

# TESTAREA IPOTEZELOR STATISTICE

## TESTAREA NORMALITĂȚII DATELOR

# Despre ...

---

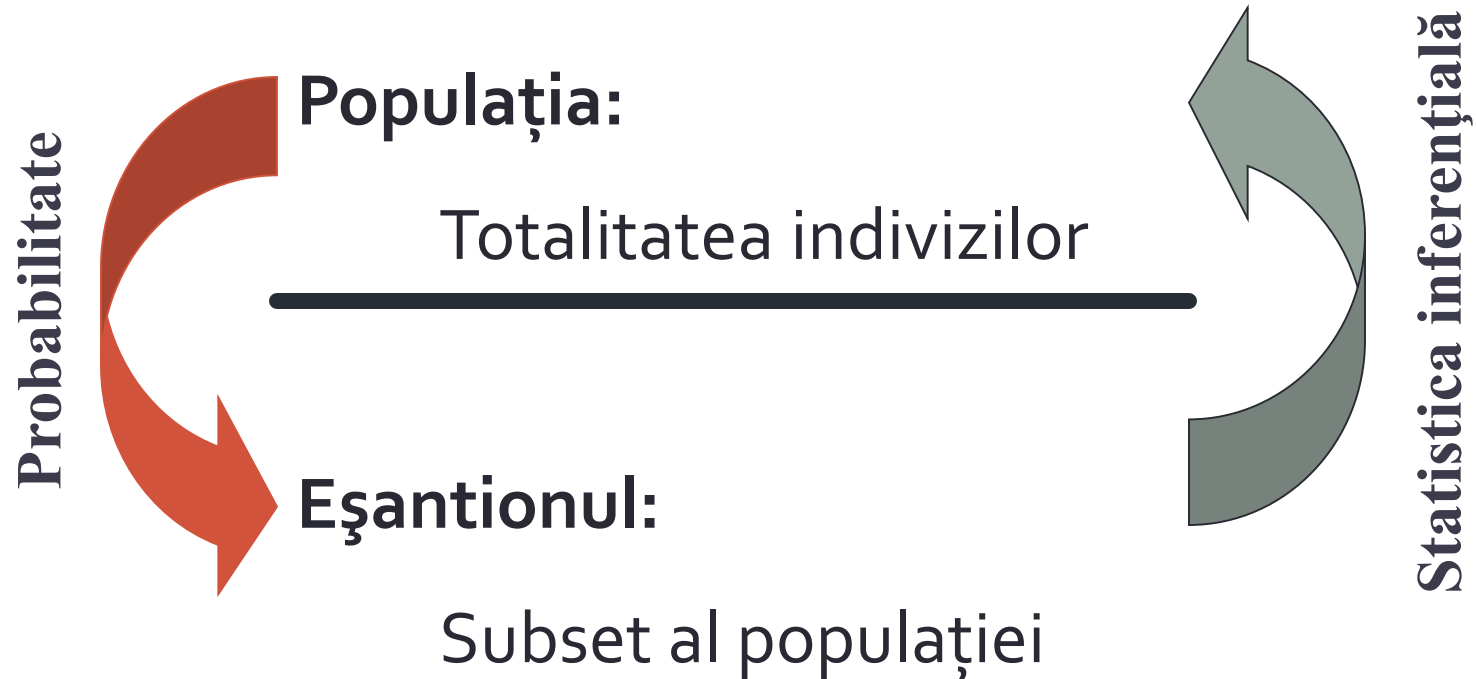
- Pașii unui test statistic
- Teste de normalitate
- Distribuția normală și alegerea testului statistic

# Definiții

- **Test statistic** = metodă a deciziei medicale prin utilizarea datelor experimentale.
- Un rezultat se numește semnificativ statistic dacă este puțin probabil să apară datorită întâmplării
- Ipoteza statistică = asumție asupra parametrului populației. Această asumție poate sau nu să fie adevărată.
- Ipoteza clinică = o idee explicativă care permite structurarea datelor cu privire la un pacient în așa fel încât să ducă la o mai bună înțelegere a patologiei sau respectiv la o decizie medicală corectă.

