Lp 2 Rezumat

3.2. Sinteza numerică univariată

Oferă măsuri obiective şi exacte ale unor aspecte esenţiale de *variabilitate* (ex. *omogenitate* versus *eterogenitate*) şi de *tendinţă centrală* sau alte tendinţe.
Gândind CANTITATIV, *variabilitatea* se manifestă ca *împrăştiere*, eventual *în jurul unei tendinţe centrale*, sau *între tendinţe extreme* sau *intermediare*.
Gândind CALITATIV, *variabilitatea* se manifestă ca *diversitate*.

3.3. Indicatori de tendinţă centrală

3.3.1. Condiţiile lui Yulle asupra unui indicator de tendinţă centrală

Cea mai importantă condiţie este: "Să se preteze uşor la calculele algebrice ulterioare."

3.3.2. Modă

Concept geometric (de gândire în continuu) aplicabil şi în cazul discontinuu (discret) al distribuţiilor de frecvenţe empirice. Important conceptual doar, pentru clasificarea distribuţiilor în unimodale (singurele care au o «tendinţă centrală»), bimodale, multimodale.
Modă = valoare cu frecvenţă maximă locală în distribuţia de frecvenţe.
Nu se pretează la calcule algebrice, ceea ce descalifică moda în teoria şi practica statistică.

3.3.3. Mediana

Concept, de asemenea geometric, aplicabil şi în cazul discret al distribuţiilor de frecvenţe empirice (seriilor statistice), cu preţul unei convenţii suplimentare pentru obţinerea unicităţii rezultatului, în cazul seriilor cu volum par.
Mediană = o valoare care împarte seria statistică **ordonată** în două subserii cu volume egale, 50%; 50%, (volumele fiind măsurate în număr de unităţi statistice şi eventual jumătăţi - fracţiuni - ale acestora).
Notăm cu $(x_i)_{i=1,2,\ldots,N}$ seria **ordonată**.

- Dacă are volum impar, *mediana* - conform definiţiei - e unică, fiind termenul din mijlocul seriei **ordonate**. Acesta are **rangul** (rg) = $N/2$ rotunjit prin adaus (adică $rg = [N/2+1]$ *). Vom scrie $Me = x_{rg}$.
- Dacă are volum par, operează convenţia: *mediana* = semisuma termenilor din mijlocul seriei **ordonate.** (Vom scrie $Me = (x_{rg} + x_{rg+1}) / 2$, unde $rg = N/2$).

Deci *mediana* este:

- termenul din mijlocul seriei **ordonate**, dacă seria are volum impar ($rg = N/2$ NU este întreg);
- semisuma termenilor din mijlocul seriei **ordonate**, dacă seria are volum par ($rg = N/2$ este întreg).

Calcul rapid:
Calculăm $rg = N/2$. Mediana Me este:

- termenul x_{rg} cu rg rotunjit prin adaus, dacă rg NU este întreg;
- semisuma termenilor x_{rg} şi x_{rg+1}, dacă rg este un întreg.

Mediana tratează valorile ca pe ranguri, fiind calculabilă şi atunci când nu se cunosc valorile extreme. Nu se pretează însă la calcule algebrice.
Alte denumiri:
LD 50 = "**L**ethal **D**ose **50**" - în toxicologie, **ED 50** = "**E**ffect **D**ose **50**" - în farmacologie, **Media de viaţă** - în biologia populaţiilor.

3.3.4. Media (aritmetică)

Concept algebric (discret), definit în cazul unei serii statistice prin:
Media (aritmetică) = suma valorilor seriei împărţită la volumul seriei.

În cazul *seriei statistice* a	În cazul grupării seriei într-o distribuţie de frecvenţe	
	absolute	relative
celor N valori distincte sau nu, x_i:	a celor p $(\le N=\sum_{j=1}^{p} N_j)$ valori distincte x_j:	
$\boxed{M=\frac{\sum_{i=1}^{N} x_i}{N}}$	$M=\frac{\sum_{j=1}^{p} N_j \cdot x_j}{N}$	$M=\sum_{j=1}^{p} F_j \cdot x_j$
Formula mediei simple	**Formula mediei ponderate**	**Sumă ponderată**

* $[x]$ înseamnă partea întreagă din x.

