



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

**Probabilități și statistică în  
sisteme medicale**

**REPREZENTAREA  
TABELARĂ ȘI GRAFICĂ  
A DATELOR MEDICALE**

prof. univ. dr. habil MANUELA ROZALIA GABOR

CURSUL 6 – 08.10.2020

# REPREZENTAREA TABELARĂ A DATELOR MEDICALE

Tabelul este cea mai simplă metodă de sumarizare a datelor unei serii statistice și se poate face atât pentru datele calitative, cât și pentru datele cantitative. Metodele de sumarizare tabelară cele mai des utilizate sunt reprezentate de tabelul de frecvență și tabelul de contingență și vor fi prezentate în această secțiune.

Prezentarea datelor brute nu va duce la o interpretare corectă a acestora, interpretarea fiind cu atât mai dificilă cu cât numărul de date din seria statistică este mai mare. Sumarizarea tabelară permite reducerea datelor și prezentarea acestora într-o formă cât mai simplă posibilă care însă să ajute la identificarea (dacă există) tiparului pe care datele îl urmează.

Există o serie de avantaje ale sumarizării datelor cu ajutorul tabelele:

- Rezultatele sunt prezentate într-o formă compactă.
- Rezultatele sunt ușor de interpretat.
- Permite observarea existenței unor modele.
- Permite compararea a două sau mai multe grupuri.
- Permite utilizarea parametrilor cantitativi.

Bineînțeles că acest proces nu este lipsit de dezavantaje. Detaliile datelor brute se pierd în momentul sumarizării și acest proces este ireversibil.

Indiferent de forma sumarizării tabelare, tabelele trebuie să aibă următoarele caracteristici:

- Trebuie să se explice singure. Citirea tabelului trebuie să fie posibilă fără a citi textul care însoțește tabelul. Astfel:
  - Semnificațiile abrevierilor și a simbolurilor utilizate în tabel trebuie definite la subsolul acestuia.
  - Rândurile și coloanele trebuie definite. De asemenea, trebuie specificat, acolo unde este cazul, care sunt unitățile de măsură utilizate.
  - Tabelul trebuie să aibă un titlu clar și la subiect.
  - Este recomandat ca tabelele să aibă rânduri și/sau coloane de sinteză (cu totaluri).
- Dacă datele nu sunt ale dvs. trebuie să obțineți acordul de utilizare al acestora de la persoana sau instituția care deține drepturile de proprietate intelectuală. În acest caz aveți, de asemenea, obligația de a menționa la subsolul tabelului sursa datelor.

Tabelul de frecvență se utilizează pentru sumarizarea unei singure variabile (serie statistică unidimensională sau univariată).

**Exemplul 2.16.**

Greutatea la naștere a primilor 120 copii cu restricție de creștere intrauterină incluși în studiu sunt redată mai jos:

740	890	1600	1600	2100	1800	2150	2140	2940	2400	1900	2460
900	990	1560	1900	1990	1660	2200	2100	2400	2490	2460	2600
990	1170	1490	1300	1870	2400	2300	2300	1900	1780	2600	2200
950	1249	1300	1560	2160	1990	2150	1850	2400	2200	2600	2300
780	1100	1400	1100	1800	2300	2200	2550	2200	2330	2490	2480
890	1300	1450	1900	1990	2300	2200	2400	1990	1800	2600	2100
900	1100	1970	1900	990	2700	2400	2500	2000	2450	2500	2650
960	1350	1490	1900	2250	2100	2350	2000	2500	1900	2400	2490
900	1350	1550	2000	1680	2280	2100	2300	2200	2490	2200	2700
800	1225	1250	2100	1600	2430	1900	2490	2450	2700	2700	2650

Care este forma cea mai potrivită de sumarizare tabelară a acestor date?

Datele brute din tabelul anterior pot fi sumarizate cu ajutorul tabelului de frecvență.

# TABELUL DE FRECVENȚĂ



---

Organizarea în tabelul de frecvență a datelor se face prin includerea valorilor distincte pe care variabila le poate lua (pentru variabilele calitative, scala nominală sau ordinală, respectiv pentru datele cantitative discrete) sau respectiv clasa de frecvență (pentru variabilele cantitative continue – se aplică obligatoriu, dar se poate aplica și celor discrete) și respectiv a frecvenței de apariție. În construirea unui tabel de frecvență, frecvențele raportate pot fi:

---

---

Absolute: număr de cazuri care iau valoarea de interes sau au valoare în clasa de frecvență de interes.

Relative: numărul de cazuri/total cazuri.

Cumulate (crescător/descrescător):

frecvența absolută cumulată crescător asociată unei valori  $x$  este egală cu suma frecvențelor absolute ale tuturor valorilor seriei care sunt mai mici sau egale cu valoarea  $x$ .

- *Frecvența absolută cumulată* descrescător asociată unei valori  $x$  este egală cu suma frecvențelor absolute ale tuturor valorilor seriei care sunt mai mari sau egale cu valoarea  $x$ .

○ *frecvența relativă cumulată* crescător asociată unei valori  $x/n$  este egală cu suma frecvențelor relative ale tuturor valorilor seriei care sunt mai mici sau egale cu valoarea  $x$ .

○ *frecvența relativă cumulată* descrescător asociată unei valori  $x/n$  este egală cu suma frecvențelor relative ale tuturor valorilor seriei care sunt mai mari sau egale cu valoarea  $x$ .

---

Clasa de frecvență	Frecvența absolută	Frecvența relativă (%)	Frecvența absolută cumulată crescător	Frecvența relativă cumulată crescător
<1000g	13	$=13/120*100 = 11$	=13	=11
1000-1499	16	$=16/120*100 = 13$	= 16+13 = 29	=13+11 = 24
1500-2499	76	$= 76/120*100 = 63$	= 76+29 =105	= 63+24 = 87
≥2500	15	$=15/120*100 = 13$	= 15+105 = 120	=13+87 = 100
Total	120	100		

Datele brute din exemplul 2.16. Se vor sumariza folosind următoarele clase de frecvență: greutate normală ( $\geq 2500\text{g}$ ), greutate mică (de la 1500g la 2499g), greutate foarte mică (de la 1000g la 1499g) și respectiv greutate extrem de mică (<1000g):

Analizând tabelul de frecvență anterior putem caracteriza seria statistică:

- Majoritatea subiecților cu restricție de creștere intrauterină au greutate mică la naștere.
- Eșantionul studiat conține un număr/procent similar de nou născuți cu greutate mică la naștere și respectiv greutate considerată în limite normale (13%).
- Numărul de copii cu greutate mai mică decât normală la naștere a fost egal cu 105.
- 24% din nou născuții din eșantion au avut greutate foarte mică sau extrem de mică la naștere.

## EXEMPLUL 2

Scorul de durere la prezentarea la fizioterapie a 100 pacienți cu artroză a fost colectate prin utilizarea scalei vizuale de durere (0=fără durere, 10 = durere extrem de puternică). Datele brute colectate sunt prezentate mai jos:

9	8	9	6	5	8	10	10	10	9
7	10	9	10	10	8	8	10	8	10
8	9	9	8	6	9	8	7	8	10
8	8	9	10	10	8	10	7	9	8
8	8	9	9	7	6	10	6	10	9
9	9	8	8	8	7	8	6	8	9
10	9	7	10	10	9	7	10	9	7
8	8	6	7	10	6	7	9	7	8
9	7	8	9	9	9	8	9	8	9
7	7	8	9	8	9	9	10	9	8



Scor durere	Nr. cazuri	%	Frecvența absolută cumulată crescător	Frecvența absolută cumulată descrescător
0	0	0	=0	=100-0=100
1	0	0	=0+0	=100-0=100
2	0	0	=0+0	=100-0=100
3	0	0	=0+0	=100-0=100
4	0	0	=0+0	=100-0=100
5	1	1	=1+0=1	=100-1=99
6	7	7	=7+1=8	=99-7=92
7	14	14	=14+8=22	=92-14=78
8	29	29	=29+22=51	=78-29=49
9	29	29	=29+51=80	=49-29=20
10	20	20	=20+80=100	=20-20=0
Total	100	100		

Analiza datelor din tabelul de frecvență evidențiază următoarele:

- Pacienții cu artroză au un scor al durerii de la moderat ( $x_{\min} = 5$ ) la durere extrem de puternică ( $x_{\max} = 10$ ).
- 99% din pacienții au un scor al durerii mai mare sau egal cu 5.