[Lazare A. The Psychiatric Examination in the Walk-In Clinic: Hypothesis Generation and Hypothesis Testing. Archives of General Psychiatry 1976;33:96-102.]

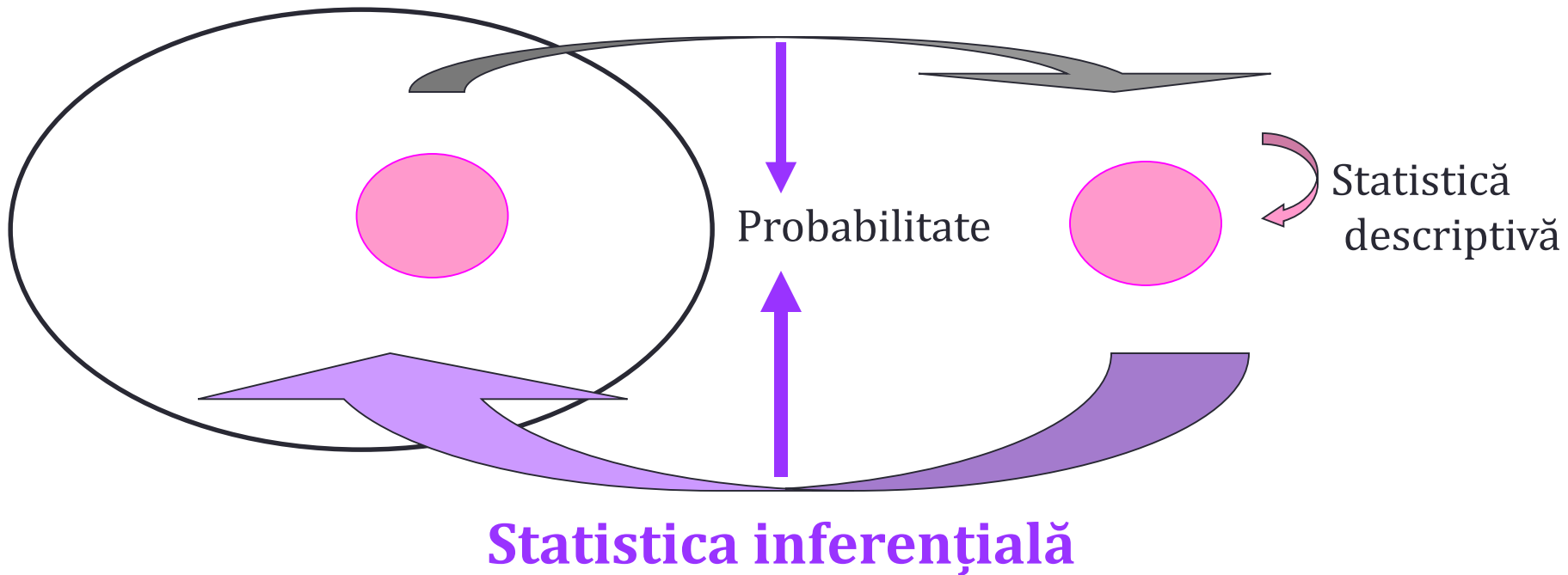
# Testarea ipotezelor



Populația  
Parametrul ( $\mu, \sigma/s$ )

Eșantionul  
Statistica ( $\bar{X}, S$ )

Eșantionare



# Statistica inferențială

- Realizăm un studiu pe un eșantion
- Întrebarea cheie în statistica inferențială este:
  - Ar putea ca întâmplarea singură să producă un eșantion ca al nostru?
- 2 interpretări ale tiparelor în date:

Inferența  
statistică  
separă



**Întâmplarea:**

Fluctuații datorate șansei

---



**Erori sistematice+ Întâmplarea:**

Diferențe adevărate în populație

Erori în design-ul experimental

# Testarea ipotezelor...

Concepte de bază:

- Se estimează parametrul populației.
- 2 ipoteze statistice, ipoteza nulă și respectiv ipoteza alternativă.
  - **Presupunem că ipoteza nulă este adevărată.**
- 2 decizii posibile:
  - Concluzionăm că există suficiente evidențe care să susțină ipoteza alternativă → respingem ipoteza nulă
  - Concluzionăm că NU există suficiente dovezi care să susțină ipoteza alternativă → ipoteza nulă nu se poate respinge

