



UNIVERSITATEA ECOLOGICĂ DIN BUCUREȘTI
FACULTATEA DE MANAGEMENT FINANCIAR



Program de studii universitare de licență: Finanțe și Bănci

Forma de învățământ: învățământ cu frecvență redusă

Suport de curs pentru autoinstruire la disciplina:
STATISTICĂ

Coordonator de disciplină:
Conf. univ. dr. Nicoleta CARAGEA

București
2021

CUPRINS

Introducere	7
a. Date privind titularul de disciplină	7
b. Date despre disciplină.....	7
c. Obiectivele disciplinei	7
d. Competențe acumulate după parcurgerea cursului.....	8
e. Resurse și mijloace de lucru	8
f. Structura cursului	9
g. Cerințe preliminare.....	10
h. Durata medie de studiu individual	10
i. Evaluarea	10
Unitatea de învățare 1.....	11
DOMENIUL ȘI METODA STATISTICII	11
1.1. Introducere	11
1.2. Conținut	11
1.2.1. Obiect și metodă.....	11
1.2.2. Concepte de bază.....	12
1.2.3. Statistica oficială în România.....	15
1.2.4. Diseminarea informației statistice.....	17
1.3. Rezumat	19
1.4. Test de autoevaluare a cunoștințelor	19
1.5. Bibliografie recomandată	20
Unitatea de învățare 2.....	21
SURSELE DE DATE STATISTICE	21
2.1. Introducere	21
2.2. Conținut	21
2.2.1. Surse statistice.....	21
2.2.2. Surse administrative.....	25
2.3. Rezumat	25
2.4. Test de autoevaluare a cunoștințelor	26
2.5. Bibliografie recomandată	27
Unitatea de învățare 3.....	28
SISTEMATIZAREA INFORMAȚIEI STATISTICE	28
3.1. Introducere	28

3.2. Conținut	28
3.2.1. Procedee de sistematizare	28
3.2.2. Gruparea datelor. Tipuri de grupări. Etapele grupării	29
3.2.3. Prezentarea datelor statistice	31
3.2.3.1. Reprezentarea datelor statistice sub formă de tabele.....	31
3.2.3.2. Reprezentarea grafică a datelor statistice – diagramele	32
3.3. Rezumat	38
3.4. Test de autoevaluare a cunoștințelor	39
3.5. Bibliografie recomandată	40
Unitatea de învățare 4.....	41
SERII DE DISTRIBUȚIE	41
4.1. Introducere	41
4.2. Conținut	41
4.2.1. Prezentare generală	41
4.2.2. Tipuri de serii de distribuție	41
4.2.3. Construcția unei serii de distribuție.....	42
4.3. Rezumat	47
4.4. Test de autoevaluare a cunoștințelor	48
4.5. Bibliografie recomandată	50
Unitatea de învățare 5.....	51
INDICATORII FRECVENȚEI	51
5.1. Introducere	51
5.2. Conținut	51
5.2.1. Frecvențe absolute și relative	51
5.2.2. Frecvențe cumulate	52
5.3. Rezumat	55
5.4. Test de autoevaluare a cunoștințelor	55
5.5. Bibliografie recomandată	58
5.6. Teme de control	59
Unitatea de învățare 6.....	60
INDICATORII TENDINȚEI CENTRALE. MEDIA	60
6.1. Introducere	60
6.2. Conținut	60
6.2.1. Media.....	61
6.2.2. Media aritmetică.....	62