- 50% din pacienți resimt o durere puternică (scor durere = 8) sau foarte puternică (scor durere = 9).
- 100% din pacienți au scorul durerii mai mic sau egal cu 10, ceea ce indică absența datelor lipsă.



# TABELE DE CONTINGENȚĂ

Tabelul de contingență permite sumarizarea în același timp a mai mult de o variabilă, cu scopul de a se observa relația dintre variabile. Cu cât numărul de variabile e mai mare complexitatea tabelului crește.

Forma generală a tabelului de contingență este  $r \times c$  unde  $r$  este numărul de coloane și  $c$  este numărul de rânduri. Pe coloane, respectiv rânduri se va sumariza una sau mai multe variabile. Numărul de coloane și respectiv rânduri este dat de tipul variabilei sumariate. Dacă variabila este calitativă ordinală sau nominală, numărul de coloane/rânduri este egal cu numărul de valori posibile pe care variabila de interes le ia în eșantion. În cazul în care variabilele de interes sunt dicotomiale (pot lua doar două valori: Da/Nu, Sănătos/Bolnav, Prezent/Absent, etc.) tabelul de contingență devine de  $2 \times 2$ , tabel frecvent utilizat în studiile medicale.

## EXEMPLU

Fie seria statistică formată din gen (F/M) și valoarea scorului de durere a 100 pacienți cu artroză:

F	9	F	9	M	4	M	9	F	7	F	10	M	9	F	10	M	9	M	6
M	9	M	6	F	4	M	5	F	7	F	5	F	9	F	9	M	8	F	9
M	7	M	9	M	5	M	8	F	6	F	9	F	8	M	7	F	9	F	8
M	10	M	8	M	9	M	8	M	6	F	10	F	8	F	8	F	9	F	8
M	10	M	7	M	5	F	10	F	9	F	8	F	8	F	7	F	6	F	9
F	7	F	8	M	8	M	8	F	10	F	8	M	8	M	7	F	8	F	7
F	10	M	9	M	8	M	7	M	7	F	5	F	2	F	9	M	6	M	8
M	8	M	10	M	9	M	10	M	5	F	7	F	4	F	10	F	10	M	6
M	7	F	9	M	10	M	8	M	6	F	9	F	10	F	10	M	7	M	7
M	7	F	8	M	9	M	7	F	6	F	9	F	10	M	8	F	8	F	9



## SOLUȚIA

Gen Scor durere	F	M	Total
2	1	0	1
4	2	1	3
5	2	4	6
6	3	6	9
7	6	11	18
8	12	12	24
9	15	9	24
10	11	5	16
Total	52	48	100

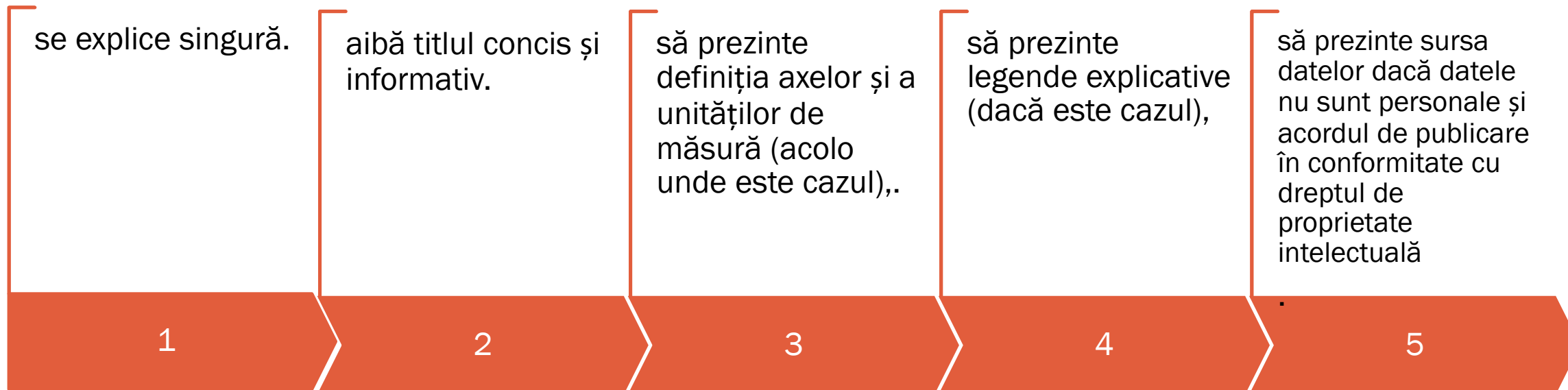
Gen	Fumat	Obezitate = prezentă	Obezitate = absentă	Total rânduri
F	da	3	28	31
	nu	7	172	179
Total F		<b>10</b>	<b>200</b>	<b>210</b>
M	da	33	39	72
	nu	61	97	158
Total M		<b>94</b>	<b>136</b>	<b>230</b>
Total coloane		<b>104</b>	<b>336</b>	<b>440</b>

**EXEMPLU DE TABEL DE CONTINGENȚĂ CU MAI MULT DE 2 VARIBILE**

# REPREZENTAREA GRAFICĂ A DATELOR MEDICALE

Reprezentarea grafică a datelor medicale este metoda care permite prezentarea rapidă și cu un bun impact vizual a rezultatelor.

Ca și la reprezentarea tabelară a datelor rezultate din cercetarea medicală, reprezentarea grafică trebuie să:



## EXEMPLU

Au fost investigate caracteristicile antropometrice ale nou născuților cu restricție de creștere intrauterină prin evaluarea:

- genului (F/M),
- greutății (g),
- lungimii (cm),
- a perimetrului cranian (cm) și a
- traumatismului la naștere.

Studiul a fost de tip caz-martor cu potrivire în ceea ce privește genul și vârsta de gestație.

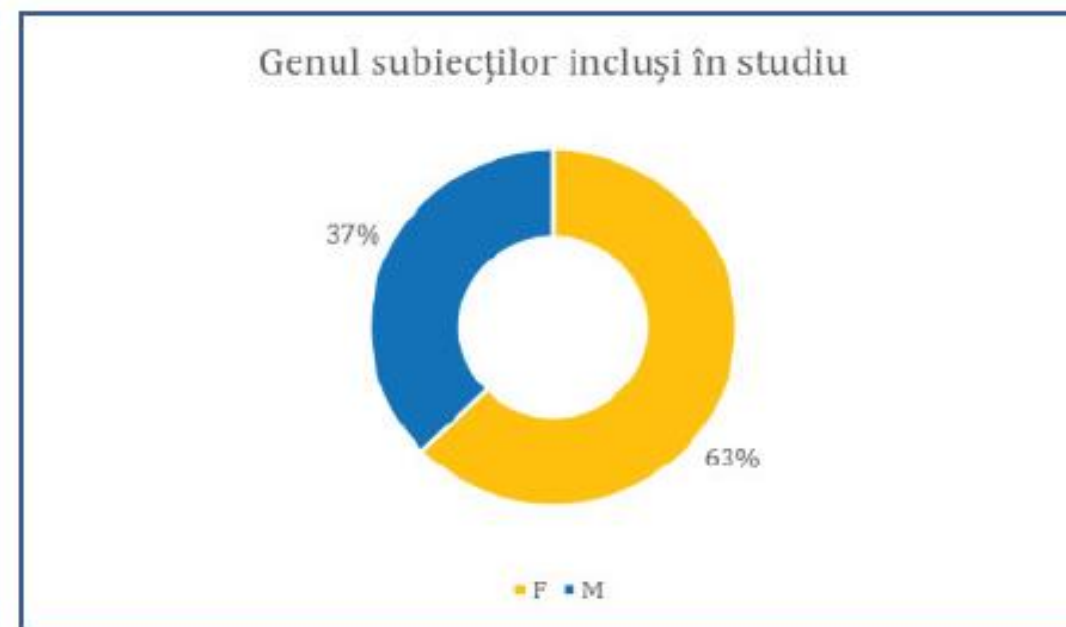
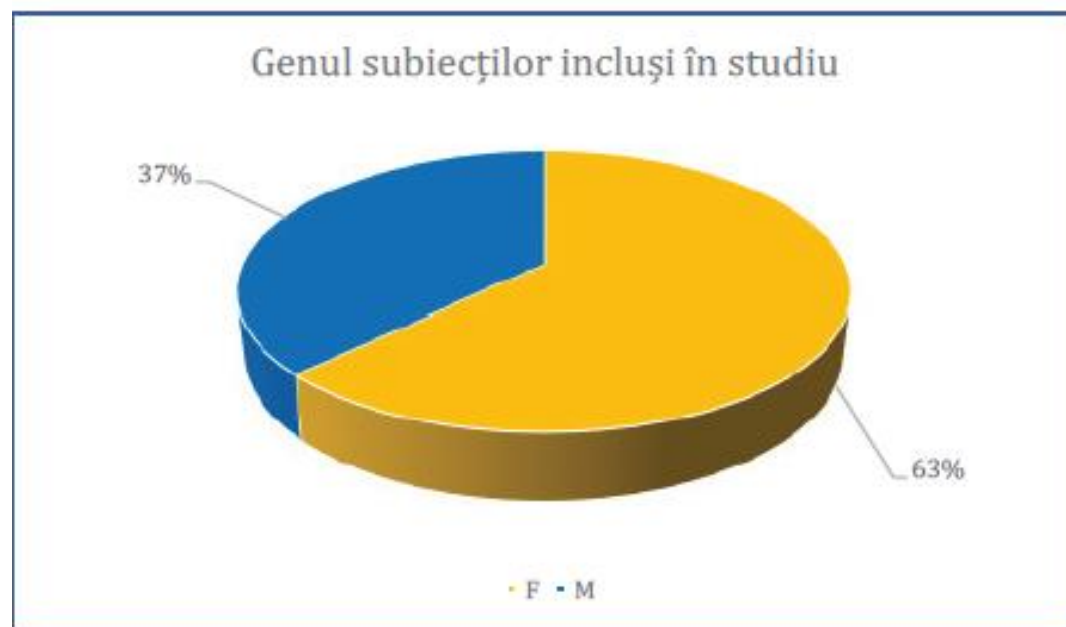
- Care sunt reprezentările grafice potrivite pentru ilustrarea datelor din acest exemplu?