# Testarea Ipotezelor: Pasul 1

- Transpune problema de cercetat în termeni statistici
  - Ipoteza nulă (ipoteza statistică care urmează a fi testată):  
abreviată ca  $H_0$ 
    - “Nimic interesant nu se întâmplă”
  - Ipoteza alternativă (ipoteza care într-un anumit sens contrazice ipoteza nulă): abreviată ca  $H_1$  sau  $H_a$ 
    - Ceea ce cercetătorul crede că se întâmplă
    - Poate să fie unilaterală sau bilaterală
- Ipotezele statistice se referă la parametrii populației

Unilateral	Bilateral
$H_0: \mu = 110$	$H_0: \mu = 110$
$H_1: \mu < 110$ SAU	$H_1: \mu \neq 110$
$H_1: \mu > 110$	



# Testarea Ipotezelor: Pasul 2

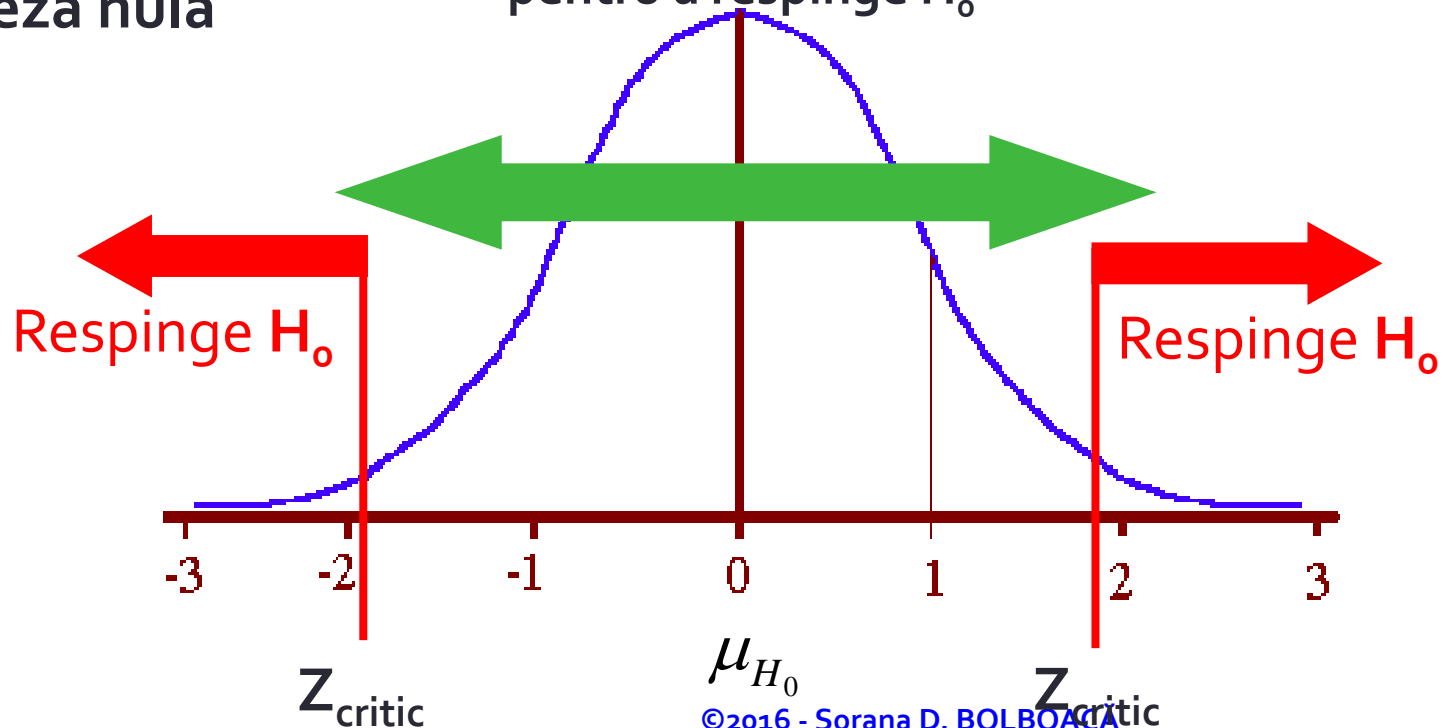
- Definiți regiunea critică:
  - Decideți care valoare  $p$  ar fi “mai puțin probabilă”
  - Această valoare prag se numește nivel de semnificație sau prag alfa
  - Atunci când probabilitatea asociată parametrului eșantionului este mai mică decât această valoare prag se spune că rezultatul este semnificativ statistic
  - De obicei nivelul alfa este de 0,05 sau 0,01
- Nivelul alfa (nivelul de semnificație) = probabilitatea erorii de tip I (probabilitatea de a respinge  $H_0$  atunci când ea este adevărată)
- Probabilitatea erorii de tip II este probabilitatea de a accepta ipoteza nulă în condițiile în care ipoteza alternativă este adevărată. Probabilitatea erorii de tip II se abreviază cu  $\beta$ .