Are extrem de importanta proprietate algebrică de *aditivitate*:
"Media unei serii este egală cu media ponderată a mediilor subseriilor compozite, ponderile fiind volumele subseriilor."

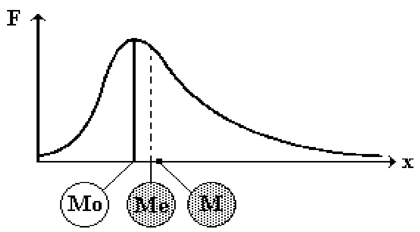
O eroare gravă - din păcate des întâlnită - este interpretarea mediei la distribuții care nu sunt unimodale, adică atunci când nu are sens. Moisil spunea "Dacă stau cu o bucă pe un cub de gheață şi cu cealaltă pe o plită încinsă, în medie, mă simt bine". Vezi coperta principală.
(Pentru detalii şi evitarea erorii vezi http://app.inthlerom.ro/histo/HistoSetup.zip)

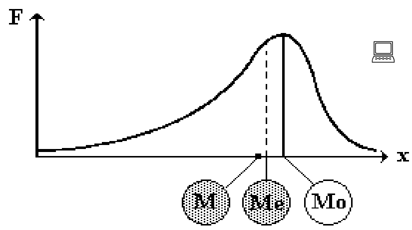
3.3.5. Indicații de preferință între principalii indicatori de tendință centrală [6]

Moda unică - omogenitate. Se foloseşte la serii cu volume mari şi/sau când vrem să ignorăm valorile extreme.
Mediana se foloseşte la serii cu volume mici şi/sau când vrem să ignorăm valorile extreme şi/sau când cunoaştem doar rangurile valorilor. Exprimă cel mai bine tendința centrală la distribuțiile asimetrice.
Media nu are sens decât la distribuții unimodale. Se calculează la distribuții cvasisimetrice, care cer prelucrări ulterioare şi când vrem să considerăm toate valorile seriei, împreună cu întreaga lor informație.

+ Observații [3]:

(1) Distribuție unimodală şi simetrică ⇒ $Mo = Me = M$.
(2) asimetrie de stânga ⇔ $Mo<Me<M$; (3) asimetrie de dreapta ⇔ $M<Me<Mo$;





3.4. Alți indicatori de localizare

- a tendințelor extreme: *valoarea minimă* (x_{min}) şi *valoarea maximă* (x_{max}),
- a tendințelor intermediare ori extreme:

3.4.1. Cuartile

Concept geometric care extinde conceptul de mediană. Ca şi mediana este aplicabil şi în cazul discret al distribuţiilor de frecvenţe empirice (seriilor statistice), de asemenea, cu preţul unei convenţii suplimentare pentru obţinerea unicităţii rezultatului, după cum urmează:

Cuartile = 3 valori - notate Q_k, (k =1,2,3) - care împart seria statistică **ordonată crescător,** $(x_i)_{i=1,2,...,N}$, în **4** subserii de volume egale (volumele fiind măsurate în număr de unităţi statistice şi eventual pătrimi – fracţiuni - ale acestora). Q_1 sn **cuartilă inferioară**, (Q_2 este mediana seriei), Q_3 sn **cuartilă superioară**.

Dacă notăm $rg = N\cdot (k / 4)$, atunci cuartila Q_k, (k =1, 2, 3) este, în **seria ordonată crescător**:

- termenul x_{rg}, cu rg rotunjit prin adaus, dacă rg NU este un număr întreg;
- semisuma termenilor x_{rg} şi x_{rg+1}, dacă rg este un număr întreg.

Se observă că această reformulare este identică cu cea pentru mediană cu excepţia valorii *rg.*

3.4.2. Decile şi (per)centile

În mod analog, se construiesc conceptele de decile şi centile.