## GRAFICUL DE TIP SECTORIAL /DE STRUCTURĂ

Reprezentarea grafică a distribuției genului în eșantionul studiat:

- stânga: grafic de tip circular (*Pie*);
- dreapta: grafic de tip 'gogoasă' (*Doughnut*)



## EXEMPLU

A fost evaluată modalitatea de tratament primit în urgență pe un eșantion de subiecți cu patologii cardiace. Datele sunt sumarizate în Tabelul 2.5.

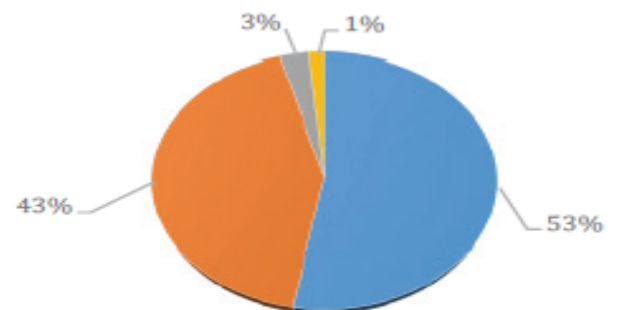
**Tabelul 2.5. Modalitatea de aplicare a tratamentului (iv = intravenos, po = per oral) în funcție de diagnostic și gen**

Diagnostic	Tratament iv			Tratament po			Total
	F	M	Total iv	F	M	Total po	
Cardiomiopatie dilatativă	1	1	2	0	0	0	2
Cardiopatie ischemică	7	5	12	12	13	25	37
Hipertensiune arterială	5	3	8	10	12	22	30
Infarct miocardic	1	0	1	0	0	0	1
Total	14	9	23	22	25	47	10

Reprezentați grafic distribuția diagnosticului în serviciul de urgență pe eșantionul studiat.

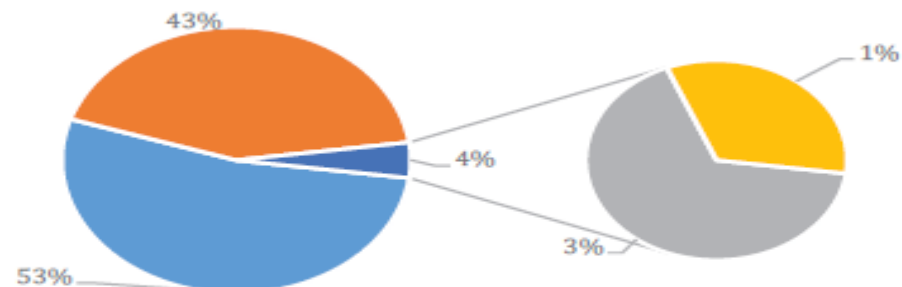
Reprezentarea grafică a distribuției diagnosticului: a) (stânga sus) grafic sectorial; b) (dreapta-sus) grafic de tip Pie to Pie; c) (jos) grafic de tip Pie to Bar

Distribuția diagnosticului la serviciul de urgență



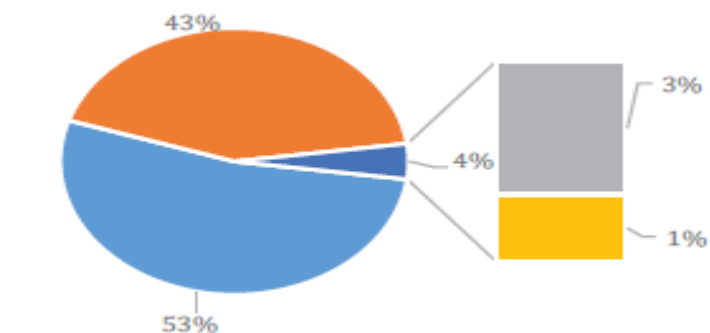
- Cardiopatie ischemică
- Hipertensiune arterială
- Cardiomiopatie dilatativă
- Infarct miocardic

Distribuția diagnosticului la serviciul de urgență



- Cardiopatie ischemică
- Hipertensiune arterială
- Cardiomiopatie dilatativă

Distribuția diagnosticului la serviciul de urgență



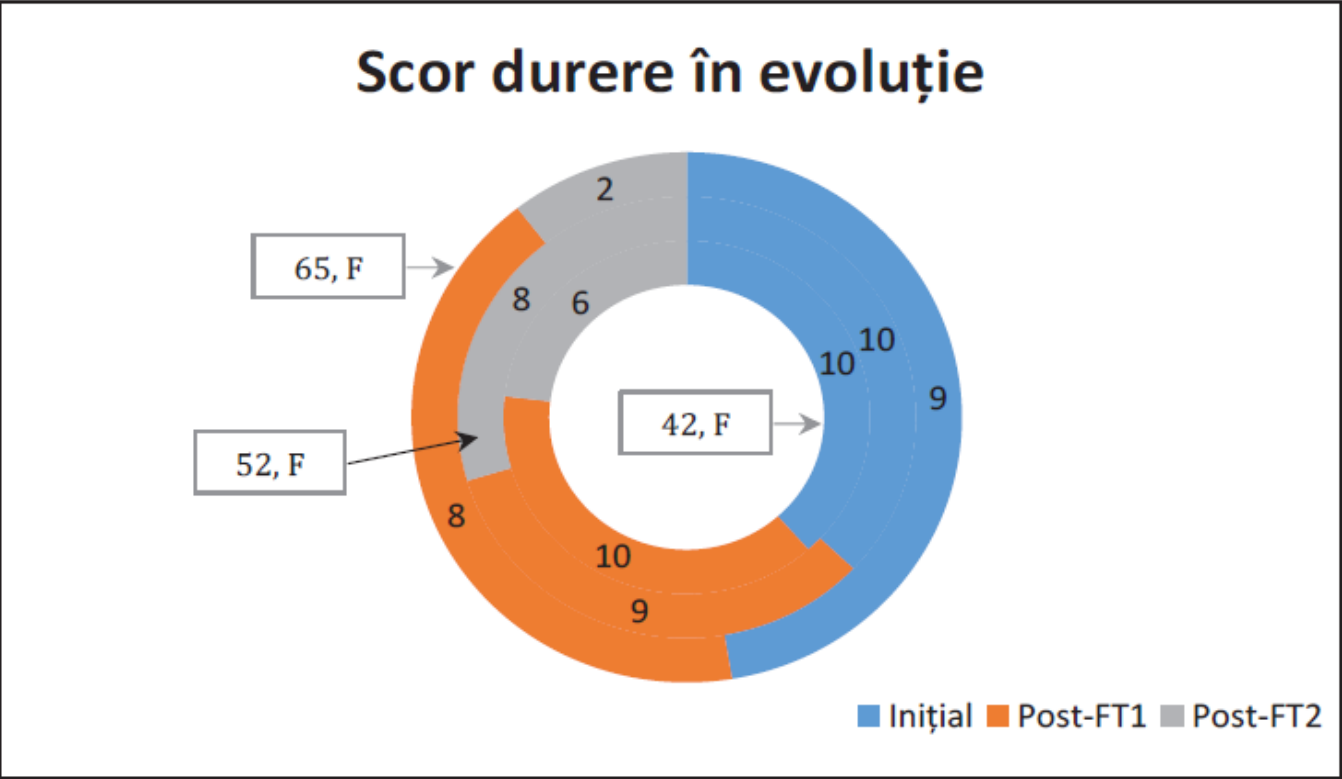
- Cardiopatie ischemică
- Hipertensiune arterială
- Cardiomiopatie dilatativă
- Infarct miocardic