# Testarea Ipotezelor: Pasul 3

- Regiunea critică:
  - Dacă valoarea parametrului statistic aparține regiunii critice, ipoteza nulă  $H_0$  va fi respinsă și va fi acceptată ipoteza alternativă  $H_1$ .
  - Dacă valoarea parametrului statistic nu aparține regiunii critice, ipoteza nulă  $H_0$  va fi acceptată.

Ipoteza nulă

Nu există suficiente dovezi  
pentru a respinge  $H_0$



# Testarea Ipotezelor: Pasul 4

- Calculează parametrul testului
- Parametrul statistic al testului aplicat (ex.  $Z_{\text{test}}$ ,  $T_{\text{test}}$ , or  $F_{\text{test}}$ ) este informația care se va utiliza pentru a decide dacă respingem sau nu ipoteza nulă.

# Testarea Ipotezelor: Pasul 5

- Concluzia statistică:
  - În principiu nu acceptăm niciodată ipoteza nulă; ipoteza nulă o respingem sau nu avem suficiente dovezi să o respingem
  - Concluzia statistică referă  $H_0$

# Testarea ipotezelor statistice

1. Scrieți ipotezele statistice ( $H_0$  și  $H_1$ )
2. Alegeți nivelul de semnificație
3. Stabiliți regiunea critică
4. Calculați statistica testului și valoarea  $p$  asociată
5. Stabiliți concluzia statistică a testului

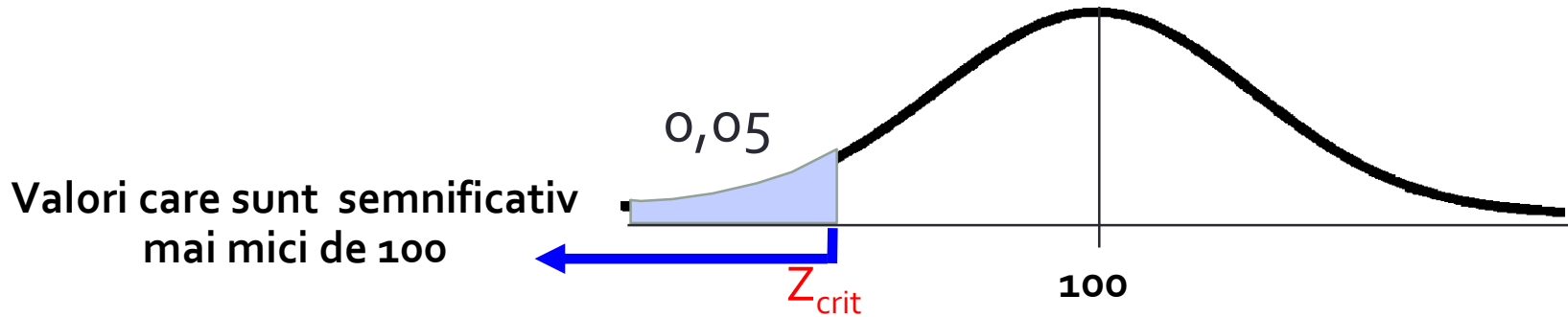
# Testul unilateral sau bilateral

- Testul unilateral se folosește când:
  1. Modificările în direcția opusă este lipsită de sens
  2. Modificările în direcția opusă nu este de interes
  3. Nici o teorie nu prezice schimbarea în direcția opusă
- Prin convenție în științele medicale se folosește testul bilateral
- De ce? Testul este mai conservativ.