6.2.2.1. Media aritmetică simplă	62
6.2.2.2. Media aritmetică ponderată	62
6.2.2.3. Centrele intervalelor	63
6.3. Rezumat	65
6.4. Test de autoevaluare a cunoștințelor	65
6.5. Bibliografie recomandată	68
Unitatea de învățare 7.....	69
ALTE TIPURI DE MEDIE. PROPRIETĂȚILE MEDIEI.....	69
7.1. Introducere	69
7.2. Conținut	69
7.2.1. Media armonică.....	69
7.2.2. Media pătratică.....	70
7.2.3. Media geometrică	71
7.2.4. Proprietățile mediei.....	72
7.3. Rezumat	76
7.4. Test de autoevaluare a cunoștințelor	76
7.5. Bibliografie recomandată	80
Unitatea de învățare 8.....	82
MEDIANA.....	82
8.1. Introducere	82
8.2. Conținut	82
8.2.1. Mediana seriilor simple	82
8.2.2. Mediana seriilor de frecvențe	85
8.3. Rezumat	88
8.4. Test de autoevaluare a cunoștințelor	88
8.5. Bibliografie recomandată	93
Unitatea de învățare 9.....	94
MODUL.....	94
9.1. Introducere	94
9.2. Conținut	94
9.2.1. Modul seriilor simple	94
9.2.2. Modul seriilor de distribuție.....	94
9.3. Rezumat	97
9.4. Test de autoevaluare a cunoștințelor	97
9.5. Bibliografie recomandată	100
9.6. Teme de control	101

Unitatea de învățare 10.....	102
INDICATORII DE NIVEL. QUANTILELE	102
10.1. Introducere	102
10.2. Conținut	102
10.2.1. Quartilele - Definiție	103
10.2.2. Algoritmul de calcul al quartilelor	103
10.3. Rezumat	105
10.4. Test de autoevaluare a cunoștințelor	105
10.5. Bibliografie recomandată	106
Unitatea de învățare 11.....	108
INDICATORII SIMPLI AI VARIAȚIEI	108
11.1. Introducere	108
11.2. Conținut	108
11.2.1. Amplitudinea.....	109
11.2.1. Abaterile individuale	110
11.3. Rezumat	111
11.4. Test de autoevaluare a cunoștințelor	111
11.5. Bibliografie recomandată	115
Unitatea de învățare 12.....	116
INDICATORII SINTETICI AI VARIAȚIEI	116
12.1. Introducere	116
12.2. Conținut	116
12.2.1. Abaterea medie.....	116
12.2.2. Dispersia.....	117
12.2.3. Abaterea standard	118
12.2.4. Coeficientul de variație.....	119
12.3. Rezumat	120
12.4. Test de autoevaluare a cunoștințelor	120
12.5. Bibliografie recomandată	123
12.6. Teme de control	123
Unitatea de învățare 13.....	124
MĂRIMILE RELATIVE	124
13.1. Introducere	124
13.2. Conținut	124
13.2.1. Mărimi relative de structură	125

13.2.2. Mărimi relative de intensitate	128
13.2.3. Mărimi relative de coordonare.....	129
13.2.4. Mărimi relative ale dinamicii și ale planului	130
13.3. Rezumat	131
13.4. Test de autoevaluare a cunoștințelor	132
13.5. Bibliografie recomandată	132
Unitatea de învățare 14.....	133
ELEMENTE DE TEORIA SELECȚIEI	133
14.1. Introducere	133
14.2. Conținut	133
14.2.1. Selecția.....	135
14.2.1.1. Metode de selecție.....	135
14.2.1.2. Tipuri de sondaj	136
14.2.2. Erorile cercetării selective.....	137
14.2.3. Mărimea eșantionului	138
14.3. Rezumat	139
14.4. Test de autoevaluare a cunoștințelor	140
14.5. Bibliografie recomandată	140
Răspunsuri la testele de autoevaluare a cunoștințelor	141

Introducere

Cursul este destinat, în principal, studenților Facultății de Științe Economice din cadrul Universității Ecologice din București – învățământ cu frecvență redusă (IFR), dar poate constitui un util material de studiu și pentru studenții la învățământ cu frecvență (IF). Cursul își propune să ofere, în primul rând studenților care se pregătesc să devină specialiști în domeniul economic, cunoștințele necesare pentru înțelegerea și familiarizarea cu conceptele de bază cu care operează statistica, precum și cu principalele instrumente și metode de analiză statistică, în vederea utilizării acestora în activitatea practică. De asemenea, prin conținutul de idei teoretice și al utilității practice, lucrarea se adresează tuturor celor care doresc să își îmbogățească cunoștințele în domeniul economic.