Decile = 9 valori care împart seria statistică **ordonată crescător** în **10** subserii de volume egale (volumele fiind măsurate în număr de unităţi statistice şi eventual zecimi – fracţiuni - ale acestora). Se notează D_1, D_2, ..., D_9. *D_1 - decilă inferioară, D_9 - decilă superioară.*

Decila D_k (k =1, 2, ..., 9) se poate calcula după aceeaşi formulă cu o mediană sau cuartilă, dar cu $rg = N\cdot (k / 10)$.

(Per)centile = 99 de valori care împart seria statistică **ordonată crescător** în **100** subserii de volume egale (volumele fiind măsurate în număr de unităţi statistice şi eventual sutimi – fracţiuni - ale acestora). Se notează C_1, C_2, ... , C_{99}. *C_1 - centilă inferioară* şi *C_{99} - centilă superioară.*

Centila C_k, (k =1,2,..,99) se poate calcula după aceeaşi formulă cu o mediană, cuartilă sau decilă dar cu $rg = N\cdot (k / 100)$. Adică este, în seria statistică **ordonată crescător**:

- termenul x_{rg}, cu rg rotunjit prin adaus, dacă rg NU este un număr întreg;
- semisuma termenilor x_{rg} şi x_{rg+1}, dacă rg este un număr întreg.

Scală de clasificare bazată pe centile [10]

Dimensiune:	foarte mică		mică		medie		mare		foarte mare
Scală centilică	$\rightarrow$	c_2	$\rightarrow$	c_{25}	$\rightarrow$	c_{75}	$\rightarrow$	c_{98}	$\rightarrow$

(Vezi şi scala sigmatică de la 3.7.5 (Lp4), precum şi http://app.inthlerom.ro/histo/HistoSetup.zip)

3.4.3. Generalizare - fractilele de ordinul *m*

Concept geometric care generalizează conceptele de cuartile, decile şi centile, aplicabil şi în cazul discret al distribuţiilor de frecvenţe empirice (seriilor statistice):

Fractile (cuantilele) de ordinul *m* = *(m-1)* valori care separă seria statistică ordonată în *m* subserii de volume egale, volumele fiind măsurate în număr de unităţi statistice şi, eventual, fracţiuni ale acestora. Evident $m = 2,3,...,n$.

Fractila (cuantila) superioară de ordinul *m* = cea mai mare fractilă de ordinul *m*. Lasă la dreapta sa *1/m* din aria distribuţiei (în cazul distribuţiilor continue).

Fractila (cuantila) inferioară de ordinul *m* = cea mai mică fractilă de ordinul *m*. Lasă la stânga sa *1/m* din aria distribuţiei (în cazul distribuţiilor continue).

Înlocuind *1/m* cu α (unde $0 < \alpha < 1$) se obţin noţiunile următoare pentru distribuţii continue.

3.4.4. α-cuantile unilaterale şi bilaterale

1° α-cuantile unilaterale

α-cuantila unilaterală superioară = punctul care lasă la DREAPTA sa proporţia α (respectiv procentul $\alpha \cdot 100$ %) din aria distribuţiei. Se notează x_α.

α-cuantila unilaterală inferioară = punctul care lasă la STÂNGA sa proporţia α (respectiv procentul $\alpha \cdot 100$ %) din aria distribuţiei. Se notează $x_{1-\alpha}$.

- $x_\alpha = -x_{1-\alpha}$ pentru orice $\alpha \Leftrightarrow$ distribuţia este simetrică faţă de zero.

Ex. distribuţia normală standard (vezi Anexa 2).

2° α-cuantile bilaterale

α-cuantila bilaterală superioară = punctul care lasă la DREAPTA sa proporţia $\alpha/2$ (respectiv procentul $\alpha/2 \cdot 100$ %) din aria distribuţiei. Se notează $x_{\alpha/2}$.