# Test unilateral

Respinge  $H_0$

Nu avem suficiente dovezi să respingem  $H_0$

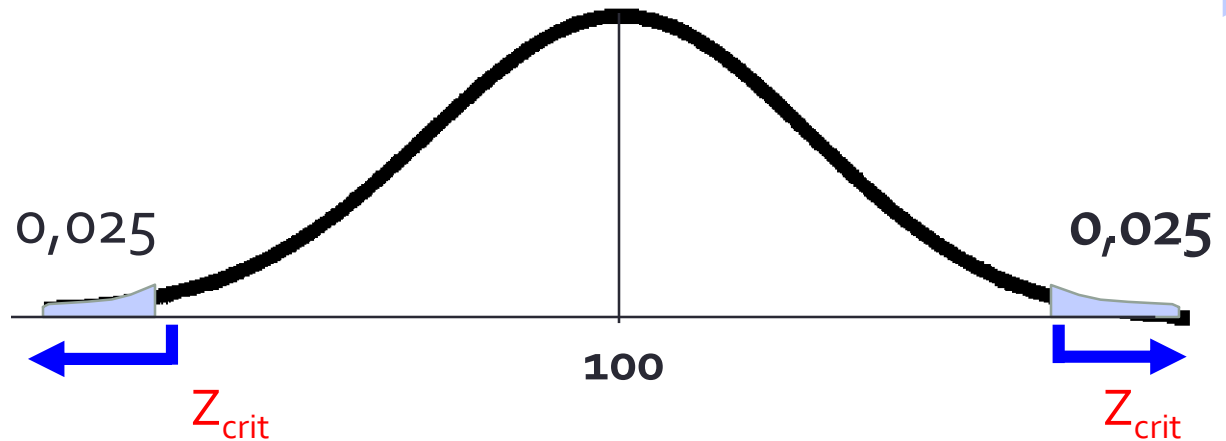


# Test bilateral

Respinge  $H_0$

Nu avem suficiente dovezi să respingem  $H_0$

Respinge  $H_0$



Valori care diferă semnificativ de 100

# Semnificația statistică ...

- $p > 0,05$  → rezultatele sunt considerate nesemnificative statistic;
- $0,05 < p \leq 0,1$  → există o oarecare tendință spre semnificație statistică;
- $0,01 < p \leq 0,05$  → rezultatele sunt semnificative statistic;
- $0,001 < p \leq 0,01$  → rezultatele sunt înalt semnificative statistic;
- $p \leq 0,001$  → rezultatele sunt foarte înalt semnificative statistic.

Realitate

Ce rezultă din  
cercetare

	$H_0 = \text{Adevărată}$	$H_0 = \text{Falsă}$
$H_0$ respins	Eroare de tip I	Decizie corectă
$H_0$ nu se respinge	Decizie corectă	Eroare de tip II

- Eroarea de tip I = respingerea ipotezei nule atunci când aceasta este adevărată
- Eroarea de tip II = nu avem suficiente dovezi srespingem  $H_0$  atunci când  $H_A$  e adevărată



# Diferența între valoarea $p$ și intervalul de încredere

- Valoarea  $p$  măsoară puterea evidenței împotriva ipotezei nule.
- $p$  este probabilitatea de a obține un rezultat extrem dacă ipoteza nulă este adevărată.
- Valoarea  $p$  măsoară semnificația statistică
- Intervalul de încredere oferă un interval de valori care permite interpretarea clinică a rezultatelor

# Intervalul de confidență vs. test statistic

- Un eșantion de 50 studenți a fost întrebat câte ore învață în medie pe zi. Studenții din eșantion învață în medie 1,2 ore pe zi cu o deviație standard de 0,6. Datele sunt ușor asimetrice spre stânga.