# GRAFICE DE TIP DOUGHNUT

Scorul de durere la trei pacienți cu artroză care s-au prezenta pentru efectuarea tratamentului fizioterapic au fost înregistrate la prezentare, după prima cură de fizioterapie (post-FT1) și respectiv după a doua cură de fizioterapie (post-FT2). Datele sunt sumarizate în Tabelul 2.6.

**Tabelul 2.6. Efectul fizioterapiei asupra durerii la pacienții cu artroză. Scorul de durere măsurat pe scala vizuală (0 = fără durere, 10 = durere extrem de puternică)**

Vârsta (ani)	Genul	Scor durere		
		Inițial	Post-FT1	Post-FT2
42	F	10	10	6
52	F	10	9	8
65	F	9	8	2



## GRAFICUL DE TIP COLOANE/BARE

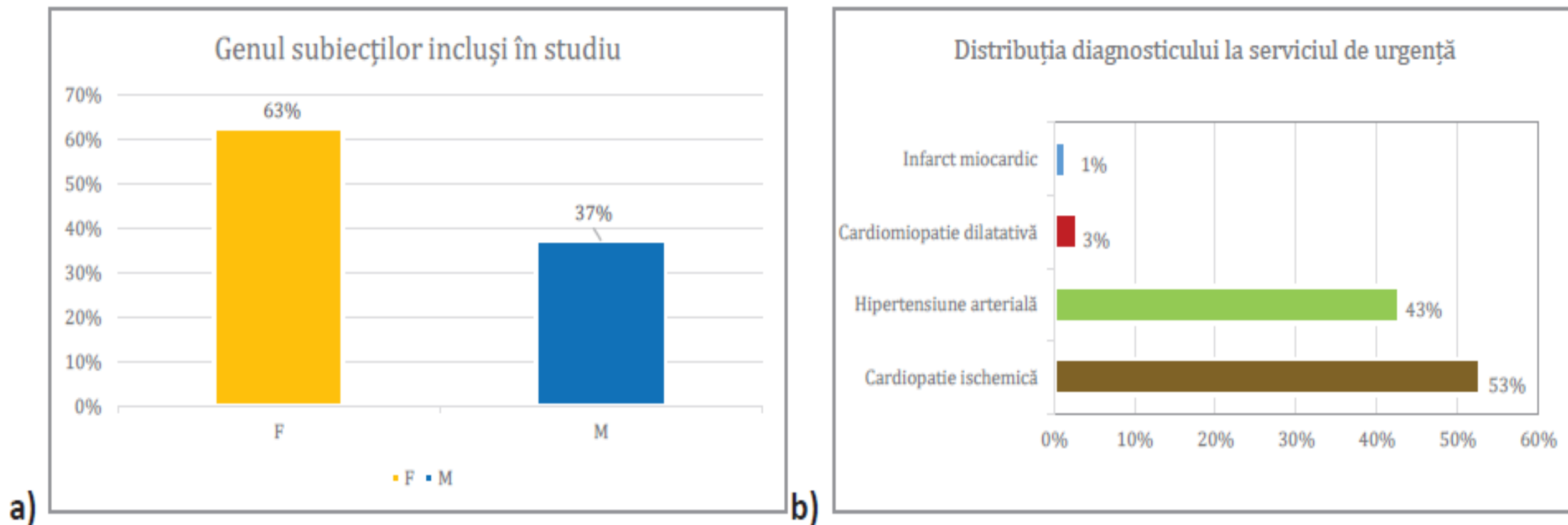


Fig. 2.10. Reprezentarea grafică de tip: a) coloane (datele din Exemplul 2.19.1, o singură variabilă calitativă dicotomială); b) bare (datele din Exemplul 2.20, o singură variabilă calitativă nominală)