α-cuantila bilaterală inferioară = punctul care lasă la STÂNGA sa proporția $\alpha / 2$ (respectiv procentul $\alpha / 2 \cdot 100$ %) din aria distribuției. Se notează $x_{1-\alpha/2}$.

- *α-cuantila bilaterală superioară* este *$\alpha / 2$-cuantila unilaterală superioară*,
- *α-cuantila bilaterală inferioară* este *$\alpha / 2$-cuantila unilaterală inferioară*.
- Cele două *α-cuantile bilaterale* sunt egale și de semne contrare $\Leftrightarrow$ distribuția este simetrică față de origine.

Anexele 2-5 conțin α-cuantile ale celor mai uzitate distribuții teoretice.

L 2 Teste, exerciții şi probleme

TG2. Durata 160'' pe calculator.

Alegeti definitia corecta pentru mediana unui sir:
1. Valoarea care imparte seria statistica in doua subserii de volume egale.
2. O valoare care imparte sirul ordonat in doua subsiruri de volume egale.
3. O valoare care imparte sirul ordonat in doua subsiruri egale.

Media de viata intr-o populatie biologica este:
1. mediana seriei de varste la deces
2. media seriei de varste la deces
3. moda seriei de varste la deces

Proprietatea de aditivitate a mediei aritmetice se enunta corect astfel:
1. Media generala este egala cu media mediilor partiale.
2. Media generala este egala cu media mediilor partiale ponderate prin volumele seriilor partiale respective.
3. Media generala este suma mediilor partiale.

Alegeti raspunsul corect in legatura cu urmatoarea definitie:
'Se numesc cuartile de ordinul m, m valori care separa seria statistica ordonata in m subserii de volume egale (volumele fiind masurate in unitati statistice si eventual fractiuni ale acestora).'
1. definitia este corecta
2. definitia se refera la cuantile de ordinul m si acestea sunt in numar de m-1
3. definitia se refera la cuantile de ordinul m

Dintre principalii indicatori de tendinta centrala se prefera:
1. media datorita proprietatilor sale algebrice, in special datorita aditivitatii si pentru ca ia in considerare toate valorile seriei impreuna cu intreaga lor informatie
2. mediana deoarece exprima cel mai bine tendinta centrala, mai ales la distributii asimetrice si pentru ca nu e sensibila la valori extreme, in particular la valori aberante
3. moda pentru ca se observa cel mai usor si este intotdeauna element al seriei si pentru ca nu e sensibila la valori extreme, in particular la valori aberante

Formula mediei in cazul unei serii statistice grupate intr-o distributie de frecvente relative este:
1. M = (Σ xi) / N
2. M = Σ Fj * xj
3. M = (Σ Nj * xj) / (Σ Nj)

Indicatorul specific al tendintei centrale pentru variabilele de tip rang este:
1. mediana
2. moda
3. media

TC2. Durata 5'.

1. Gândind cantitativ, variabilitatea este concepută ca o ______________, iar gândind calitativ variabilitatea "devine" __________________.
2. Se foloseşte _________ la o serie mare pentru o orientare rapidă.
3. ___________ este indicatorul specific al tendinței centrale al variabilelor de tip rang, deci aceasta tratează valorile ca pe ____________.
4. Se notează cu _________ doza care omoară 50% din indivizii care au fost intoxicați cu doza respectivă.
5. Cea mai importantă proprietate a mediei este cea de _________________.
6. Q_3 lasă la stânga sa ________ din termeni şi fracțiuni ale acestora, iar Q_1 lasă la dreapta ________.
7. Decila superioară, notată ____, lasă la dreapta ____ şi deci este centila ____.
8. Cuantila inferioară de ordinul *m* lasă la dreapta sa _____________, iar la stânga ______.
9. Se ştie că $x_{0,8}$ în tabela cu α-cuantile inferioare ale distribuției normale standard este - 0,84. Cât este $x_{0,2}$? _______.
10. α-cuantila bilaterală superioară este ______ -cuantila unilaterală superioară.
11. Dacă 10%-cuantila unilaterală superioară este 1,28, atunci 20%-cuantila bilaterală inferioară este _______.