Estimați numărul adevărat de ore pe zi dedicate studiului utilizând intervalul de confidență de 95%

$$n=50, \text{ media} = 1,20, s=0,6$$

$$\text{media} \pm z * ES \text{ (ES} = s/\sqrt{n}) \rightarrow 1,20 \pm 1,96 * 0,07 \rightarrow [1,06; 1,34]$$

→ Suntem 95% siguri că studenții învață în medie între 1,06 și 1,34 ore pe zi

# Intervalul de încredere vs. test statistic

- Un eșantion de 50 studenți a fost întrebat câte ore învață în medie pe zi. Studenții din eșantion învață în medie 1,2 ore pe zi cu o deviație standard de 0,6. Datele sunt ușor asimetrice spre stânga. Este această valoare semnificativ diferită de zero?
  - IC95% [1,06; 1,34] → Valoarea este semnificativ diferită de 0 deoarece valoarea 0 nu este cuprinsă în intervalul de încredere.
  - Test:  $H_0: \mu = 0$  vs.  $H_1: \mu \neq 0$
  - $Z = (\text{media}-0)/ES = (1,20-0)/0,07 = 17,14$  ( $p < 0,00001$ )
    - Respingem ipoteza nulă cu un risc de eroare de 5%

# Semnificația statistică vs. Semnificația clinică

Studiu de cohortă n=34,079 femei → greutatea câștigată de grupul care a efectuat exercițiu fizic >21 ore/săptămână față de grupul cu <7,5 ore de exercițiu fizic/săptămână (p<0,001)

**Table 2.** Mean (SD) Differences in Weight Over Any 3-Year Period by Physical Activity Level, Women's Health Study, 1992

Group	No. of Women <sup>b</sup>	Physical Activity, MET Hours per Week			P Value for Trend
		<7.5	7.5 to <21	≥21	
All women Analytical model <sup>c</sup> 1		0.15 (0.04)	0.12 (0.04)	0 [Reference]	<.001

- Mărimea efectului ? <7,5 ore/săpt → 0,15 kg în plus față de grupul cu ≥21
- Extrapolare: dacă studiul se continuă 13 ani grupul cu ≥21 va acumula în greutate cu 0,635 kg mai puțin comparativ cu grupul cu < 7,5!

# De reținut!

- Structură generală de urmat în testarea ipotezelor statistice este aceeași indiferent de testul care se dorește a fi aplicat.
- Alegerea testului statistic se face în funcție de ipoteza de cercetare.
- Test statistic vs. Interval de încredere
  - → ambele acceptă sau nu  $H_0$

# TESTAREA NORMALITĂȚII DATELOR CANTITATIVE

De ce? Alegerea testului statistic se face în funcție de distribuția datelor care se testează

