UMF
UNIVERSITATEA DE MEDICINĂ ȘI FARMACIE
IULIU HAȚIEGANU
CLUJ-NAPOCA

# Statistici Descriptive (2)

# Statistici descriptive

| Măsuri de centralitate | Măsuri de dispersie / împrăştiere |
|---|---|
| ◆ Media<br>◆ Mediana<br>◆ Modulul<br>◆ Valoarea centrală | ◆ Amplitudine<br>◆ Intervalul intercvartilic [Q1, Q3]<br>◆ **Variaţia (sau varianţa)**<br>◆ **Deviaţia standard**<br>◆ **Coeficientul de variaţie**<br>◆ **Eroarea standard** |
| **Măsuri de simetriei** | **Măsuri de localizare** |
| ◆ Asimetria<br>◆ Boltirea | ◆ Cvartile (decile; percentile) |

# Statistici descriptive de dispersie

- Calcularea parametrilor de dispersie/împrăştiere:
  - Amplitudinea
  - Intervalul intercvartilic [Q1, Q3]: IQR
  - Distanța intercvartile: IQR=Q3-Q1
  - Deviaţia standard de la medie
  - Eroarea standard
  - Coeficientul de variație

# Amplitudinea

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

- Nu ne spune nimic despre modalitatea în care datele variază în jurul valori centrale
- Valorile extreme afectează semnificativ valoarea amplitudinii
- Excel: RANGE (Descriptive Statistics)

| | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $X_5$ | $X_6$ | $X_7$ | $X_8$ | $X_9$ | $X_{10}$ |
| 2 | 3 | 9 | 9 | 5 | 8 | 6 | 10 | 7 | 8 | 5 |

$$A = X_{max} - X_{min} = 10 - 3 = 7$$

# Abaterea de la medie (de la mediană)

- Măsoară media distanţelor dintre fiecare valoare a variabilei şi valoarea centrală (media aritmetică sau mediana).
- Dă o pondere egală fiecărei observaţii
- Sensibilitate mai mare decât amplitudinea sau intervalul intercuartilic dintre cuartila 3 şi 1 (IQR)
- Se utilizează în calculul variaţiei şi respectiv a deviaţiei standard

$$AD_m = \frac{|X_1 - m| + |X_2 - m| + |X_3 - m| + \cdots + |X_n - m|}{n} \qquad AD_{Me} = \frac{|X_1 - Me| + |X_2 - Me| + |X_3 - Me| + \cdots + |X_n - Me|}{n}$$

- Unde $\bar{X}$ este media aritmetică (Ma), Me=mediana datelor

# Variația standard

- **VARIAȚIA (descriptivă) a EȘANTIONULUI** = media sumei pătratelor abaterilor de la medie

$$s^2 = \frac{(x_1-m)^2+(x_2-m)^2+\cdots(x_n-m)^2}{n}$$ unde n = talia eșantionului

## Avantaje

✓ pozitivă

## Dezavantaje

✗ unitățile varianței = pătratul unităților variabilei

✗ varianță mare => dispersie mare în jurul mediei => greu de inteles și de facut comparații

✗ este sensibilă la valori extreme

✗ are sens doar pentru o variabilă CANTITATIVĂ

# Deviația standard

**DEVIAȚIA STANDARD** (abaterea standard) **a eșantionului** se notează cu **s**

- Are aceeași unitate de măsură ca și media și datele seriei

$$s = \sqrt{\frac{\sum_{i=1}^{n}(X_i - m)^2}{n}} = \sqrt{\frac{(x_1 - m)^2 + (x_2 - m)^2 + \cdots + (x_n - m)^2}{n}}$$

unde n = talia eșantionului

$$s = \sqrt{s^2}$$

- **DEVIAȚIA STANDARD (abaterea standard) a populației** se notează cu **σ**

unde N = talia populației; μ = media pe populație

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}} = \sqrt{\frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + \cdots + (x_N - \mu)^2}{N}}$$

# Deviația standard de eșantionare

- **<u>Deviația standard de eșantionare</u>** (deviația standard a eșantionului optimizată pentru a aproxima cel mai bine deviația standard a populației:

- $S = \sqrt{\frac{n}{n-1}}\, s$ unde n = talia eșantionului

$$S = \sqrt{\frac{(x_1-m)^2+(x_2-m)^2+\cdots+(x_n-m)^2}{n-1}}$$

- Va fi utilizată în statistica inferențială
- Notație S (majuscule) pentru deviația standard de eșantionare
- respectiv $S^2$ = variația de eșantionare

# Proprietățile deviației standard

**AVANTAJE**

- ✓ Aceleași unități de măsură ca și variabila în sine.
- ✓ Pozitivă
- ✓ Dacă înmulțim valorile unei serii statistice cu o constantă, abaterea standard se înmulțește cu aceeași constantă
- ✓ Dacă adăugăm o valoare la fiecare valoare a unei serii statistice, abaterea standard nu se modifică

**DEZAVANTAJE**

- ✓ este sensibilă la valori extreme

# Variația și deviația standard: exemplu de calcul

| Id_student | Nota | Deviația de la medie ($X_i$-m) | ($X_i$-m)$^2$ |
|---|---|---|---|
| $X_1$ | 9 | =9-7,3 = 1,67 | =1,67$^2$ = 2,78 |
| $X_2$ | 6 | -1,33 | 1,78 |
| $X_3$ | 4 | -3,33 | 11,11 |
| $X_4$ | 9 | 1,67 | 2,78 |
| $X_5$ | 4 | -3,33 | 11,11 |
| $X_6$ | 8 | 0,67 | 0,44 |
| $X_7$ | 8 | 0,67 | 0,44 |
| $X_8$ | 9 | 1,67 | 2,78 |
| $X_9$ | 7 | -0,33 | 0,11 |
| $X_{10}$ | 4 | -3,33 | 11,11 |
| $X_{11}$ | 10 | 2,67 | 7,11 |
| $X_{12}$ | 10 | 2,67 | 7,11 |

**Variația (descriptivă) a eșantionului**

$$s^2 = \frac{(9-7{,}33)^2 + (6-7{,}33)^2 + \cdots (10-7{,}33)^2}{12}$$

**Media aritmetică ($m$): = 7,33**

$s^2 = 58{,}77/12 = 4{,}90$

$s = \sqrt{4.90}$

**Deviația standard a eșantionului**

s = 2,21

# Variația și deviația standard: exemplu de calcul

| Id_student | Nota | Deviația de la medie ($X_i$-m) | ($X_i$-m)$^2$ |
|---|---|---|---|
| $X_1$ | 9 | =9-7,3 = 1,67 | =1,67$^2$ = 2,78 |
| $X_2$ | 6 | -1,33 | 1,78 |
| $X_3$ | 4 | -3,33 | 11,11 |
| $X_4$ | 9 | 1,67 | 2,78 |
| $X_5$ | 4 | -3,33 | 11,11 |
| $X_6$ | 8 | 0,67 | 0,44 |
| $X_7$ | 8 | 0,67 | 0,44 |
| $X_8$ | 9 | 1,67 | 2,78 |
| $X_9$ | 7 | -0,33 | 0,11 |
| $X_{10}$ | 4 | -3,33 | 11,11 |
| $X_{11}$ | 10 | 2,67 | 7,11 |
| $X_{12}$ | 10 | 2,67 | 7,11 |