a. Date privind titularul de disciplină

Nume și prenume:	Conf. univ. dr. Nicoleta CARAGEA
E-mail:	nicoletacaragea@gmail.com

b. Date despre disciplină

Anul de studiu:	1
Semestrul:	2

c. Obiectivele disciplinei

Obiectivul general al disciplinei	Cunoașterea conceptelor de bază, a metodelor și utilizarea instrumentelor oferite de statistică în domeniul economic
Obiectivele specifice	Observarea statistică, prelucrarea rezultatelor observării și prezentarea lor în tabele și grafice Identificarea și utilizarea corectă a datelor din surse statistice și administrative Estimarea indicatorilor care definesc parametrii de nivel, structură și distribuție a fenomenelor sociale și factorii care le determină Înțelegerea semnificației rezultatelor estimării pe baza cercetărilor selective Înțelegerea argumentației și concluziilor din studii care aplică instrumente de analiză statistică și econometrică Analiza caracteristicilor cantitative ale fenomenelor și proceselor din domeniul economic și social Analiza influenței factorilor care determină nivelul variabilelor economice Utilizarea corectă a metodelor de măsurare și analiză statistică; înțelegerea semnificației rezultatelor și utilizarea testelor de verificare a ipotezelor statistice.

d. Competențe acumulate după parcurgerea cursului

Competențe profesionale	C1.1 Identificarea și definirea conceptelor, teoriilor, metodelor și instrumentelor de natură financiară în entitățile/organizațiile private și publice C1.2 Explicarea și interpretarea conceptelor, teoriilor, metodelor și instrumentelor de natură financiară în entitățile/organizațiile private și publice C1.3 Aplicarea conceptelor, teoriilor, metodelor și instrumentelor de natură financiară în entitățile private și publice pentru rezolvarea de probleme specifice C2.4 Identificarea criteriilor de selecție și aplicarea variantei adecvate pentru culegerea și analiza datelor economico- financiare C2.5 Efectuarea de analize economico-financiare curente pe baza datelor și informațiilor culese C3.1 Descrierea lucrărilor de natură economico-financiară realizate la nivelul entităților/organizațiilor private și publice C3.2 Explicarea lucrărilor de natură economico-financiară realizate la nivelul entităților/organizațiilor private și publice C3.3 Aplicarea cunoștințelor, metodelor, tehnicilor și instrumentelor pentru realizarea lucrărilor economico-financiare C3.4 Utilizarea unor indicatori de calitate pentru evaluarea realizării lucrărilor de natură economico-financiară C3.5 Elaborarea unui studiu referitor la realizarea lucrărilor economico- financiare
Competențe transversale	

e. Resurse și mijloace de lucru

Pentru o pregătire temeinică, vă sugerăm să efectuați toate testele de evaluare, astfel încât rezultatul pregătirii dumneavoastră să fie cât mai obiectiv. De asemenea, vă sugerăm să vă alcătuiți propriile dumneavoastră planuri și scheme, acestea ajutându-vă la o mai bună sistematizare a cunoștințelor dobândite.

Fiecare curs debutează cu prezentarea obiectivelor pe care trebuie să le atingeți – din punctul de vedere al nivelului de cunoștințe – prin studierea temei respective. Vă recomandăm să le citiți cu atenție și apoi, la sfârșitul cursului, să le revedeți pentru a verifica dacă le-ați atins în întregime.

Testele de evaluare prezente la sfârșitul fiecărui curs vă vor ajuta să verificați modalitatea specifică de învățare și să vă îmbunătățiți cunoștințele. Timpul recomandat rezolvării testelor este de 30 de minute. Rezolvați testele numai după studierea în întregime a unui curs și nu vă uitați la răspunsuri decât după rezolvarea testului și numai dacă nu reușiți să găsiți răspunsul prin recitirea cursului.