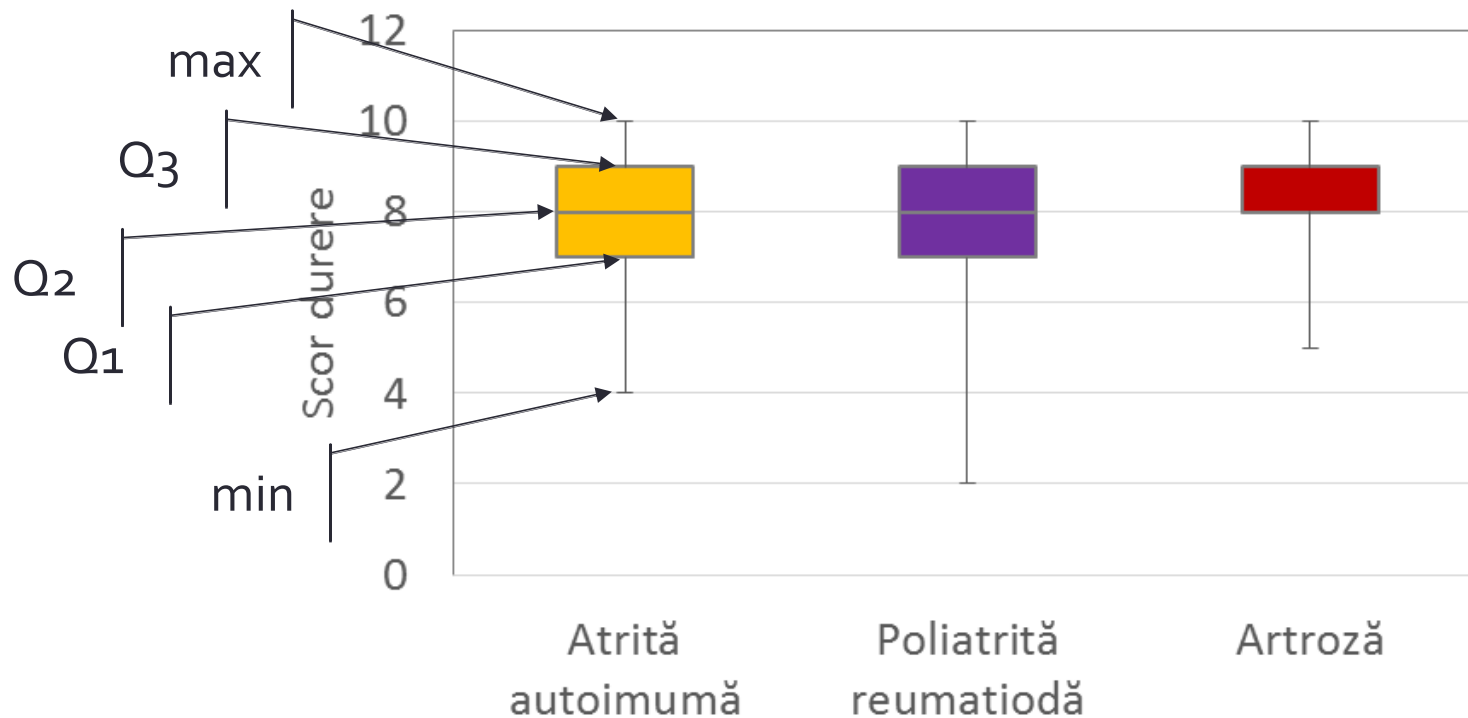
- Intervalul de încredere pentru medie are sens a fi calculat dacă datele cantitative urmează distribuția normală

# Modalități de evaluare a normalității

- Reprezentarea grafică:
  - grafic de tip cutie cu mustăți (boxplot)
  - histograma
  - ~~graficul cuantilă-cuantilă~~
- Parametrii statistici descriptivi
  - Media, mediana și modulul au valori apropiate
  - Asimetria are valoare în intervalul  $(-1, 1)$
  - Boltirea are valoare în intervalul  $(-0,5; +0,5)$
- Teste statistice:
  - $H_0$ : Datele sunt normal distribuite
  - $H_A$ : Datele nu urmează distribuția normală

# Normalitate: Grafic cutie cu mustăți

- S-a evaluat scorul de durere cu scala vizuală la 3 grupuri de pacienți: aprite autoimune ( $n_1=30$ ), poliartrită reumatoidă ( $n_2=40$ ), și artroză ( $n_2=50$ )