## GRAFICUL DE TIP COLOANE/BARE

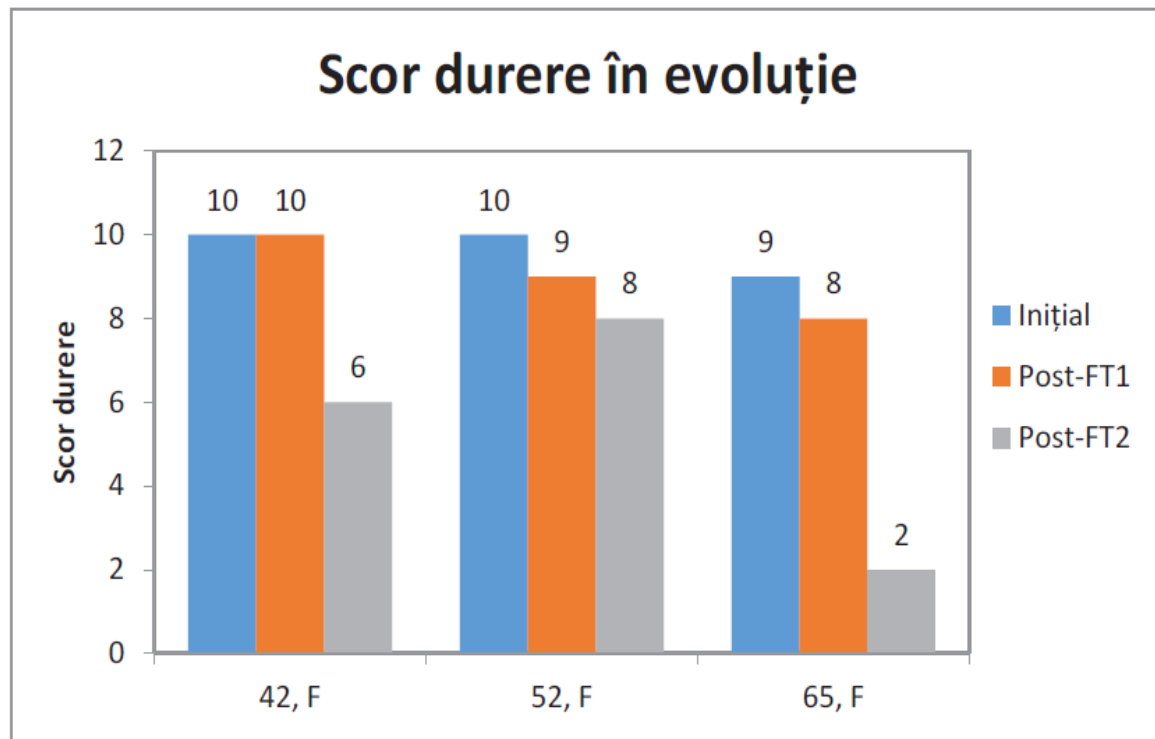


Fig. 2.11. Evoluția scorului de durere pentru datele din exemplul 2.22

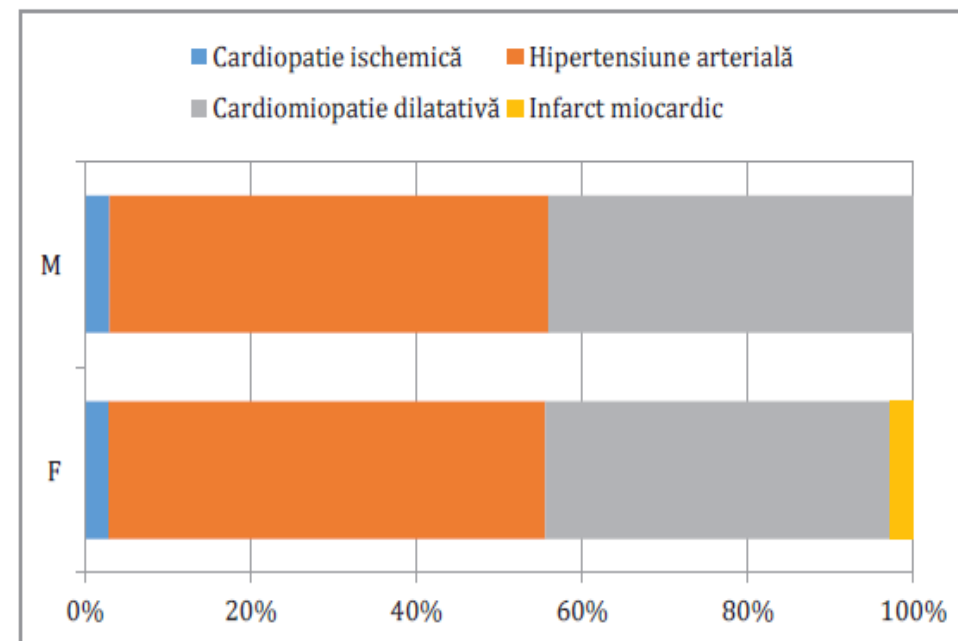
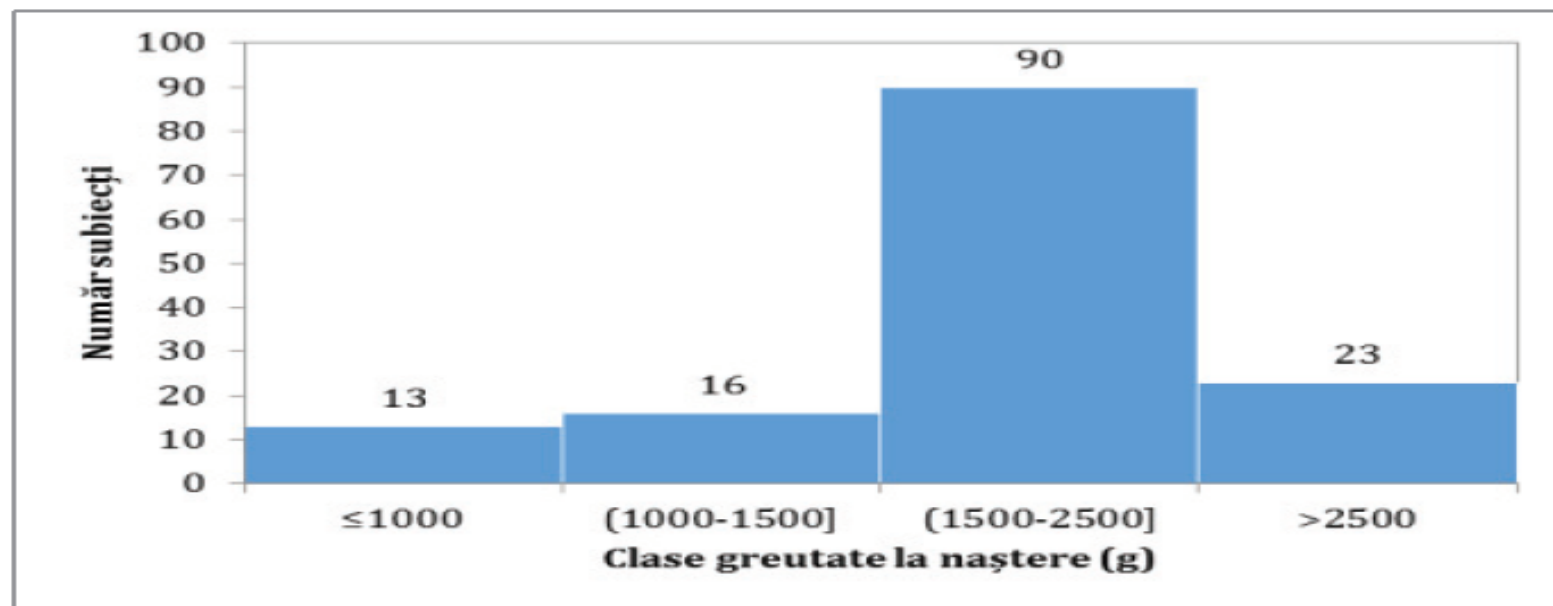


Fig. 2.12. Diagnosticul în funcție de gen pentru datele din Exemplul 2.21

# HISTOGRAMA

Reprezentați grafic distribuția greutății la naștere a subiecților cu restricție de creștere intrauterină din exemplul 2.20.

Reprezentarea grafică se va face pentru următoarele clase de frecvență exprimate în grame:  $\leq 1000$ ;  $(1000-1500]$ ;  $(1500-2500]$  și  $>2500$  (figura 2.12). Putem opta să reprezentăm grafic frecvențele absolute (numărul de cazuri) sau relative (%) pentru fiecare clasă în parte.



**Fig. 2.13. Histograma greutății la naștere a copiilor cu restricție de creștere intrauterină din exemplul 2.20**

## GRAFICE DE TIP POLIGON DE FRECVENȚĂ

Reprezentați grafic distribuția greutății la naștere a subiecților cu (grup caz) și respectiv fără (grup martor) restricție de creștere intrauterină din exemplul 2.20.

Reprezentarea grafică se va face utilizând următoarele clase de frecvență: 1000, (1000-1500], (1500-2000], (2000-2500], (2500-3000], (3000-3500], >3500 (vezi figura 2.14).

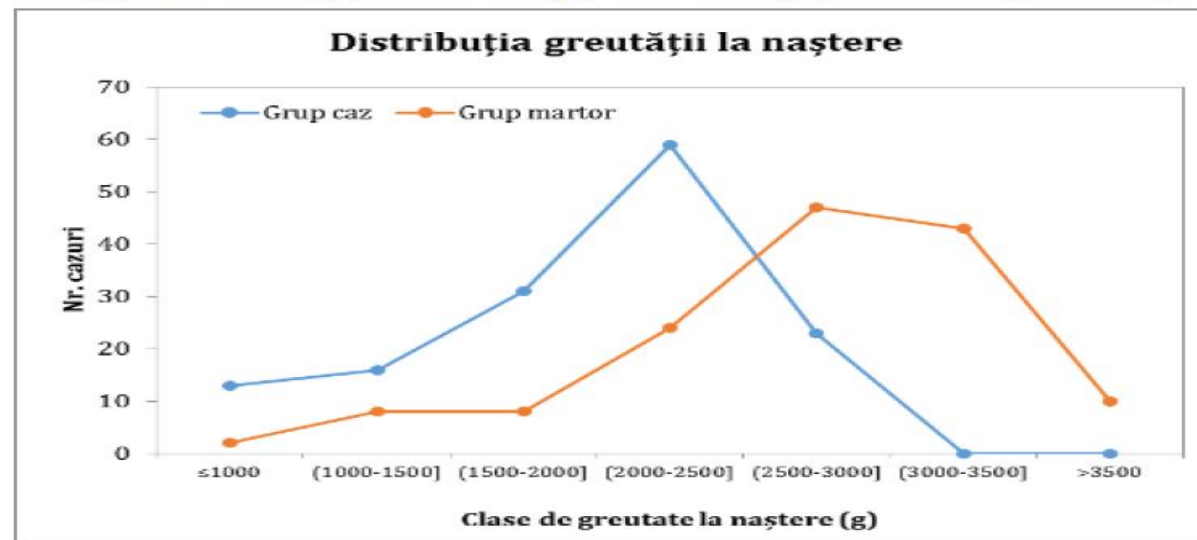


Fig. 2.14. Poligoane de frecvență pentru variabila greutate la naștere pentru nou-născuții cu (Grup caz) și fără (Grup martor) restricție de creștere intrauterină din exemplul 2.20

Reprezentarea grafică a distribuției unei serii statistice cantitative se poate face și cu ajutorul poligonului de frecvență. Spre deosebire de histogramă, poligonul de frecvență permite reprezentarea simultană a două grupuri folosind aceleași clase de frecvență. Această modalitate de reprezentare grafică prezintă un interes deosebit deoarece: forma sa dă o imagine globală a distribuției observate (care poate fi "comparată" cu o distribuție teoretică) și permite vizualizarea observațiilor "aberrante" (apărute prin erori sistematice de codificare și de măsurare sau introduse de subiecți cu anumite particularități) chiar dacă acestea sunt atenuate prin gruparea în clase.