La fiecare curs există indicată o bibliografie selectivă pe care vă sfătuim să o parcurgeți. Dacă într-o anumită unitate de învățare se regăsesc referințe către alte resurse și materiale tipărite sau în format electronic, pe Internet, vă recomandăm să le parcurgeți.

f. Structura cursului

Cursul de Statistică este alcătuit din 14 capitole (fiecare capitol fiind aferent unei **unități de învățare**):

- Capitolul I. DOMENIUL ȘI METODA STATISTICII
- Capitolul II. SURSELE DE DATE STATISTICE
- Capitolul III. SISTEMATIZAREA INFORMAȚIEI STATISTICE
- Capitolul IV. SERII DE DISTRIBUȚIE
- Capitolul V. INDICATORII FRECVENȚEI
- Capitolul VI. INDICATORII TENDINȚEI CENTRALE
- Capitolul VII. MEDIA
- Capitolul VIII. MEDIANA
- Capitolul IX. MODUL
- Capitolul X. INDICATORII DE NIVEL. CUANTILE
- Capitolul XI. INDICATORII SIMPLI AI VARIAȚIEI
- Capitolul XII. INDICATORII SINTETICI AI VARIAȚIEI
- Capitolul XIII. MĂRIMILE RELATIVE
- Capitolul XIV. ELEMENTE DE TEORIA SELECȚIEI

La sfârșitul fiecărui capitol beneficiați de modele de teste de autoevaluare care vă vor ajuta să stabiliți singuri ritmul de învățare și necesitățile proprii de repetare a unor teme. La sfârșitul suportului de curs sunt prezentate răspunsurile corecte la testele de autoevaluare a cunoștințelor.

Totodată, sunt formulate **3 teme de control**, în vederea evaluării pe parcurs, astfel:

Tema 1 – Capitolul 2-5: „Realizați un referat de maximum 5 pagini în care să tratați sursele de date și sistematizarea datelor statistice, utilizând datele și metadatele din baza de date TEMPO”;

Tema 2 – Capitolul 6-9: „Realizați un referat de maximum 5 pagini în care să calculați indicatorii tendinței centrale ai unei serii de distribuție”;

Tema 3 – Capitolul 11-12: „Realizați un referat de maximum 5 pagini în care să calculați indicatorii simpli și sintetici ai unei serii de distribuție”.

Temele de control se prezintă de către fiecare student, iar rezultatele acestei evaluări se comunică de către cadrul didactic în mod direct și nemijlocit studenților.

Cursul de STATISTICĂ poate fi studiat atât în întregime, potrivit ordinii prestabilite a capitolelor, dar se poate și fragmenta în funcție de interesul propriu mai accentuat pentru anumite

teme. Însă, în vederea susținerii examenului este obligatorie parcurgerea tuturor celor 14 capitole și efectuarea testelor prezentate.

g. Cerințe preliminare

Pentru o mai bună înțelegere a materiei, este necesară coroborarea cunoștințelor dobândite în cadrul acestei discipline cu cele acumulate la matematică și la microeconomie, informatică, având în vedere că disciplina se studiază în anul I, semestrul 2.

h. Durata medie de studiu individual

Durata medie de învățare estimăm a fi de aproximativ 14 de ore, iar pentru examen ar fi necesar un studiu de o săptămână.

i. Evaluarea

La încheierea studierii cursului, este obligatorie realizarea celor trei teme de control.

Înainte de examen este indicat să parcurgeți din nou toată materia, cu atenție, durata estimată pentru această activitate fiind de aproximativ o săptămână.

Forma de evaluare (E-examen, C-colocviu/test final, LP-lucrări de control)		
Stabilirea notei finale (procentaje)	- evaluarea finală	50%
	- activități aplicative /laborator/lucrări practice/proiect etc.	20%
	- teste pe parcursul semestrului	20%
	- teme de control	10%

Unitatea de învățare 1.