**Variația de eșantionare**

$$S^2 = \frac{(9 - 7{,}33)^2 + (6 - 7{,}33)^2 + \cdots (10 - 7{,}33)^2}{11}$$

**Media aritmetică ($m$): = 7,33**

$S^2 = 58{,}77/11 = 5{,}34$

$S = \sqrt{5{,}34}$

**Deviația standard de eșantionare**

$S = 2{,}31$

# Variația și deviația standard: exemplu de calcul

| Id_subiect | Greutate(kg) | Deviația de la medie ($X_i$-m) | ($X_i$-m)$^2$ |
|---|---|---|---|
| 1 | 71 | -0.92 | 0.85 |
| 2 | 76 | 3.57 | 12.72 |
| 3 | 62 | -9.87 | 97.45 |
| 4 | 72 | -0.26 | 0.07 |
| 5 | 75 | 3.49 | 12.19 |
| 6 | 71 | -1.36 | 1.85 |
| 7 | 70 | -1.58 | 2.51 |
| 8 | 70 | -1.80 | 3.24 |
| 9 | 70 | -2.41 | 5.83 |
| 10 | 66 | -5.84 | 34.06 |
| 11 | 60 | -12.44 | 154.85 |
| 12 | 81 | 8.99 | 80.91 |
| 13 | 67 | -4.80 | 23.02 |
| 14 | 65 | -6.96 | 48.49 |
| 15 | 83 | 10.56 | 111.53 |
| 16 | 68 | -4.35 | 18.89 |
| 17 | 77 | 4.70 | 22.12 |
| 18 | 74 | 1.74 | 3.01 |
| 19 | 65 | -7.43 | 55.22 |
| 20 | 75 | 2.65 | 7.00 |
| 21 | 81 | 9.25 | 85.55 |
| 22 | 69 | -2.74 | 7.50 |
| 23 | 73 | 1.21 | 1.47 |
| ... | ... | ... | .. |