Exerciții sau probleme rezolvate

1.

Pentru 1962 de subiecți s-a determinat concentrația în sânge (în μg / 100 l) a unui principiu activ la 30 minute după administrarea sa orală şi s-a obținut seria de mai jos. Calculați indicatorul de tendință centrală *cel mai rapid* de calculat şi motivați alegerea.

x_j	N_j	x_j	N_j	x_j	N_j
0,15	5	0,33	92	0,48	62
0,17	15	0,34	96	0,49	54
0,18	22	0,35	99	0,50	46
0,20	26	0,37	102	0,51	42
0,22	28	0,39	106	0,52	36
0,23	32	0,40	101	0,53	31
0,24	45	0,41	98	0,55	26
0,26	59	0,42	87	0,57	24
0,27	61	0,44	85	0,58	19
0,29	67	0,45	78	0,59	15
0,31	73	0,46	72	0,61	9
0,32	84	0,47	65		N= 1962

Rezolvare:
În cazul unei distribuţii de frecvenţe cel mai uşor se observă modele. În acest caz, existând o singură modă Mo = 0,39 (corespunde frecvenţei maxime, 106), aceasta este indicatorul de tendinţă centrală cel mai rapid de calculat (observat).

2.

(a) Să se reprezinte ca histograme distribuţiile de frecvenţe din primele două coloane ale tabelelor următoare. (b) Să se calculeze pentru fiecare distribuţie mediana şi media şi (c) să se studieze legătura dintre egalitatea acestora şi simetria distribuţiei respective.

Distribuţia 1.

x_j	N_j	*Frecvenţe cumulate:*	$x_j \cdot N_j$
4	1	$N_1 = 1$	4
5	3	$N_1 + N_2 = 4$	15
6	1	$N_1 + N_2 + N_3 = 5$	6
	$N = \sum N_j = 5$		$T_1 = \sum x_j \cdot N_j = 25$

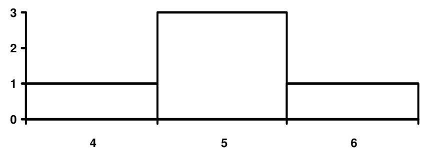
Distribuţia 2.

x_j	N_j	*Frecvenţe cumulate:*	$x_j \cdot N_j$
4	2	$N_1 = 2$	8
5	3	$N_1 + N_2 = 5$	15
7	1	$N_1 + N_2 + N_3 = 6$	7
	$N = \sum N_j = 6$		$T_1 = \sum x_j \cdot N_j = 30$

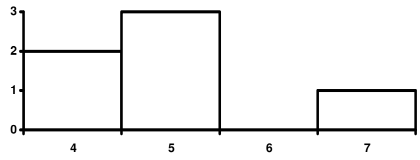
Rezolvare:

(a)

Distribuția 1:



Distribuția 2:



(b)

Distribuția 1:

Determinare mediană: *rg* = *N*/2 = 5/2 = 2,5 nu este întreg. Rotunjit prin adus, *rg* = 3. Deci mediana este termenul de rang 3. Acesta se determină ușor cu ajutorul coloanei de frecvențe cumulate, alăturată primelor două coloane din tabel. Se observă că termenul de rang 3 se află între primii 4 termeni (inventariați prin frecvența cumulată 4), deoarece rangul 3 este **mai mic sau egal** cu frecvența cumulată (4) dar **mai mare strict** decât frecvență cumulată anterioară (1). Prin urmare, termenul de rang 3 este un x_j de pe linia frecvenței cumulate 4 (vezi coloana 1). Deci *Me* = 5.

Media se calculează cu ajutorul ultimei coloane atașate tabelului. $M = T_1 / N$ = 25 / 5 = 5.