DOMENIUL ȘI METODA STATISTICII



1.1. Introducere

În cuprinsul acestui capitol sunt prezentate cunoștințele generale referitoare la studiul statisticii, precum și conceptele, principiile și legitățile fundamentale ale statisticii.

Durata medie de parcurgere a acestui capitol este de aproximativ o oră.



1.2. Conținut

1.2.1. Obiect și metodă

Termenul de "statistică" este unul familiar, pe care îl auzim sau folosim în fiecare zi. Este asociat adesea cu calcule, cu un rezultat exprimat procentual sau cu un tabel de date. Cu toate acestea, termenul trebuie privit într-o accepțiune mai largă.

Poate v-ați întrebat deseori care este valoarea medie a cifrei de afaceri sau care a fost creșterea medie a acestora într-un interval de timp, ce valoare a vânzărilor ar putea avea o firmă de telefonie mobilă care ar lansa pe piață un nou produs sau serviciu, sau care ar fi cel mai potrivit loc de muncă disponibil când veți absolvi facultatea. Cine va fi ales noul președinte al țării la următoarele alegeri prezidențiale? Sau poate vă întrebați care ar fi contribuția creșterii nivelului de educație al populației ocupate la dezvoltarea economică a țării.

Răspunsurile la astfel de întrebări, dar și la multe altele, vin din înțelegerea corectă a valorilor numerice înregistrate de o variabilă căreia îi vom spune generic "de interes", a fluctuațiilor sau a tendinței de creștere sau de scădere a acestor valori într-o anumită perioadă de timp.

Conform dicționarilor publicate de institutele de cercetare din România, statistica este o știință care, folosind calculul probabilităților, studiază fenomenele și procesele de tip colectiv (din societate, natură etc.) din punct de vedere cantitativ în scopul descrierii acestora și a descoperirii legilor care guvernează manifestarea lor. Altfel definită, statistica este știința care se ocupă cu

descrierea și analiza numerică a fenomenelor de masă, dezvăluind particularitățile lor de volum, structură, dinamică, conexiune, precum și regularitățile sau legile ce le guvernează .

Sensul termenului statistică poate avea mai multe conotații. Una dintre ele este aceea de date sau colecție de date, referitoare la un anumit fenomen sau domeniu. Într-o altă accepțiune, statistica înseamnă activitatea de producere a datelor statistice, de culegere, prelucrare și prezentare a informației statistice sub diverse forme (diseminare), la dispoziția utilizatorilor. Mai mult, termenul de statistică se utilizează și pentru a desemna metodele statistice aplicate în cunoașterea fenomenelor de masă, sau teoria statistică, implicând expunerea sistematică a conceptelor și metodelor de cercetare statistică. Statistica are ca obiect studierea aspectelor cantitativ-numerice ale fenomenelor de masă, a dimensiunii, dinamicii și structurii acestora, a raporturilor de interdependență și a altor aspecte care pot fi caracterizate numeric. Fenomenele de masă sunt fenomene care se produc sub acțiunea comună și repetată a unui număr mare de factori, cu caracter sistematic sau întâmplător. Fenomenele de masă sunt caracterizate printr-o mare diversitate, formele individuale de manifestare diferind de la o unitate la alta în funcție de modul în care se combină acțiunea acestor factori. Esența acestor fenomene poate fi pusă în evidență numai prin studierea unui număr mare de cazuri. De exemplu, productivitatea muncii, la nivelul economiei naționale, poate fi cunoscută numai prin analiza informațiilor privind rezultatele activității tuturor unităților din economia națională, respectiv cantitatea de forță de muncă utilizată (dintr-o ramură de activitate sau dintr-o unitate teritorială) sau a unei părți semnificative a acestora.