**Variația de eșantionare**

$$S^2 = \frac{(71-72)^2 + (76-72)^2 + \cdots (80-72)^2}{99}$$

**Media aritmetică ($m$): = 72 kg**

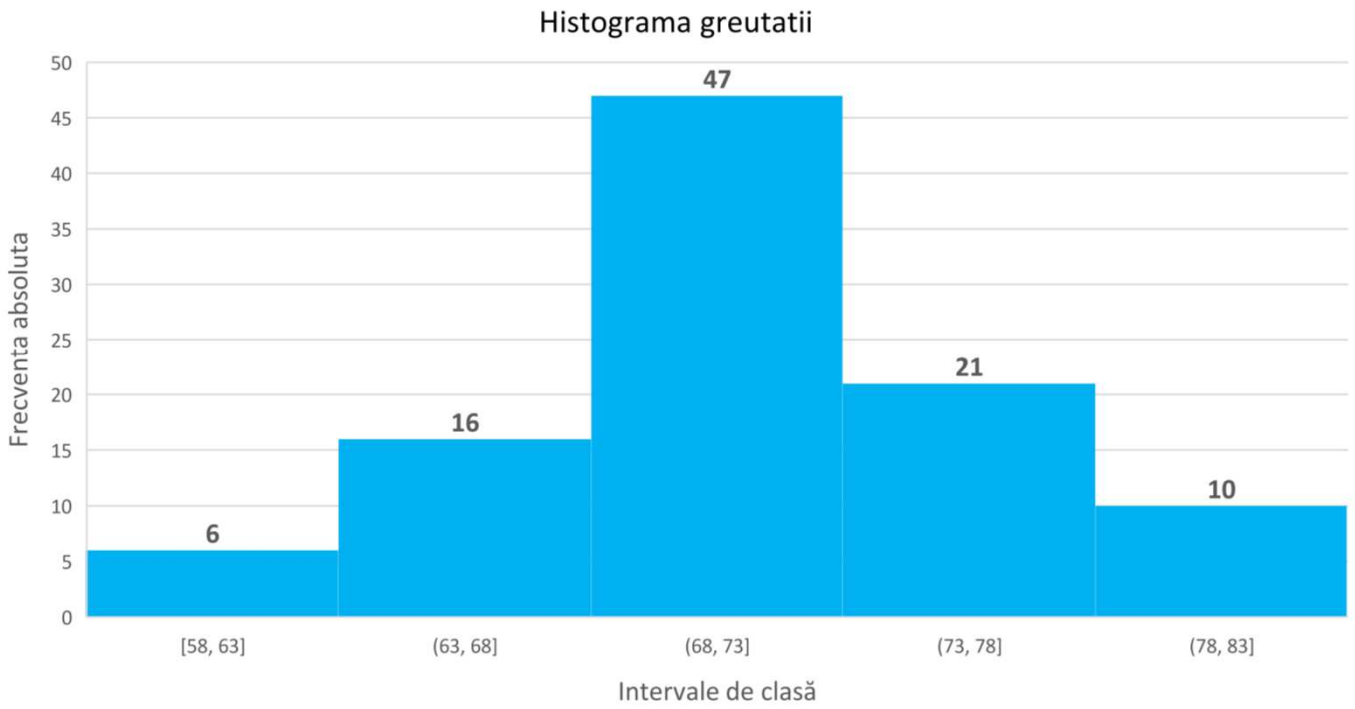
$S^2 = 2414{,}10/99 = 24{,}38$

$S = \sqrt{24{,}38}$

**Deviația standard de eșantionare**

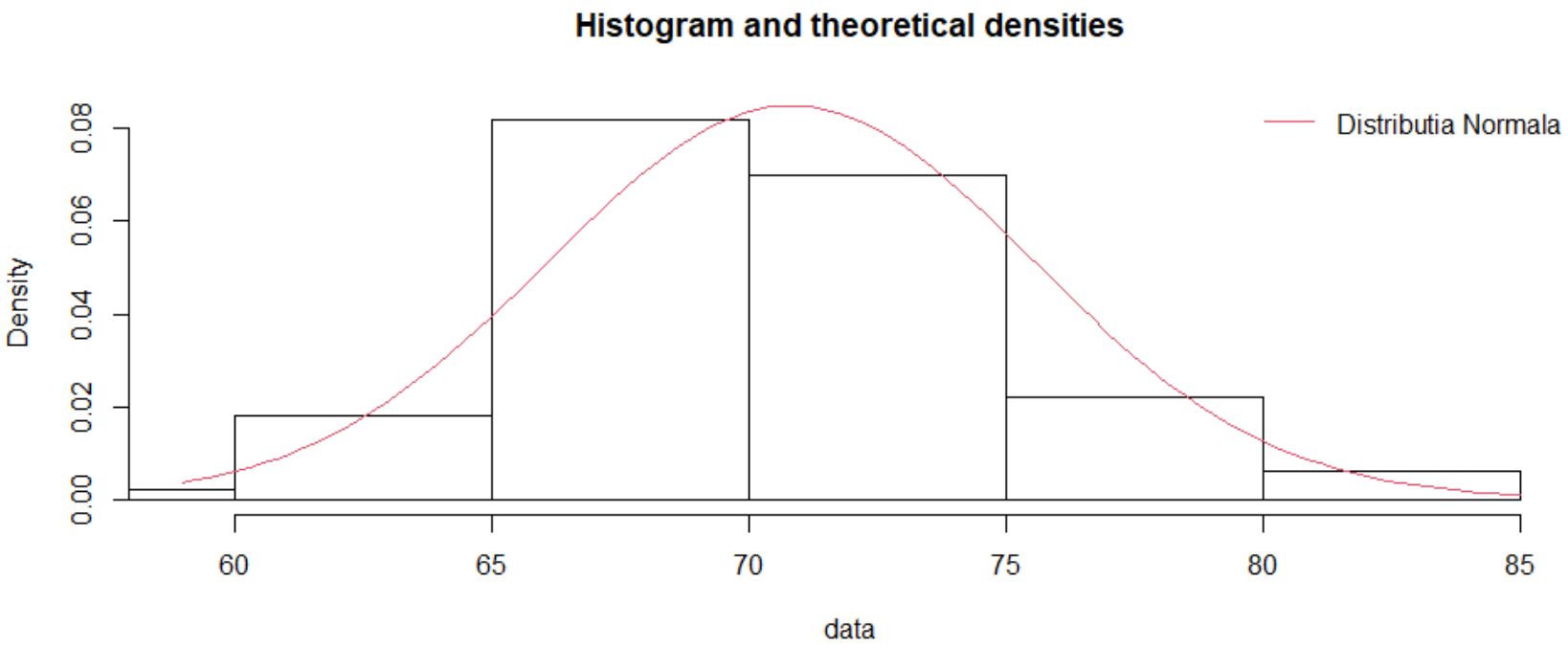
S = 4,94 kg

# Histograma și curba normală (gaussiană)



# Histograma și curba normală (gaussiană)

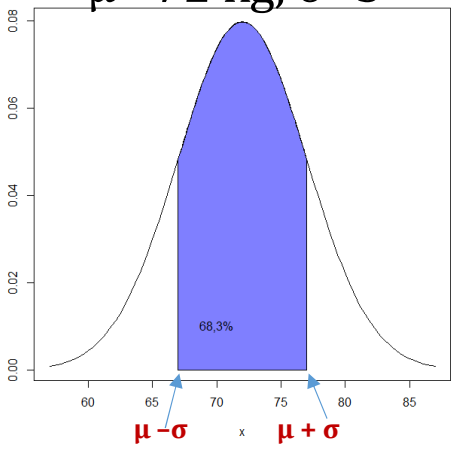
**Histograma greutății**



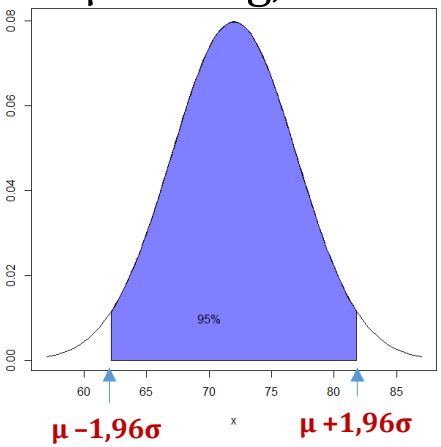
# Curba Normală

- Distribuţie simetrică în jurul mediei

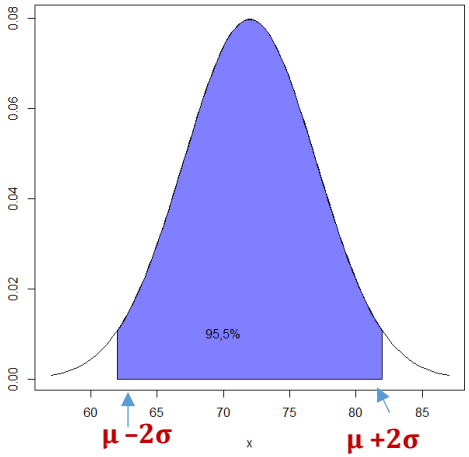
μ =72 kg, σ=5



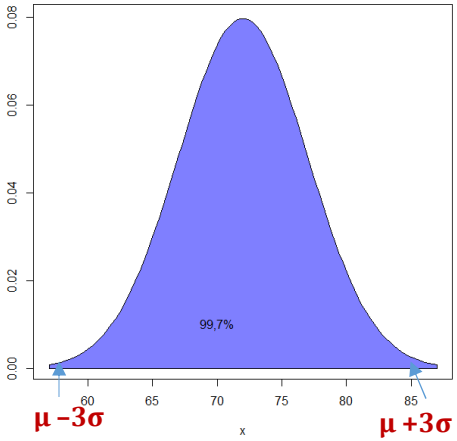

μ =72 kg, σ=5




μ =72 kg, σ=5



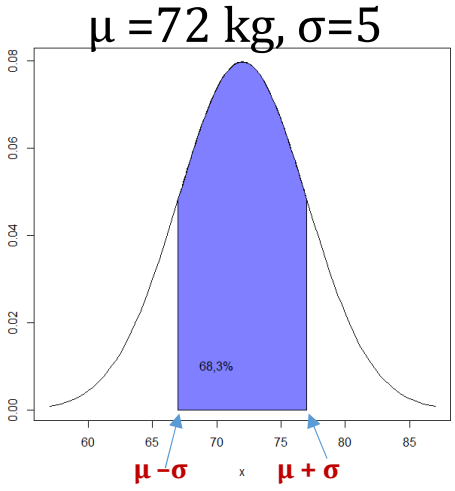

μ =72 kg, σ=5

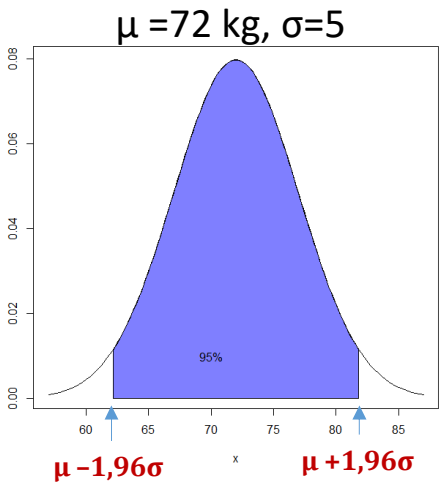


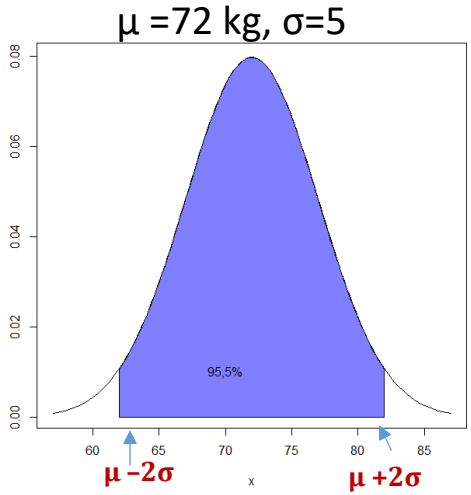

- Intervalul de valori **[72 –5; 72 +5] = [67, 77]** conţine aproximativ **~68,3 %** din valorile distribuţiei.
- Intervalul de valori **[72 – 2•5; 72 + 2•5] = [62, 82]** conţine aproximativ **~95,5 %** din valorile distribuţiei.
- Intervalul de valori **[72 –1,96•5; 72 + 1,96•5] = [62.2, 81.8]** conţine **95 %** din valorile distribuţiei.
- Intervalul de valori **[72 –3•5; 72 +3•5] = [57, 87]** conţine aproximativ **~99,7 %** din valorile distribuţiei.

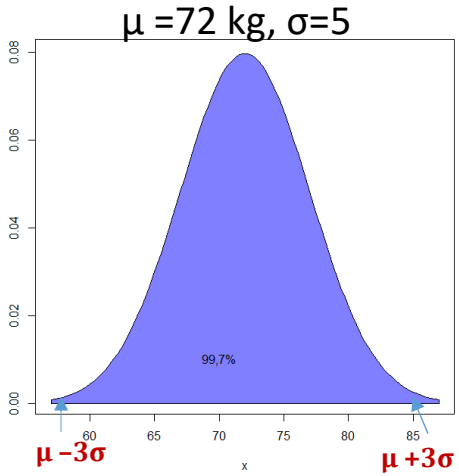
# Curba Normală

- Distribuţie simetrică în jurul mediei




- Intervalul de valori **[μ –σ; μ +σ]** conţine aproximativ **~68,3 %** din valorile distribuţiei.
- Intervalul de valori **[μ –2σ; μ +2σ]** conţine aproximativ **~95,5 %** din valorile distribuţiei.
- Intervalul de valori **[μ –1,96σ; μ +1,96σ]** conţine **95 %** din valorile distribuţiei.
- Intervalul de valori **[μ –3σ; μ +3σ]** conţine aproximativ **~99,7 %** din valorile distribuţiei.

# Coeficientul de variație

- Măsură relativă a dispersiei datelor
- Formula de calcul:

$$CV(\%) = \frac{s}{m} \times 100$$

Reguli empirice de interpretare a coeficientului de variație [100]:

- CV<10% ➔ populația poate fi considerată omogenă
- 10%≤ CV < 20% ➔ populația poate fi considerată relativ omogenă
- 20%≤ CV < 30% ➔ populația poate fi considerată relativ eterogenă
- CV ≥ 30% ➔ populația poate fi considerată eterogenă

- Evaluare a abaterii standard în raport cu valoarea medie
- Are avantajul de a fi un indicator independent de unitățile de măsură
- Se poate exprima și/sau procentual

# Coeficientul de variație

- Măsură a variabilității relative utilizată pentru:
  - Măsurarea modificărilor care au apărut în populație în timp
  - Compararea variabilității a două populații când unitățile de măsură sunt diferite (mg/dL vs mmol/L – colesterol)
  - Frecvent exprimat procental

| | Greutate (kg) | Înălțime (cm) |
|---|---|---|
| Media aritmetică | 72,6 | 168 |
| Deviația standard | 13,6 | 10,2 |

» Care din variabilele de mai sus are împrăștierea mai mare?

> Nu se poate răspunde la întrebare

» Care din variabilele de mai sus are împrăștierea **relativă la medie** mai mare ?

Greutate: CV = 13,6/72,6*100 = 19%

Înălțime: CV = 10,2/168*100 = 6,1%

# Dispersia vs. dispersia relativă

- Aplicabilitate: dispersia în două seturi de date
  - *A*: 12, 13, 16, 18, 18, 20
  - *B*: 120, 130, 160, 180, 180, 200

| | Grup A | Grup B |
|---|---|---|
| Media | 16 | 162 |
| Deviaţia standard | 3 | 29 |
| CV (%) | 18 | 18 |

- Aplicabilitate: dispersia în două seturi de date
  - *A*: 12, 13, 16, 18, 18, 20
  - *B*: 2, 3, 160, 18, 200, 300

| | Grup A | Grup B |
|---|---|---|
| Media | 16 | 114 |
| Deviaţia standard | 3 | 114 |
| CV (%) | 18 | 100 |

# Eroarea standard

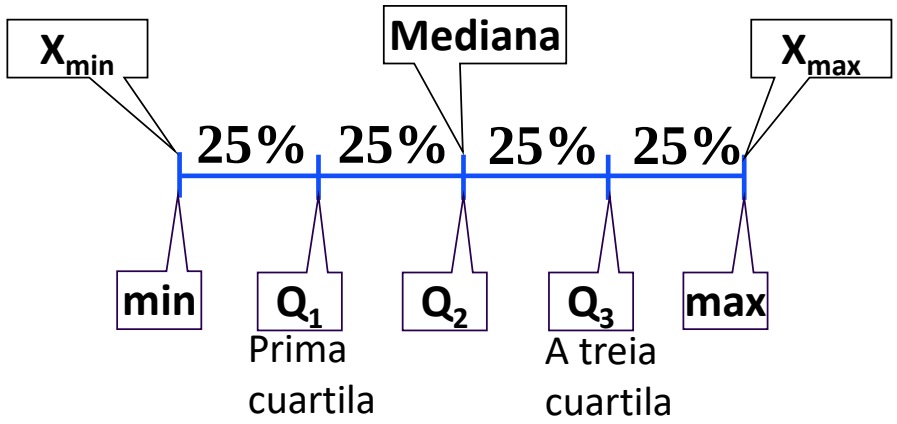
- este utilizată pentru a măsura precizia statistică a unei estimări
- este utilizată în statistica inferențială (teste de ipoteze și estimarea intervalului de încredere)