Distribuția 2:

Determinare mediană: *rg* = *N*/2 = 6/2 = 3. Deci mediana este semisuma dintre termeni de rang 3 și 4 ($Me = (x_3 + x_4)/2$). În coloana de frecvențe cumulate observăm că termenii de rang 3 și rang 4 se află între primii 5 termeni (inventariați prin frecvența cumulată 5), deoarece rangurile 3 și 4 sunt **mai mici sau egale** cu frecvența cumulată 5, dar **mai mari strict** decât frecvența cumulată anterioară (2). Prin urmare, termenii de rang 3, respectiv, 4 sunt valori x_j de pe linia frecvenței cumulate 5, adică au valoarea 5 (vezi coloana 1). Deci *Me* = (5+5)/2= 5.

Media se calculează cu ajutorul ultimei coloane atașată tabelului. $M = T_1 / N$ = 30 / 6 = 5.

(c) Se observă că la ambele distribuții cei doi indicatori coincid. Distribuția 1 ilustrează implicația "simetria unei distribuții ⇒ egalitatea între mediană și medie (*Me = M*)". Distribuția 2 demonstrează că implicația reciprocă nu este valabilă, adică deși cei doi indicatori sunt egali această distribuție nu este simetrică.

În concluzie, egalitatea celor doi indicatori este doar o condiție necesară (dar nu și suficientă) pentru simetria unei distribuții. Altfel spus, dacă cei doi indicatori sunt egali distribuția POATE fi simetrică.

Alcătuind contrara reciprocei pentru implicația de mai sus obținem că "inegalitatea celor doi indicatori ⇒ asimetria distribuției respective". Altfel spus, inegalitatea celor doi indicatori este o condiție suficientă (nu și necesară) pentru asimetria unei distribuții unimodale.

3.

S-a determinat concentraţia în sânge (în µg / 100 ml) a unui principiu activ la 30 minute după administrarea sa orală la 26 de pacienţi. S-a obţinut distribuţia de frecvenţe din primele două coloane ale tabelei de mai jos. Calculaţi indicatorul de tendinţă centrală adecvat şi motivaţi alegerea.

x_j	N_j	*Frecvenţe cumulate* (*Fc*):
<	5	*5*
0,1	1	*6*
0,2	3	*9*
0,3	4	*13*
0,4	6	*19*
0,5	3	*22*
0,7	2	*24*
0,8	1	*25*
0,9	1	*26*
1,0		
	$N = 26$	

Rezolvare:
Deoarece nu se cunosc toate valorile şirului, nu putem calcula media. 0,5 este o modă iar intervalul (0; 0,1) este un interval modal (provenit din grupare). Abandonăm modele fiind instabile – depind de grupări. Rămâne *Me*. Pentru mediană calculăm $rg = N/2 = 26 / 2 = 13$. Fiind un număr întreg mediana va fi semisuma termenilor de rang 13 şi 14. Pentru determinarea acestora calculăm coloana de frecvenţe cumulate în cadrul căreia se observă că termenul al 13-lea este ultimul din cei 13 termeni cu valoarea ≤ 0,4, adică are valoarea 0,4 iar termenul al 14-lea este primul termen cu valoarea 0,5. Deci $Me = (0,4 + 0,5) / 2 = 0,45$.

4.

Se dă următoarea distribuţie de frecvenţe absolute, formată din primele două linii următoare:

x_j	49	51	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	$N =$
N_j	1	3	8	16	25	32	39	36	30	25	20	14	8	4	2	1	1	265
Fc:	*1*	*4*	*12*	*28*	*53*	*85*	*124*	*160*	*190*	*215*	*235*	*249*	*257*	*261*	*263*	*264*	*265*	

Să se calculeze:

(a) Mediana.
(b) Cuartilele.
(c) Decila inferioară şi superioară.
(d) Centila inferioară şi superioară.
(e) Centila C_5
(f) Centila C_{95}.
(g) Cuantila inferioară de ordinul 5.
(h) Cuantila superioară de ordinul 6.

Rezolvare:
Seria fiind ordonată ascendent, deoarece este prezentată ca distribuţie de frecvenţe, calculăm linia de frecvenţe cumulate alipită tabelului de mai sus. Această coloană ne va ajuta să determinăm termenii cu anumite ranguri, după modelul din problema 2.