Metoda statistică este constituită dintr-un ansamblu de operații, tehnici, procedee și metode de cercetare statistică. Specificitatea metodei statistice constă în faptul că se bazează pe observarea unui număr mare de cazuri și utilizează instrumente derivate din teoria probabilităților, dezvoltate în cadrul statisticii matematice.

1.2.2. Concepte de bază

Pentru organizarea unei cercetări corecte este necesar să se folosească un limbaj specific fiecărei disciplinei științifice. Acest lucru este valabil și pentru statistică, care a reușit să-și elaboreze propriile noțiuni, concepte de bază pe care să le folosească pe parcursul întregului demers statistic. În acest sens, cercetarea statistică operează cu câteva concepte specifice: colectivitatea sau populația statistică, unitatea statistică, caracteristicile sau variabilele statistice, datele statistice și indicatorii statistici.

Colectivitatea sau populația statistică reprezintă totalitatea elementelor sau a manifestărilor de aceeași natură, asupra cărora se efectuează cercetarea. În funcție de obiectivul și particularitățile

cercetării, populația statistică poate fi formată din persoane sau grupuri de persoane, din instituții, din obiecte, din tranzacții economice, din evenimente etc.

De exemplu:

- O cercetare privind veniturile populației se realizează pe o colectivitate formată din gospodării;
- O cercetare referitoare la nivelul de educație a populației – se realizează pe o colectivitate formată din persoane;
- Colectivitatea care face obiectul unei investigații privind producția și eficiența economică este alcătuită din unități economice;
- Colectivitatea cercetată în cadrul unui studiu privind activitatea sistemului de sănătate cuprinde unități sanitare (toate spitalele, policlinicile, centrele de tratament etc.);
- Studiul statistic al calității producției se realizează pe colectivități formate din obiecte (produse finite, semifabricate etc.);
- Studiul statistic al natalității sau mortalității se realizează pe colectivități compuse din evenimente (nașteri sau decese).

Colectivitatea supusă cercetării poate fi totală (cercetare exhaustivă), atunci când se fac înregistrări referitoare la toate elementele care formează obiectul studiului, sau parțială (eșantion), atunci când se fac înregistrări referitoare numai la o parte din această colectivitate.

Unitățile statistice reprezintă mulțimea numărabilă de elemente care compun colectivitatea statistică. De exemplu, unitatea statistică a unei cercetări referitoare la veniturile populației poate fi gospodăria sau persoana. Unitatea statistică a unei cercetări privind calitatea producției este produsul finit, semifabricatul sau piesa căreia i se testează caracteristicile.

Unitățile statistice pot fi simple sau complexe. Acestea din urmă sunt formate din mai multe unități simple. O astfel de unitate este gospodăria¹. Unitățile pe care se realizează cercetarea salariilor sunt, de asemenea unități complexe: unitățile economice la care se face înregistrarea datelor. Fiecare element al colectivității este purtătorul unei *caracteristici* supuse observării statistice. Caracteristica statistică reprezintă acea proprietate/însușire care este comună tuturor unităților unei colectivități statistice cercetate. Formele sau nivelurile concrete ale acestora, denumite *variante* sau *valori*, diferă de la o unitate la alta (sau în timp, în cazul aceleiași unități) sub influența unui complex de

¹ Prin **gospodărie** se înțelege grupul de două sau mai multe persoane care locuiesc împreună în mod obișnuit, având, în general, legături de rudenie și care se gospodăresc (fac menajul) în comun, participând în totalitate sau parțial la formarea veniturilor și la cheltuirea lor. Persoana care nu aparține de o gospodărie și care declară că locuiește și se gospodărește singură se consideră gospodărie formată dintr-o singură persoană. Se consideră membri ai gospodăriei și persoanele plecate din localitate pentru o perioadă mai mare de 6 luni, care se află în țară sau străinătate, dacă acestea păstrează legături familiale cu gospodăria.

factori. Numărul unităților la care se înregistrează aceeași variantă sau valoare poartă denumirea de *frecvență* a variantei/valorii respective.