$$\boldsymbol{ES = \frac{S}{\sqrt{n}}}$$

unde S= deviația standard de eșantionare-vezi cursul viitor-intervale de încredere

n=talie eșantion

# Măsuri de localizare:
# Cuartilele și Intervalul intercuartilic (IQR)



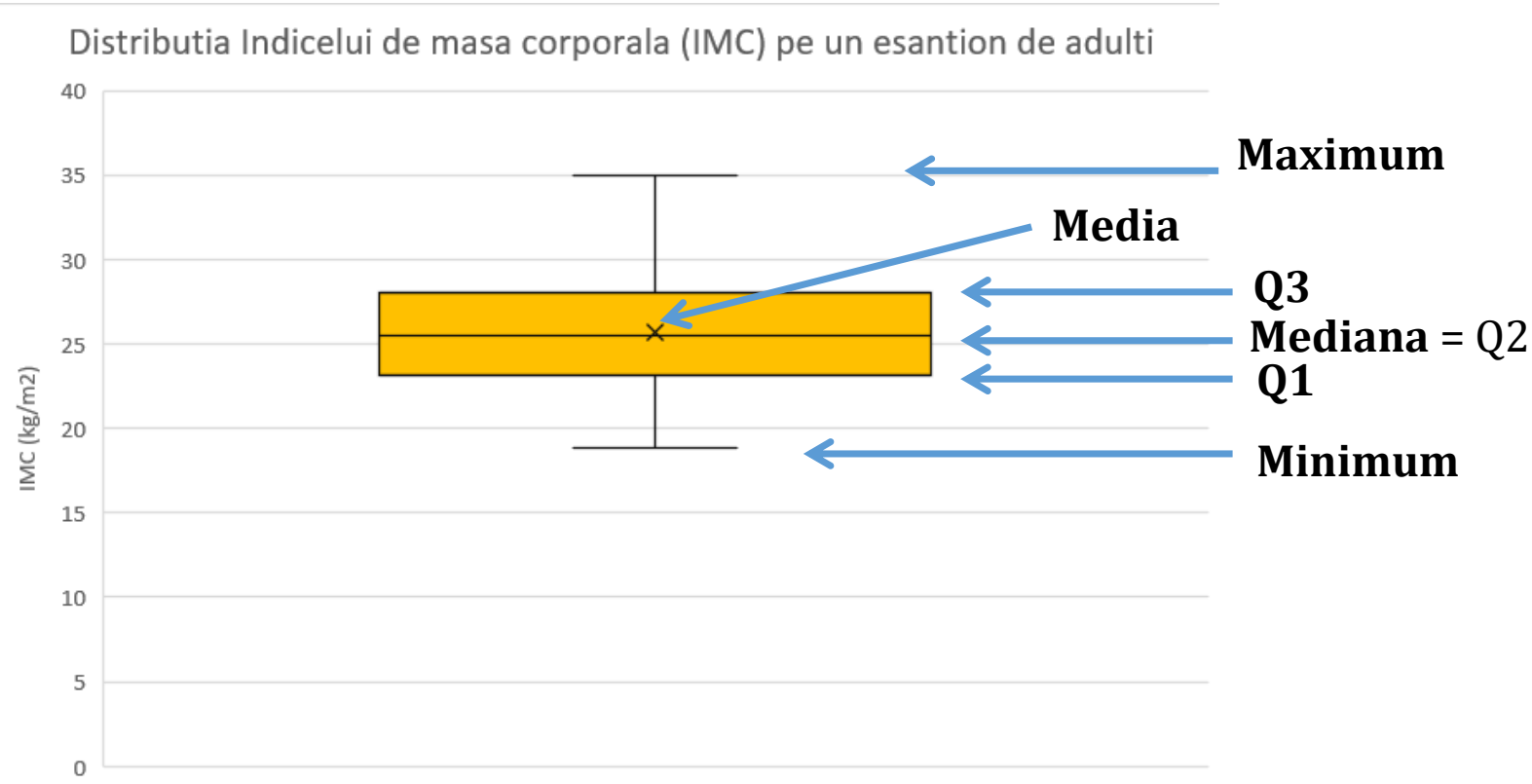

Măsură a dispersiei pentru 50% din datele de mijloc.