(a) Pentru *mediană* calculăm $rg = N / 2 = 265 / 2 = 132,5$. *Rg* nefiind întreg, îl rotunjim prin adaus şi deci Me = termenul de rang 133 = 58.
(b) Pentru *cuartila inferioară* calculăm $rg = N / 4 = 265 / 4 = 66,25$. Rangul fiind un număr fracţionar, Q_1 = termenul de rang 67 = 56. $Q_2 = Me = 58$.

Pentru *cuartila superioară* calculăm $rg = 3 \cdot N / 4 = 3 \cdot 265 / 4 = 198{,}75$. Rangul fiind un număr fracționar, Q_3 = termenul de rang 199 = 60.

(c) Pentru *decila inferioară* calculăm $rg = N / 10 = 265 / 10 = 26{,}5$. Rangul fiind un număr fracționar, D_1 = termenul de rang 27 = 54. Pentru *decila superioară* calculăm $rg = 9 \cdot N / 10 = 9 \cdot 265 / 10 = 238{,}5$. Rangul fiind un număr fracționar, D_9 este termenul de rang 239 = 62.

(d) Pentru *centila inferioară* calculăm $rg = N / 100 = 265 / 100 = 2{,}65$. Rangul fiind un număr fracționar, C_1 este termenul de rang 3 = 51. Pentru *centila superioară* calculăm $rg = 99 \cdot N / 100 = 99 \cdot 265 / 100 = 262{,}35$. Rangul fiind un număr fracționar, C_{99} este termenul de rang 263 = 65.

(e) Pentru C_5 calculăm $rg = 5 \cdot N / 100 = 5 \cdot 265 / 100 = 13{,}25$. Rangul fiind un număr fracționar, C_5 este termenul de rang 14 = 54.

(f) Pentru C_{95} calculăm $rg = 95 \cdot N / 100 = 95 \cdot 265 / 100 = 251{,}75$. Rangul fiind un număr fracționar, C_{95} este termenul de rang 252 = 63.

(g) Pentru *cuantila inferioară de ordinul 5* calculăm $rg = N / 5 = 265 / 5 = 53$. Rangul fiind un număr întreg, cuantila va fi semisuma dintre termenii de rang 53 și 54 = (55 + 56) / 2 = 55,5.

(h) Pentru *cuantila superioară de ordinul 6* calculăm $rg = 5 \cdot N / 6 = 5 \cdot 265 / 6 = 220{,}83$. Rangul fiind un număr fracționar, cuantila va fi termenul de rang 221 = 61.

5.

Se dau următoarele trei distribuții de frecvențe, definite prin asocierea primei coloane cu fiecare din următoarele trei coloane:

	S	**S'**	**S''**
x_i	N_i	N_i'	N_i''
2	3	1	1
3	5	1	2
4	7	2	3
5	6	2	5
6	3	3	8
7	2	6	5
8	2	7	3
9	1	5	2
10	1	3	1
N:	30	30	30

Să se calculeze pentru fiecare distribuție moda, mediana și media și să se observe relațiile de ordine dintre cei trei indicatori de tendință centrală. Să se pună diagnosticele de simetrie-asimetrie utilizând calea cea mai rapidă.

Rezolvare:

S: $Mo = 4$, $Me = (4 + 5) / 2 = 4{,}5$; $M = 146 / 30 = 4{,}87$.

$Mo < Me < M \Rightarrow$ distribuție unimodală asimetrică de stânga.

S': $Mo = 8$, $Me = (7 + 8) / 2 = 7{,}5$; $M = 214 / 30 = 7{,}13$.

$M < Me < Mo \Rightarrow$ distribuție unimodală asimetrică de dreapta.

S'': $Mo = 6$, $Me = 6$, $M = 180 / 30 = 6$.

$Mo = Me = M \Rightarrow$ POATE FI distribuție unimodală simetrică, ceea ce se verifică grafic (construind histograma).

