, $\Pr(A) = 0,25$
- $B = \{\text{TAS tată} > 140 \text{ mmHg}\}$, $\Pr(B) = 0,15$
- Care este probabilitatea ca într-o familie să avem un părinte hipertensiv?

$$\Pr(A \cup B) = 0,25 + 0,15 - 0 = 0,40$$



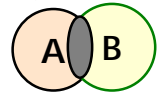
Reguli de adunare a probabilităților

- Într-o cafenea există 20 de persoane; la 10 le place ceaiul, la alți 10 cafeaua și la 2 le place și ceaiul și cafeaua.
- Care este probabilitatea de a extrage la întâmplare din populație o persoană careia să-i placă și ceaiul și cafeaua?

$$\Pr(\text{ceai} \cup \text{cafea}) = \Pr(\text{ceai}) + \Pr(\text{cafea}) - \Pr(\text{ceai} \cap \text{cafea}) = 0,50 + 0,50 - 0,10 = 0,90$$



Reguli de înmulțire a probabilităților



- Fie A și B două evenimente:
 $\Pr(A \cap B) = \Pr(A) \cdot \Pr(B|A)$
- Evenimente independente $\Pr(B|A) = \Pr(B)$



Reguli de înmulțire a probabilităților

- $A = \{\text{TAS mamă} > 140 \text{ mmHg}\}$, $\Pr(A) = 0,10$
- $B = \{\text{TAS tată} > 140 \text{ mmHg}\}$, $\Pr(B) = 0,20$
- $\Pr(A \cap B) = 0,05$
- Evenimentele A și B sunt dependente sau independente?

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B) - \text{evenimente independente}$$

$$0,05 \neq 0,10 \cdot 0,20 \rightarrow \text{evenimente dependente}$$



De reținut! Operații cu probabilități

- **Adunare:**
 - $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$
 - $\Pr(A \cup B) = \Pr(A) + \Pr(B)$: **evenimente mutual exclusive**
- **Înmulțire:**
 - $\Pr(A \cap B) = \Pr(A) \cdot \Pr(B|A)$
 - $\Pr(A \cap B) = \Pr(A) \cdot \Pr(B)$: **evenimente independente**