**$IQR=Q_3 - Q_1$,** unde Q3 = cvartila 3 (percentila 75%), Q1 = cvartila 1 (percentila de 25%)

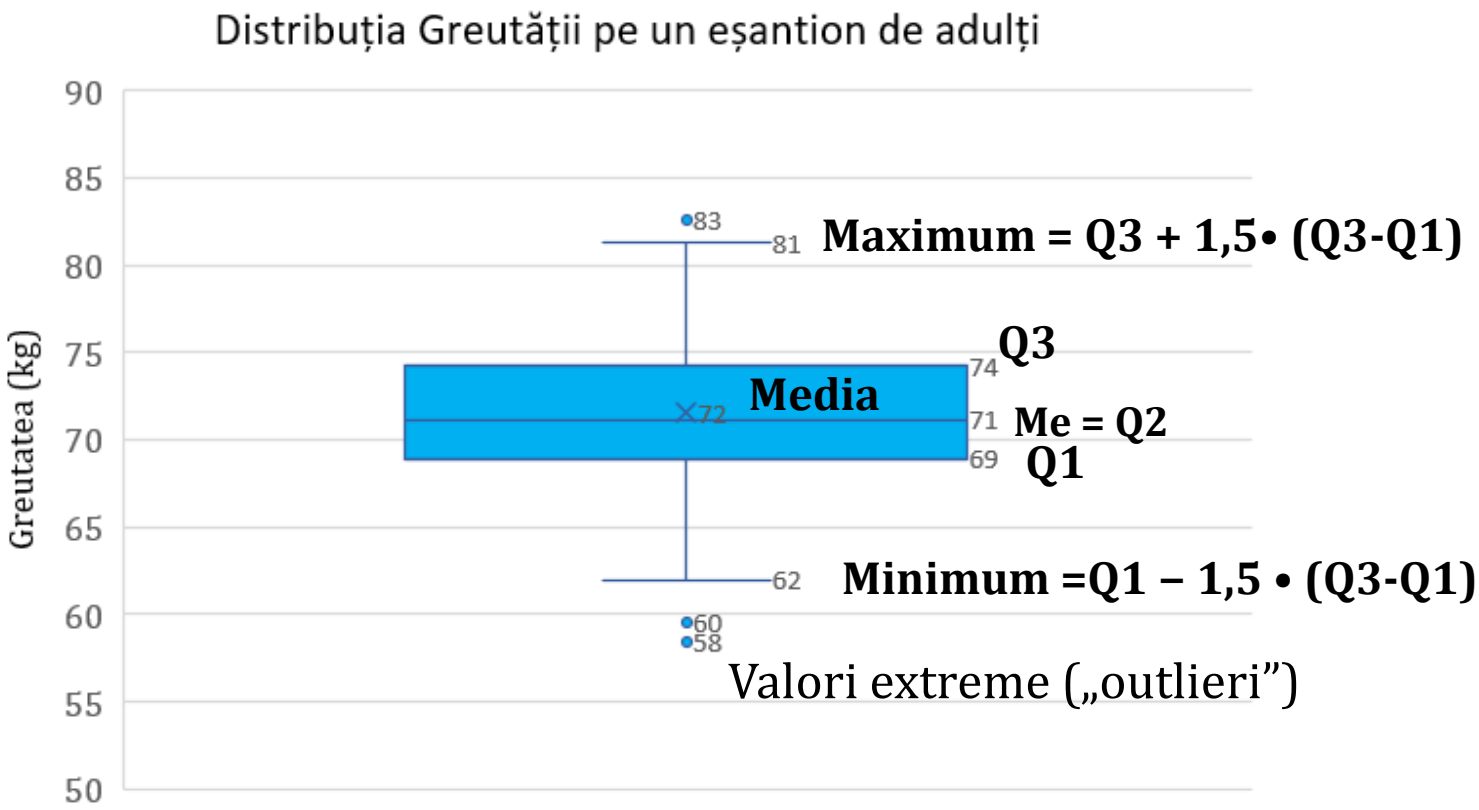
Sau

**IQR: $[Q_1; Q_3]$**

# Reprezentare grafică: box-plot

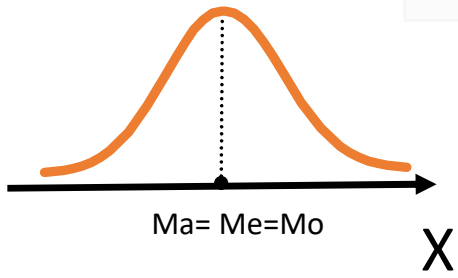


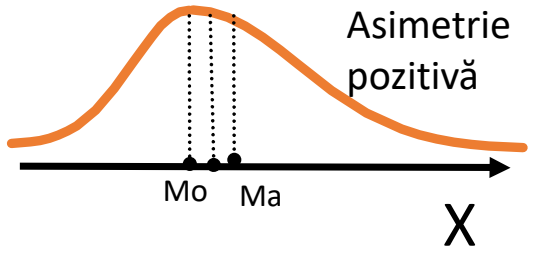
# Reprezentare grafică: box-plot

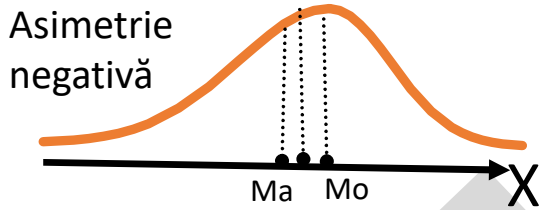


# Măsuri ale simetriei: Coeficientul de asimetrie

- Coeficientul de asimetrie (notație: $\alpha_3$): Măsoară gradul de asimetrie al unei distribuții de date
- **Semnul acestuia indică direcția asimetriei: pozitivă sau negativă**
  - ✓ $\alpha_3 \approx 0$ => distribuție simetrică.
  - ✓ $\alpha_3 > 0$ => distribuție alungită la dreapta– asimetrie pozitivă
  - ✓ $\alpha_3 < 0$ => distribuție alungită la stanga– asimetrie negativă
- **Valoarea acestuia indică gradul asimetrie:**
  - ✓ $\alpha_3 \in [-0,5; 0,5]$ distribuție aproximativ simetrică
  - ✓ $\alpha_3 \in [-1 ; -0,5]$ sau $[0,5; 1]$ distribuție cu asimetrie moderată
  - ✓ $\alpha_3 < -1$ sau $>1$ distribuție cu asimetrie importantă

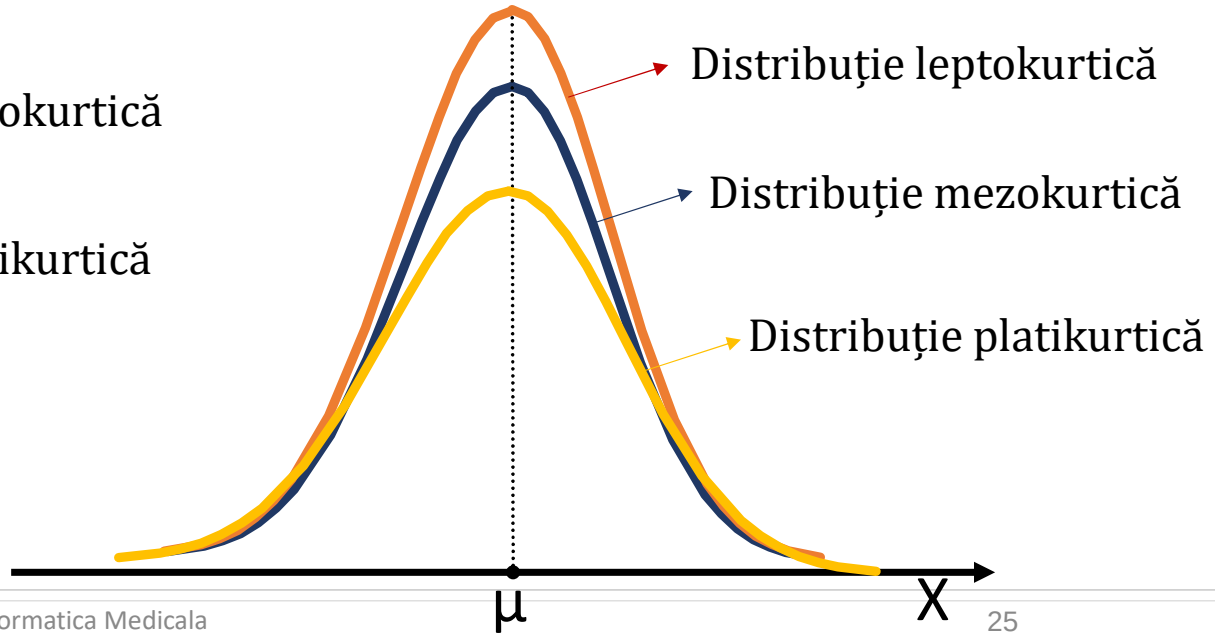


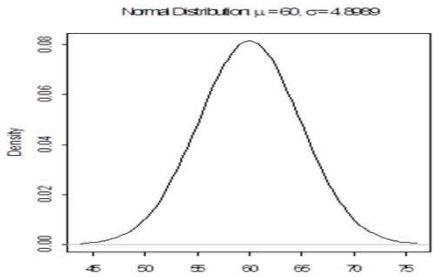
# Măsuri ale simetriei și boltirii: Coeficientul de boltire