Exerciții sau probleme propuse

6.

S-a determinat concentrația în sânge a unui principiu activ la 30 minute după administrarea sa orală, cu ajutorul unei metode cu sensibilitatea cuprinsă în domeniul (0,1 , 0,9) μg/100 ml și s-a obținut distribuția de frecvențe de mai jos. Indicați valoarea mediei fără a realiza calculul ei. Motivați rezultatul obținut.

x_j	0,20	0,25	0,30	0,35	0,40	0,45	0,50	0,55	0,60
N_j	1	2	3	5	7	5	3	2	1

7.

Pentru a se determina efectul antiagregant plachetar al unui nou medicament în tratamentul trombozelor arteriale și venoase s-au administrat, la 248 de pacienți, diferite doze din medicamentul studiat și s-a notat în cazul fiecăruia doza minimă la care a răspuns (exprimată în mg / kg corp). A rezultat următoarea distribuție de frecvențe absolute din coloanele notate x_j și N_j. (Pentru facilitarea rezolvării am alipit și coloanele cu frecvențe absolute cumulate.)

x_j	N_j	*Fc*	x_j	N_j	*Fc*	x_j	N_j	*Fc*
129,0	4	*4*	131,5	20	*69*	134,0	22	*219*
129,5	7	*11*	132,0	23	*92*	134,5	15	*234*
130,0	9	*20*	132,5	35	*127*	135,0	14	*248*
130,5	13	*33*	133,0	39	*166*		$N = 248$	
131,0	16	*49*	133,5	31	*197*			

Stabiliți prin sinteză numerică tipul de asimetrie al acestei distribuții, fără a calcula media.

8.

Pentru a se determina efectul analgezic al unui nou medicament în tratamentul bolilor de stomac s-a luat în studiu un lot de 437 pacienți. Subiecților li s-au administrat diferite doze din medicamentul studiat și s-a notat în cazul fiecăruia doza minimă la care a răspuns (exprimată în mg / kg corp). A rezultat următoarea distribuție de frecvențe și de frecvențe cumulate:

x_j	N_j	*Fc*	x_j	N_j	*Fc*	x_j	N_j	*Fc*
9,8	1	*1*	10,9	33	*201*	12,0	7	*416*
9,9	3	*4*	11,0	31	*232*	12,1	5	*421*
10	5	*9*	11,1	30	*262*	12,2	4	*425*
10,1	7	*16*	11,2	28	*290*	12,3	3	*428*
10,2	10	*26*	11,3	26	*316*	12,4	3	*431*
10,3	14	*40*	11,4	23	*339*	12,5	2	*433*
10,4	18	*58*	11,5	20	*359*	12,6	1	*434*
10,5	23	*81*	11,6	17	*376*	12,7	1	*435*
10,6	26	*107*	11,7	13	*389*	12,8	1	*436*
10,7	29	*136*	11,8	11	*400*	12,9	1	*437*
10,8	32	*168*	11,9	9	*409*		$N=437$	

Stabiliți prin sinteză numerică tipul de asimetrie al acestei distribuții, fără a calcula media.

9.

Se dă următoarea distribuție de frecvențe absolute la care am adăugat și frecvențele cumulate:

x_j	N_j	*Fc*	x_j	N_j	*Fc*	x_j	N_j	*Fc*
49	1	*1*	57	39	*124*	64	8	*257*
51	3	*4*	58	36	*160*	65	4	*261*
53	8	*12*	59	30	*190*	67	2	*263*
54	16	*28*	60	25	*215*	69	1	*264*
55	25	*53*	61	20	*235*	70	1	*265*
56	32	*85*	62	14	*249*			
							$N=265$	

Să se calculeze:

I. (a) Centila superioară.
   (b) Decila inferioară de ordinul 100.

II. (a) Cuantila inferioară de ordinul 50.
   (b) Cuartila superioară de ordinul 60.

III. (a) Cuantila superioară de ordinul 5.
   (b) Cuartila inferioară de ordinul 50.