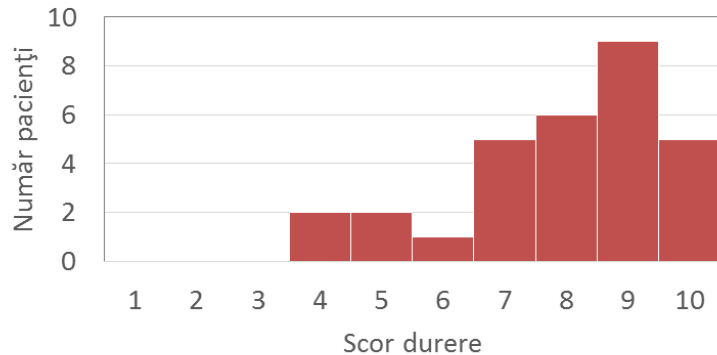




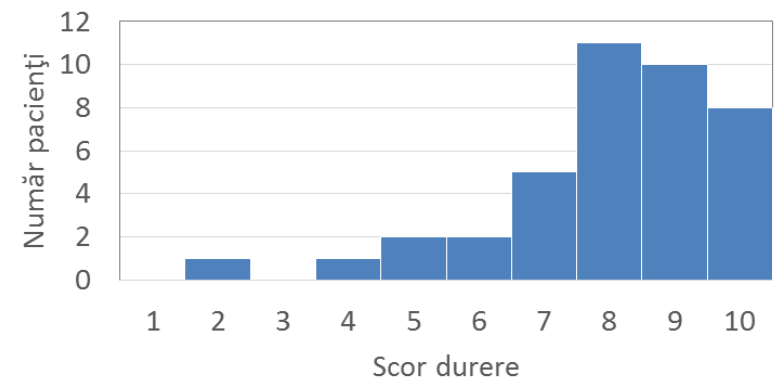
# Normalitate: Histograma

- S-a evaluat scorul de durere cu scala vizuală la 3 grupuri de pacienți: artrite autoimune ( $n_1=30$ ), poliartrită reumatoidă ( $n_2=40$ ), și artroză ( $n_3=50$ )

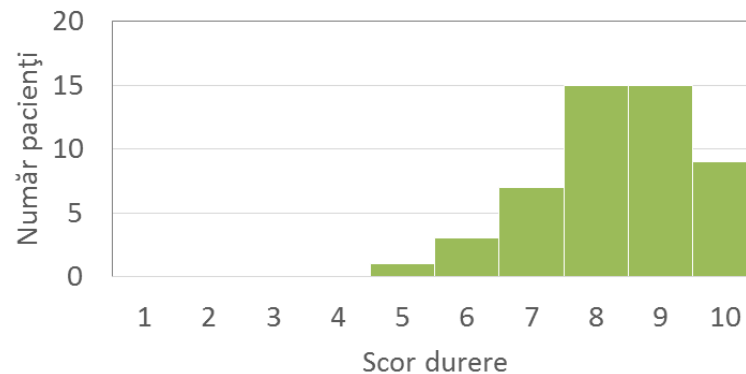
Artrită autoimună



Poliartrită reumatoidă



Artroză



# Normalitate: Parametrii statistici descriptivi

	Arită autoimună	Poliarită reumatoidă	Artroză
media	7,93	8,03	8,34
mediana	8	8	8
modulul	9	8	8,9
asimetria	-0,90	-1,37	-0,55
boltirea	0,12	2,34	-0,01

# Normalitate: Teste statistice

1933

	Mai puțin conservativ	conservativ
Small sample (5-50)	Shapiro-Wilk	Kolmogorov-Smirnov
Medium to large sample (>50)	Shapiro-Wilk	Chi-Square Goodness-of-Fit

1965

# Normalitate: Teste statistice

Pentru fiecare din bolile studiate

- $H_0$ : Scorul de durere urmează distribuția normală
- $H_A$ : Scorul de durere nu urmează distribuția normală

Variable	Tests of Normality (Spreadsheet1)					
	N	max D	K-S p	Lilliefors p	W	p
Artrita autoimuna	30	0.196661	p < .20	p < .01	0.885503	0.003781
Poliartrita reumatoida	40	0.219429	p < .05	p < .01	0.865100	0.000214
Artroza	40	0.219429	p < .05	p < .01	0.865100	0.000214

Interpretare

- Dacă  $p < 0,05 \rightarrow$  se respinge  $H_0$  (datele nu urmează distribuția normală)

# Test parametric vs. Test neparametric

	Parametric	Non-Parametric
Distribuția	Normală	oricare
Varianța	Omogene	oricare
Scala de măsură	Rație / Interval	oricare
Valoarea centrală	Media	mediana
Dispersia	Deviația standard	(Q1; Q3)

	Parametric	Non-Parametric
2 grupuri independente	Independent t-test	Mann-Whitney
2 grupuri dependente	Paired t-test	Wilcoxon test
> 2 grupuri	ANOVA	Kruskal-Wallis test Friedman's ANOVA
Corelația	Pearson	Spearman, Kendall, etc.
...	...	...

# MULȚUMESC PENTRU ATENȚIE!



Statistics is the grammar of science.

— *Karl Pearson* —

AZ QUOTES