- **Coeficientul de boltire** (notație: $\alpha_4$): Măsoară gradul de aplatizare al unei distribuții de date în raport cu distribuția normală (Gaussiană): proprietea distribuției de a fi mai ascuțită sau mai aplatizată decât curba normală.
- Dacă:
  - ✓ $\alpha_4 \approx 0$ => distribuție mezokurtică (normală)
  - ✓ $\alpha_3 > 0$ => distribuție leptokurtică
  - ✓ $\alpha_3 < 0$ => distribuție platikurtică
-



# Verificarea condiției de normalitate a datelor (existența distribuției Normale)

- **Datele unei variabile cantitative sunt considerate a avea o distribuție Normală (de probabilitate) dacă frecvențele valorilor acestei variabile sunt similare cu frecvențele teoretice generate de o funcție matematică (legea Normală).**

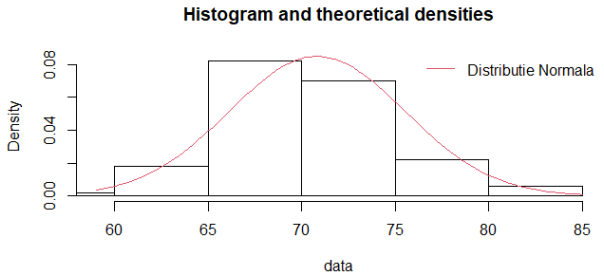



- **Utilitate:**

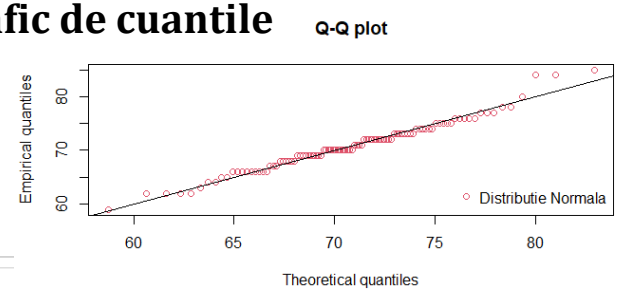
✓ Important pentru alegerea statisticilor descriptive

✓ De exemplu: pentru datele normale distribuite folosim media și abaterea standard, pentru datele distribuite nenormale folosim mediana și quartilele

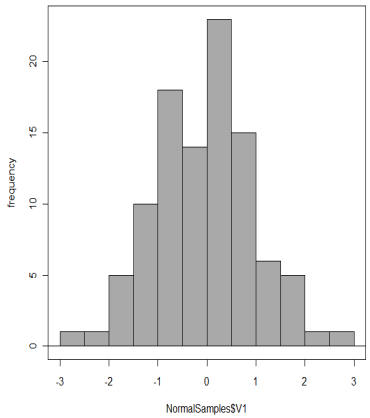
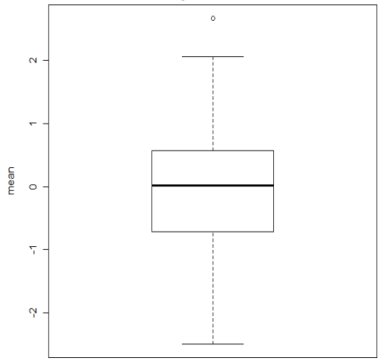
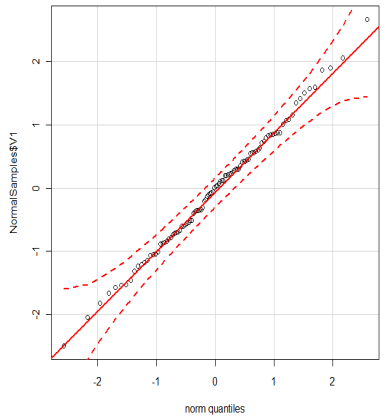
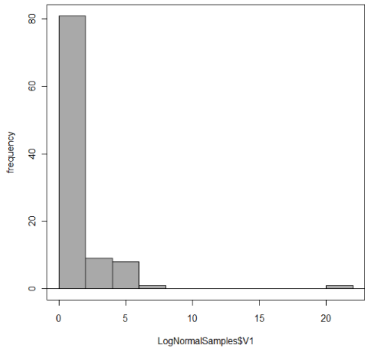
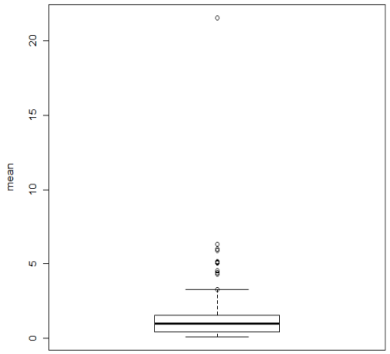
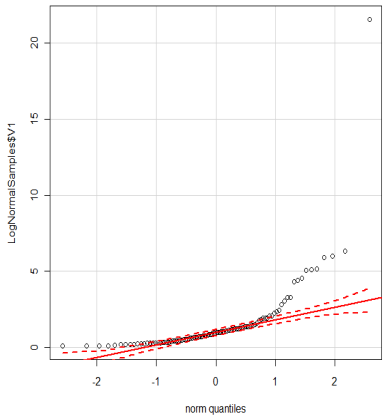
✓ Important pentru alegerea tehnicilor statistice analitice/inferențiale




**Grafic de cuantile**



# Date normal distribuite versus Date care au deviații de la normalitate

| | Histograma | Box plot | Grafic de cuantile |
|---|---|---|---|
| **Date normal distribuite** |  |  |  |
| **Date care <u>NU sunt</u> normal distribuite** |  |  |  |

# Verificarea condiției de normalitate a datelor (existența distribuției Normale)

**Metode de verificare:**

**1. modalități grafice (cele mai bune)**

- Histograma (simetrică, ca o pălărie)
- Boxplot (simetric în jurul mediei)
- **Density plot**
- **Graficul cuantilelor Q-Q plot**

**2. statistici descriptive (nu foarte fiabile)**

- ✓Dacă media este ~ = mediană
- ✓Dacă coeficientul de aplatizare/boltire ~ = 0 / aparține intervalului [-1, 1] (kurtoză)
- ✓Dacă coeficientul de simetrie ~ = 0 / aparține intervalului [-1, 1] (asimetrie)

