



UNIVERSITATEA DE MEDICINĂ,  
FARMACIE, ȘTIINȚE ȘI TEHNOLOGIE  
„GEORGE EMIL PALADE”  
DIN TÂRGU MUREȘ

# PROBABILITĂȚI ȘI STATISTICĂ ÎN SISTEME MEDICALE

Cursul 3 - completare, 08.10.2020

## STATISTICA DESCRIPTIVĂ

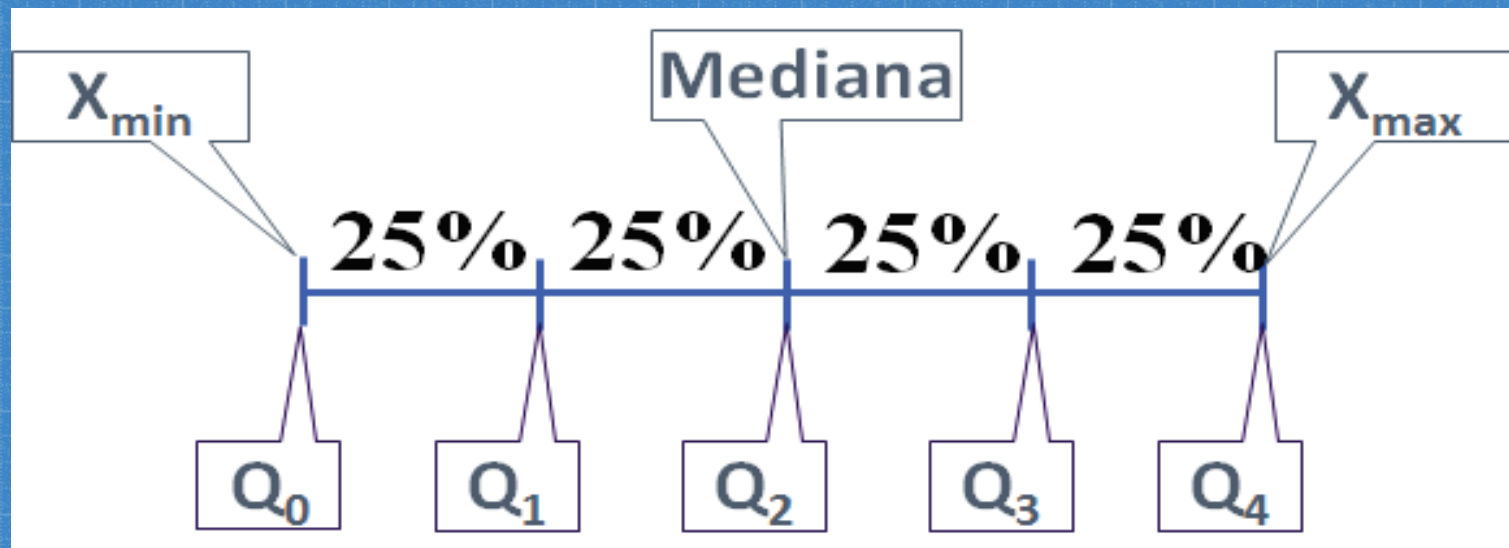
prof. univ. dr. habil Manuela Rozalia GABOR



# INDICATORI AI DISPERSIEI. QUARTILE/PERCENTILE

În cazul percentilelor, seria statistică este împărțită în părți egale în funcție de volumul eșantionului. Cele mai utilizate percentile sunt cvartilele ( $Q_1$  = percentila de 25%,  $Q_2$  = percentila de 50% (mediana) și  $Q_3$  = percentila de 75%) și decilele (percentilele de 10%, 20%, ..., 90%).

În cazul cvartilelor, seria statistică este împărțită în 4 părți egale:





- Cuartila zero (**Q0**) are valoarea egală cu valoarea minimă din seria statistică.
- Cuartila 1 (**Q1**) este acea valoare care împarte seria statistică în două, astfel încât 25% dintre datele seriei să fie mai mici sau egale cu valoarea Q1, iar 75% mai mari decât valoarea Q1.
- Cuartila 2 (**Q2**) este acea valoare care împarte seria statistică în două părți egale. Are aceeași valoare ca și mediana.
- Cuartila 3 (**Q3**) este acea valoare care împarte seria statistică în două, astfel încât 75% din datele seriei sunt mai mici sau egale cu valoarea Q3, iar 25% mai mari decât valoarea Q3.
- Cuartila 4 (**Q4**) este egală cu valoarea maximă.



Cuartilele și respectiv percentilele sunt utilizate în sumarizarea datelor atunci când parametrul de centralitate care descrie cel mai bine seria statistică este mediana. Are sens calcularea cuartilelor și/sau percentilelor pentru variabile ordinale (calitative), interval (cantitative) sau raport (cantitative).

## EXEMPLU

Fie greutatea la naștere a primilor 20 nou-născuți cu restricție de creștere intrauterină cu vârstă gestațională egală cu 40 și respectiv 41 săptămâni incluși într-un studiu.

i	1	2	3	4	5	6	7	8	9	10
$x_i$	2450	2500	1900	2700	2650	2190	2600	2800	2600	2490

i	11	12	13	14	15	16	17	18	19	20
$x_i$	2200	2350	2700	2650	2660	2700	2600	2900	2350	2490

Ce putem spune despre dispersia datelor acestei serii statistice?



## SOLUȚIA

$i$	$x_i$
1	2450
2	2500
3	1900
4	2700
5	2650
6	2190
7	2600
8	2800
9	2600
10	2490
11	2200
12	2350
13	2700
14	2650
15	2660
16	2700
17	2600
18	2900
19	2350
20	2490

ordonare

$x_i$
1900
2190
2200
2350
2350
2450
2490
2490
2500
2600
2600
2600
2650
2650
2660
2700
2700
2700
2800
2900

$$Q1 = 2425$$

$$Me = (x_{10} + x_{11}) / 2$$
$$Me = 2600$$

$$Q3 = 2670$$

$$A = 2900 - 1900 = 1000$$



Pentru această serie statistică, interpretarea parametrilor este următoarea:

- Jumătate din subiecți au avut greutatea la naștere mai mică sau egală cu 2600g.
- 25% din subiecți au avut greutatea la naștere mai mică sau egală cu 2425g.
- 75% din subiecți au valoarea greutatea la naștere mai mică sau egală cu 2670g.
- Diferența dintre valoarea maximă și minimă la naștere a fost de 1000g.



# ALEGEREA PARAMETRILOR DE CENTRALITATE ȘI DISPERSIE

Parametrii de centralitate și dispersie se pot calcula pe orice tip de variabilă indiferent de scala de măsură. Dar calcularea lor are sens doar pentru anumite variabile și respectiv scale de măsură

Tabelul prezintă tipurile de date și recomandările pentru calcularea parametrilor de centralitate și dispersie

Parametrul	Date calitative		Date cantitative	
	Nominale	Ordinale	Interval	Raport
Media	Nu	Nu	Da*	Da*
Mediana	Nu	Da	Da	Da
Modulul	Da	Da**	Da**	Da**
Amplitudinea	Nu	Da	Da	Da
Deviația standard	Nu	Nu	Nu	Da*

\* Dacă datele sunt normal distribuite

\*\* Nu este recomandat