# GRAFICE DE TIP CUTIE CU MUSTĂȚI (BOXPLOT)

Reprezentați grafic utilizând diagrama de tip *boxplot* greutatea la naștere a subiecților cu și respectiv fără restricție de creștere intrauterină din Exemplul 2.20.

Graficele de tip cutie cu mustăți, *boxplot* în limba engleză, asociază categoriilor reprezentate segmente de dreaptă și dreptunghiuri cu diferite semnificații aparținând parametrilor statistici descriptivi. În cazul acestui tip de reprezentare grafică, putem avea o variabilă cantitativă pentru un eșantion (rar utilizat) sau o variabilă cantitativă pentru mai mult de un grup, caz în care avem variabila cantitativă de interes și o variabilă de grupare (calitativă). Sunt frecvent utilizate pentru a arăta diferențele între grupuri în ceea ce privește o variabilă cantitativă.

Obținerea acestui tip de reprezentare grafică este foarte ușoară cu ajutorul programelor dedicate de statistică dar există și implementări în ultimele versiuni ale Microsoft Excel (vezi figura 2.15).

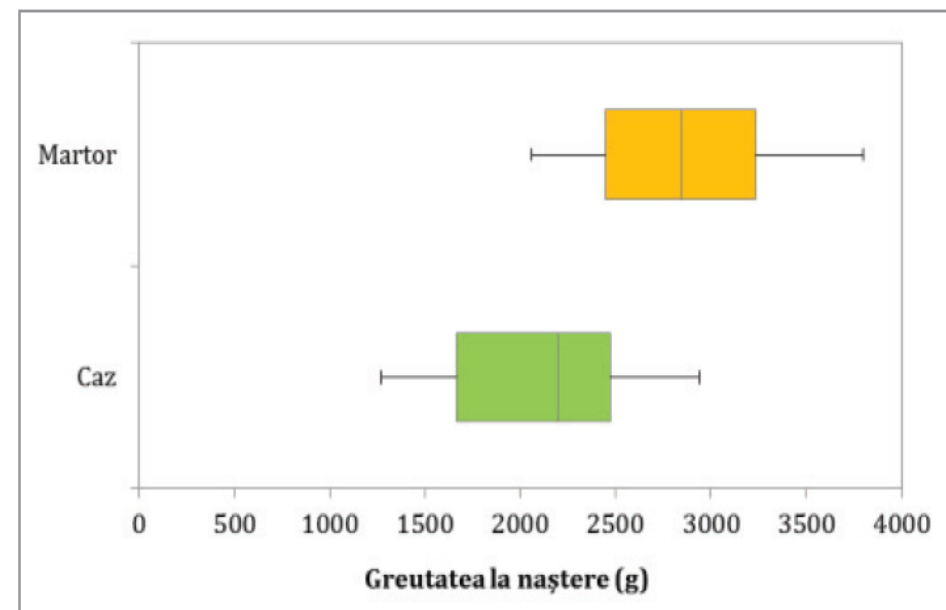
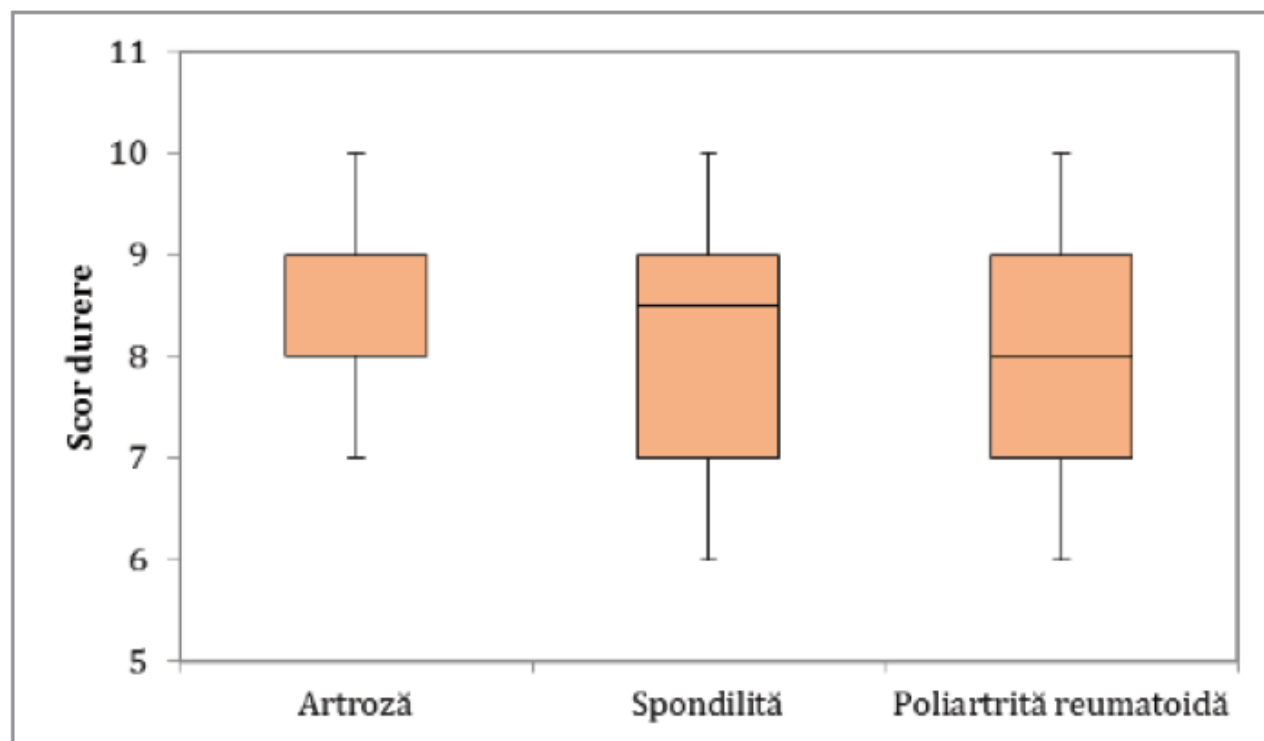


Fig. 2.15. Grafic de tip boxplot pentru variabila greutate la naștere pentru nou-născuții cu (grup caz) și fără (grup martor) restricție de creștere intrauterină din exemplul 2.20. Valoarea din mijlocul dreptunghiului este mediana, dreptunghiul este dat de valorile cvartilei 1 și respectiv 3 iar barele extreme sunt date de valoarea minimă și respectiv maximă

## GRAFICE DE TIP CUTIE CU MUSTĂȚI (BOXPLOT)



Pentru utilizarea acestui tip de grafic se utilizează media dacă datele cantitative urmează distribuția normală. În caz contrar, valoarea minimă, valoarea primei cvartile, mediana, valoarea celei de a 3-a cvartile și valoarea maximă se utilizează pentru construirea graficului de tip *boxplot*. Orientarea graficului poate să fie orizontală (figura 2.15) sau verticală (figura 2.16).

**Fig. 2.16.** Grafic de tip boxplot pentru scorul durerii (scala de la 0 = fără durere la 10 = durere extrem de puternică) la pacienții cu artroză, spondilită anchilozantă și respectiv cu poliartrită reumatoidă. Valoarea din mijlocul dreptunghiului este mediana și lipsește la pacienții cu artroză deoarece valoarea medianei este egală cu valoarea cvartilei 1



## GRAFIC DE TIP NOR DE PUNCTE

Pentru prezentarea grafică a relațiilor sau asocierilor dintre două variabile cantitative perechi (de exemplu, lungime la naștere – perimetru cranian la naștere) se utilizează diagrama de tip nor de puncte (*scatter*).

Fiecare punct din reprezentarea grafică este caracterizat prin două valori, o valoare pe axa OX și cea de-a doua valoare pe axa OY. Dacă există o relație de dependență cunoscută între cele două variabile atunci pe axa OX (lungimea în cazul datelor din figura 2.17) se vor pune valorile variabilei independente iar pe axa OY (perimetrul cranian în cazul datelor din figura 2.17) se vor pune valorile variabilei dependente.

Pe graficul de tip nor de puncte se poate prezenta unul (figura 2.17a) sau mai multe grupuri (figura 2.17b) pentru care au fost studiate aceleași variabile perechi.

Pe graficul de tip nor de puncte pot exista informații cu privire la dreapta de regresie și a coeficientului de determinare ( $r^2$ ) noțiuni care vor fi introduse în capitolul *Corelații și Regresii*.