**3. Teste de normalitate:** testul Shapiro-Wilk

# De reținut....

| | |
|---|---|
| **Amplitudinea** | • se calculează pentru Variabile cantitative discrete sau continue<br>• Slab informativă deoarece nu ne spune nimic cu provire la distribuția datelor |
| **Deviația standard** | • se calculează pentru Variabile cantitative<br>• Cel mai frecvent raportată relativ la medie<br>• Informații cu privire la cât de departe sunt datele față valoare centrală (cel mai frecvent media aritmetică – date normal distribuite) |
| **Cvartilele** | • se calculează pentru Variabile cantitative si calitative ordinale<br>•Utile ca mărimi ale dispersiei dacă datele nu urmează distribuția normală |
| **Coeficientul de variație** | • se calculează pentru Variabile cantitative<br>• Mărime a variabilității relative<br>• Permite compararea variabilității pe două eșantioane diferite |

# STATISTICI DESCRIPTIVE: DATE CALITATIVE
# RAPORTUL / PROPORȚIA / RATA

# Raportul

- Numere raționale pozitive a și b, b≠0
- **Raportul**: a/b
- Simbolistică
  - a:b
  - a/b
- Numitorul nu include în mod obligatoriu subiecții numărătorului
- Într-o grupă de 12 studenți avem 4 fumători.
  - Raportul nefumători/fumători = 8/4 = 2/1 = 2 ⇨ adică la 2 nefumători există un fumător
  - Raportul fumători/nefumători = 4/8 = 0,5

# Proporţia

» O **<u>proporţie</u>** este un raport în care numărătorul face parte din numitor. Astfel o proporţie are forma generală:

$$a/(a+b)$$

» Ia valori între:

> 0 și 1

> 0 și 100 dacă se exprimă procentual

» Toți indivizii de la numărător sunt incluși la numitor

» Frecvența relativă este o proporție (Prevalența unei boli este o proporție)

» Într-o grupă de 12 studenți avem 4 fumători.

» Proporția fumătorilor = 4/12=0,33

» Proporția nefumătorilor = 8/12

# Proporția:exemplu

- La serviciul de urgențe ale unui spital județean s-au prezentat pentru consultație 1200 pacienți. Dintre aceștia 420 au fost internați (200 femei și 220 bărbați):

  » Proporția pacienților internați = 420/1200*100 = 35%
  » Proporția pacienților de sex feminin internați = 200/420*100 = 48%
  » Proporția pacienților de sex masculin internați = 220/420*100 = 52%

# Prevalența

- = proporţia de indivizi dintr-o populaţie care au boala la un moment dat
- estimează probabilitatea ca un individ să aibă boala la un moment dat
- Formula:

Prevalența = (numărul de cazuri de boală) / (total populație)

# Rata

- O rată calculată reflectă riscul de a surveni în timp un anumit eveniment.
- Ia valori de la 0 la infinit
- Număr de indivizi raportat la unitatea de timp (oră / zi / săptămână / lună / an  etc.)
    - Rate de morbiditate
    - Rate de mortalitate
    - Rate de natalitate

# Rata

- Într-un oraș cu populație de 100000 locuitori s-au înregistrat 200 născuți vii în anul 1999
- Rata de natalitate = 200/100000 * 1000 = 2 nou născuți la mia de locuitori
- Rata de fertilitate = (nr. nașteri)/(nr. femei cu vârsta între 15-45 ani) * 1000
- Rata de morbiditate = frecvența cazurilor de îmbolnăvire pentru o populație specificată într-o perioadă de timp
- Rata de mortalitate = (nr. decese)/(populație)*1000

# Exemple de probleme pentru examen

**E1. Valorile tensiunii arteriale sistolice (mmHg) măsurate pe un eșantion de 10 pacienti sunt următoarele: 120, 100, 110, 120, 130, 160, 130, 120, 140, 160.**

**Media aritmetica, mediana, modulul si valoarea centrala sunt următoarele:**

A. 129 – 125 – 120 – 130

B. 130 – 125 – 130 – 125

C. 120 – 130 – 120 – 125

D. 129 – 130 – 120 – 130

E. 125 – 125 – 120 – 130

**R1: A**

*Rezolvare*-> **Media aritmetică(Ma)** =? Ma=(**120+100+110+120+130+160+130+120+140+160**) / 10 = 129 mmHg

-> **Mediana (Me)**=? Seria statistica ordonata: 100, 110, 120,120,120, 130,130,140,160,160

Talie esantion: numar par = 10=> $\frac{X_{\frac{10}{2}}+X_{\frac{10}{2}+1}}{2}=\frac{X_5+X_6}{2}=125$

-> **Modulul (Mo)**=? Valoarea din seria statistică a TAS cu cea mai mare frecvență (cea mai repetitiva valoare)=120

-> **Valoare centrala**=? $\frac{x_{max}+x_{min}}{2}=130$

# Exemple de probleme pentru examen

**E2. A fost evaluat un eșantion de studenți AMG1 cuprinzând 40 de băieți si 40 de fete. Pentru fiecare student s-au colectat înălțimea si greutatea si s-a calculat indicele de masă corporală (IMC). Valoarea medie și deviația standard ale IMC au fost de 20 kg /$m^2$ respectiv 4 kg/$m^2$. Din următoarele afirmații, selectați cele pe care le considerați adevărate:**

A. Volumul eșantionului în acest studiu a fost de 40.

B. Dacă valorile IMC au avut o distribuție simetrică atunci jumătate dintre studenți au avut valorile IMC ≤ 20 kg/$m^2$.

C. Coeficientul de variație a fost egal cu 20%.

D. IMC este relativ eterogen.

E. Coeficientul de variație a fost independent de unitățile de măsură

**R2. B, C, D, E**

# Exemple de probleme pentru examen

**E4. Valorile tensiunii arteriale sistolice (mmHg) măsurate pe un eșantion de 10 pacienți sunt următoarele: 120, 100, 110, 120, 130, 160, 130, 120, 140, 160. Se cunoaște ca prima cvartilă (Q1) este egala cu 120, a doua cvartilă (Q2) este egala cu 125 si a treia cvartilă (Q3) este egala cu137.5. Care dintre următoarele afirmații sunt corecte**

A. Q2 - Q1 = 5

B. Q3 - Q2 = 12

C. Q3 - Q2 = 12,5

D. Distribuția este aproximativ simetrica

E. Distribuția este simetrica

**R4: A, C, E**

*Rezolvare:* **->** Într-o distribuție simetrică distanța dintre cvartilele (Q1 si Q3) si mediana (Q2) este aceeasi=> Q2 - Q1 =Q3 - Q2 ; in cazul nostru, cei doi termeni (Q3-Q2 si Q2-Q1) nu sunt egali deci seria statistica a valorilor TAS nu este simetrica.

# Exemple de probleme pentru examen

**E5. Pentru o variabila cu distribuție simetrica:**

A. Media (aritmetică) este egala cu prima cvartilă

B. 50% din date sunt superioare mediei

C. Q3 – Q2 = Q2– Q1

D. Q3 – Q2 = 3*(Me – Q1)

E. Modul≈ Medie ≈ Mediana

**R5: B, C, E**

# MULȚUMESC PENTRU ATENȚIE!