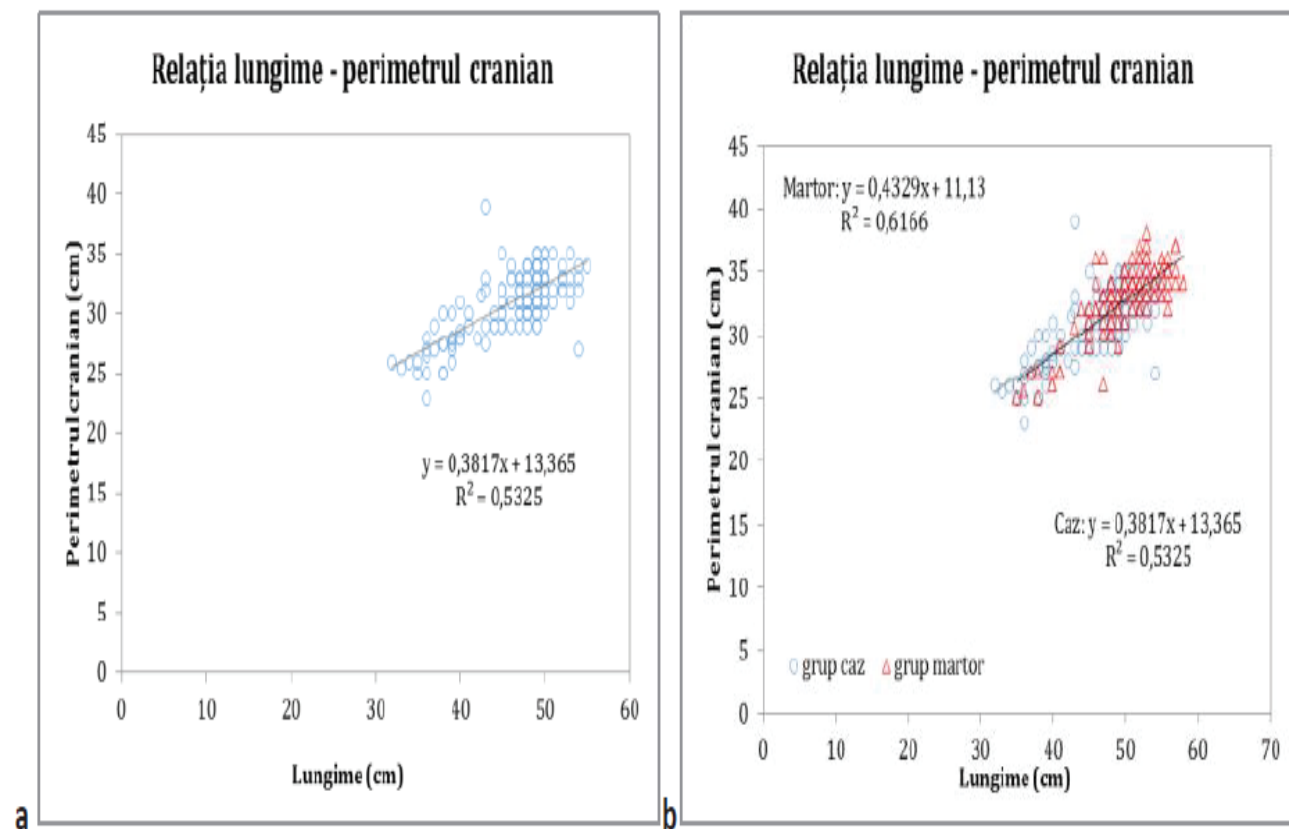


Fig. 2.17. Relația dintre lungimea și perimetrul cranian la naștere pentru datele studiului prezenta în exemplul 2.20: a) grupul cu restricție de creștere intrauterină (caz); b) grupul cu (caz) și fără (martor) restricție de creștere intrauterină.

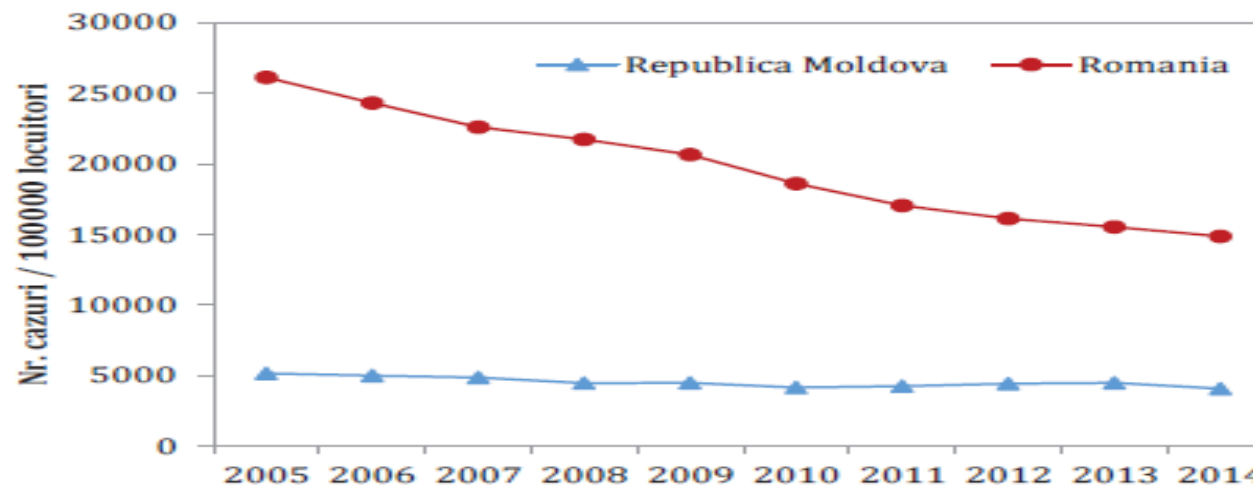
## GRAFIC DE TIP LINIE (CRONOGRAME)

S-a investigat numărul de cazuri de tuberculoză (TBC) din Republica Moldova și România în perioada 2004-2014. Sumarizarea datelor colectate este prezentată în Tabelul 2.7.

**Tabelul 2.7. Numărul de cazuri de tuberculoză raportate la 100000 locuitori**

An Țara	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Republica Moldova	5141	4990	4857	4464	4471	4135	4233	4409	4485	4058
România	26106	24295	22590	21724	20643	18590	17045	16107	15523	14861

Reprezentarea grafică a evoluției cazurilor de TBC în timp pentru cele două țări este redată în figura 2.18.



**Fig. 2.18. Evoluția pe parcursul a 10 ani a cazurilor de TBC în România și Republica Moldova**

# GRAFIC DE TIP ARIE

Graficul de tip arie permite vizualizarea sumarizărilor cantitative și arată importanța relativă a valorilor fiind utilizat pentru a compara doua sau mai multe cantități.

S-a evaluat scorul de durere al unui eșantion de 7 pacienți de gen feminin cu modificări artrozice la includerea în studiu, după prima cură de fizioterapie, la 6 luni de la prima cură de fizioterapie și respectiv la după cea de-a doua cură de fizioterapie. Scala vizuală a fost utilizată pentru a cuantifica durerea (scor 0 = fără durere, scor 10 = durere extrem de puternică) iar datele brute sunt prezentate în tabelul 2.8.

Tabelul 2.8. Scorul durerii la pacienți cu artroză

Vârsta (ani)	Scor durere			
	Inițial	Post fizioterapie 1	6 luni	Post fizioterapie 2
65	9	8	2	2
73	7	6	1	2
72	8	7	2	2
50	7	6	5	4
52	10	9	9	8
41	8	8	2	2
76	9	8	3	2

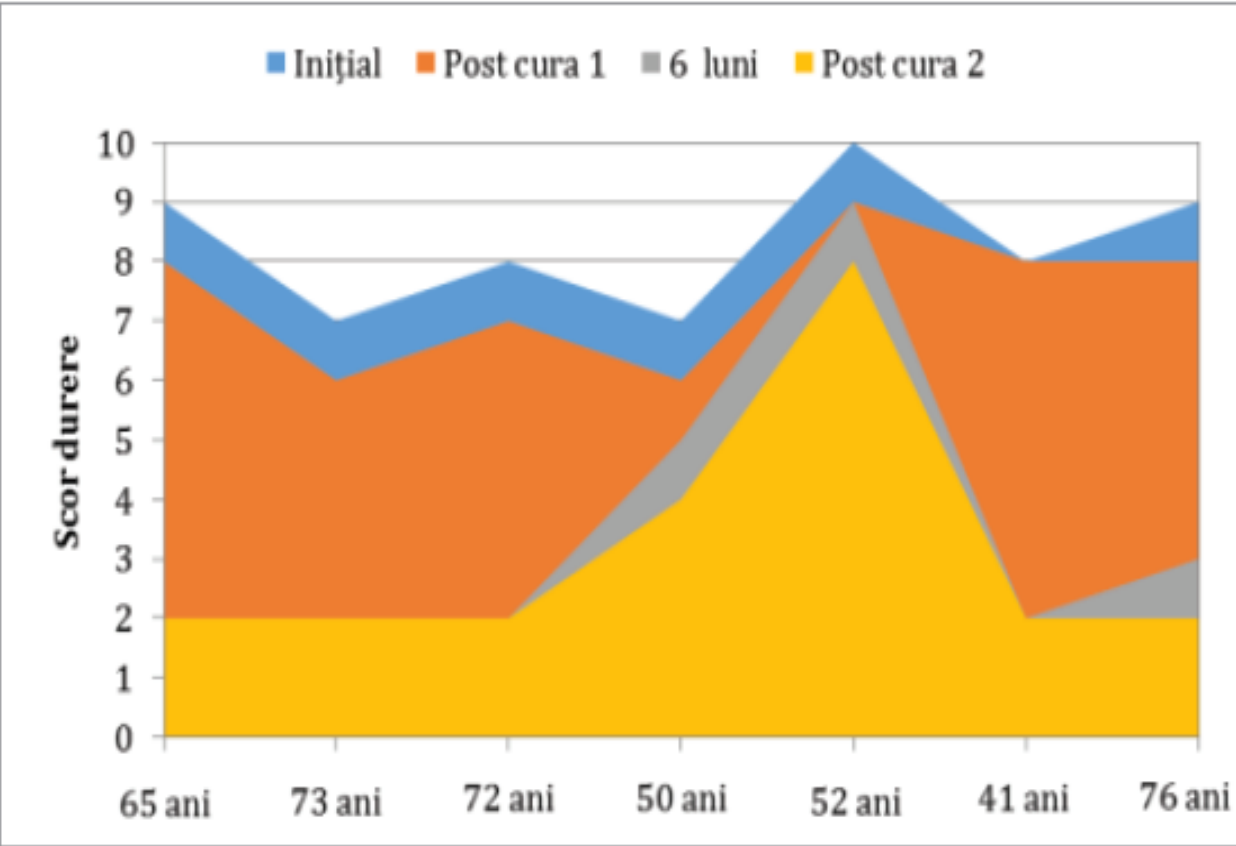


Fig. 2.19. Evoluția durerii în contextul curei de fizioterapie

## GRAFIC DE TIP BALOANE

Graficul de tip baloane permite reprezentarea grafică simultană a 3 dimensiuni. Diametrul balonului indică valoarea celei de a treia dimensiuni în timp ce prima și cea de-a doua dimensiune sunt reprezentate pe axa OX și respectiv OY de poziția centrului acestuia.

S-a realizat un studiu pentru a evalua costul tratamentului cu antibiotice pentru tusea convulsivă pe o durată de 5 ani (2008-2012). Datele sunt prezentate în Tabelul 2.9.

**Tabelul 2.9. Costuri ale antibioterapiei în tusea convulsivă: 2008-2012**

Anul	Nr. cazuri tuse convulsivă	Costul mediu al antibioterapiei
2012	82	2952
2011	86	3096
2010	29	1044
2009	10	360
2008	51	1836

Reprezentarea grafică asociată datelor din Tabelul 2.9 este redată în figura 2.20. Graficul nu este foarte intuitiv deoarece aria fiecărei sfere este proporțională cu costul mediu, deci variază cu radicalul valorii.

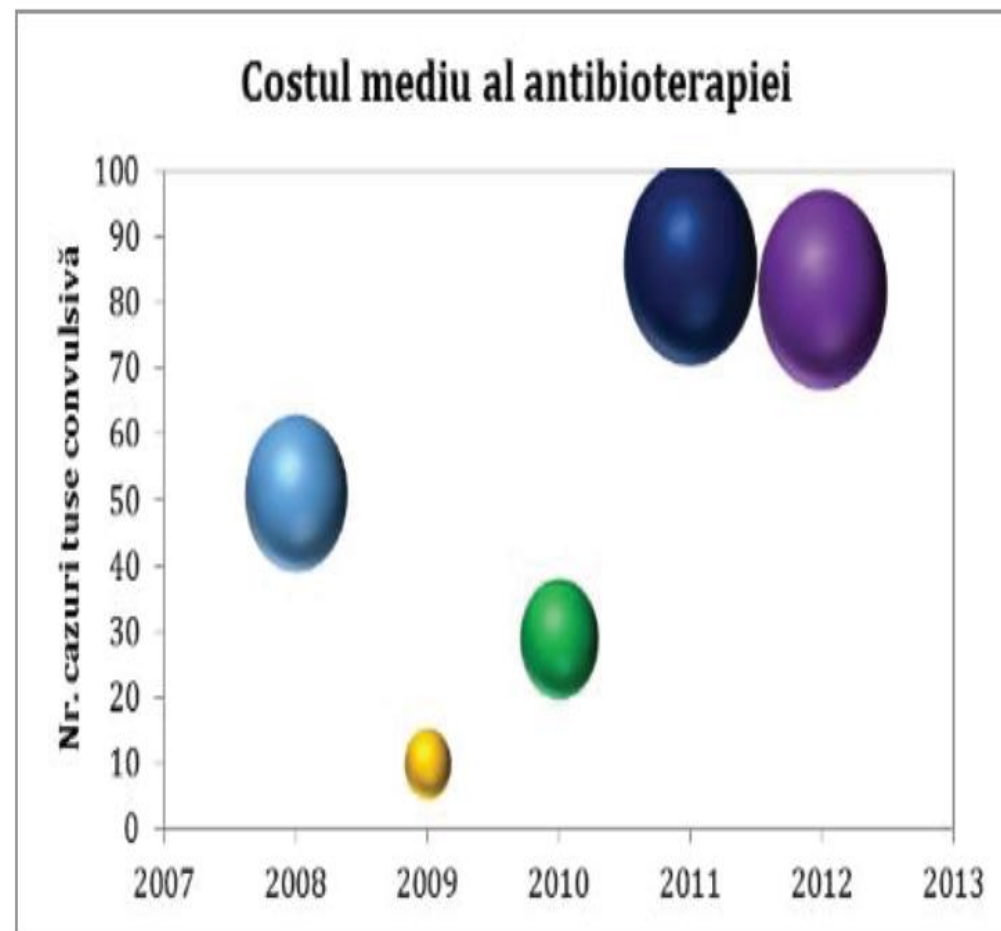


Fig. 2.19. Costul mediu al antibioterapiei per ani pentru tusea convulsivă

## NOTIUNI DE RETINUT!

Proporția, rata, procentul sunt utilizate pentru sumarizarea (rezumarea, sintetizarea) datelor cantitative (nominale sau ordinale).



Modalitatea de sumarizare a datelor este dictată de tipul variabilei și scala de măsură. Același principiu se aplică și în cazul calculării parametrilor statistici descriptivi



Media aritmetică este un parametru de centralitate caracteristic datelor cantitative

Mediana este parametrul de centralitate care caracterizează datele ordinale sau cele cantitative preferată pentru acele date care nu urmează distribuția normală



Deviația standard este un parametru de dispersie care caracterizează modalitatea în care datele seriei statistice se distribuie în jurul mediei.



Coeficientul de variație este o mărime de dispersie relativă care permite compararea a două sau mai multe variabile cu sau fără diferite unități de măsură.



## NOTIUNI DE RETINUT!

Tabelele de frecvență arată numărul de observații sau % observațiilor care au o anumită caracteristică.



Histograma, poligonul de frecvență și boxplot-ul permit reprezentarea și vizualizarea distribuției datelor cantitative.



Graficul de tip nor de puncte prezintă relația dintre două caracteristici cantitative

Rata descrie numărul de evenimente care apar într-o perioadă definită de timp, de spațiu și dintr-o populație definită



Relația dintre două caracteristici calitative se descrie cu ajutorul indicatorilor de tip raport sau rație (ex. riscul relativ, rația șansei, etc.).