Statistica operează, în principal, cu frecvențele diferitelor variante ale caracteristicilor calitative, iar analiza statistică a acestor caracteristici se realizează prin metode ce diferă de cele utilizate în studiul variabilelor cantitative.

În cazul unei cercetări referitoare la activitatea economică, de exemplu, se vor înregistra informații privind volumul producției, cheltuielile de producție și profitul, numărul de salariați și dotarea cu active corporale etc. din fiecare unitate economică.

Evaluarea situației existente pe piața forței de muncă, evoluția fenomenelor de ocupare, șomaj și inactivitate va înregistra informații referitoare la populația în vârstă de muncă, populația activă, populația ocupată, numărul șomerilor etc. Un alt exemplu este cel al unei cercetări referitoare la venituri, care va cuprinde, în principal, informații numerice despre nivelul veniturilor de care beneficiază gospodăriile (din activități salariale, din agricultură, din activități pe cont propriu, din prestații sociale etc.), despre impozitele și contribuțiile de asigurări sociale plătite de gospodării, despre mărimea gospodăriei, vârsta membrilor acesteia etc.

Prin natura lor, variabilele cantitative se grupează în:

- variabile discrete - pot lua numai anumite valori (exprimate prin numere întregi);
- variabile continue - pot lua orice valori într-un interval finit sau infinit.

Numărul salariaților unei firme sau numărul de persoane care compun o gospodărie sunt variabile discrete, în timp ce valoarea producției, nivelul salariilor sau nivelul veniturilor sunt variabile continue (chiar dacă nivelurile înregistrate ale acestora apar ca valori ale unor variabile discrete). În practică, variabilele discrete sunt tratate ca fiind continue, dacă valorile lor posibile sunt numeroase (de ex., numărul salariaților unităților economice).

Datele statistice exprimă valori ale unor caracteristici cantitative ale unităților statistice, precum și ale unor grupe ale colectivității sau ale colectivității în ansamblul ei. Informațiile numerice sunt date statistice numai dacă sunt definite sub aspectul conținutului, al unității de măsură, al identificării unității/grupului/colectivității la care se referă, precum și al localizării în spațiu și în timp. Datele se obțin fie direct prin observare, fie sunt rezultate ale prelucrării datelor primare. De exemplu, date statistice sunt nivelurile cifrei de afaceri înregistrate de firmele X și Y din București, în luna septembrie 2012.

Indicatorii statistici reprezintă expresia numerică ce exprimă manifestările unor fenomene economice sau sociale, caracterizând nivelul și structura, modificările în timp și variația spațială, precum și interdependențele dintre ele.

Indicatorii statistici sunt rezultate ale prelucrării informațiilor primare înregistrate în cursul unei observări statistice sau a unor informații cuprinse în publicații statistice sau baze de date. De exemplu, modificarea cifrei de afaceri a firmei X în luna septembrie 2009 față de aceeași lună a anului precedent este un indicator statistic.

Pentru cunoașterea fenomenelor de masă, indicatorii statistici îndeplinesc mai multe funcții și anume: de măsurare, de comparare, de analiză, de sinteză; de estimare, de verificare a ipotezelor și/sau de testare a semnificației parametrilor utilizați.

Simpla enumerare a principalelor funcții² ale indicatorilor statistici pune în evidență o multitudine de aspecte care trebuie avute în vedere la elaborarea și folosirea acestora în analiză, inclusiv stabilirea condițiilor și limitelor în care pot fi utilizați indicatorii statistici în raport cu conținutul specific al fenomenelor, al surselor de informație de care se dispune, cu scopul cercetării.

Indicatorii statistici se pot grupa în indicatori primari și derivați:

1. Indicatori primari (mărimi absolute) – exprimă direct, general nivelul caracteristicii cercetate. Se pot obține prin înregistrarea directă, centralizarea datelor sau prin însumarea parțială sau totală a datelor individuale; prezintă o capacitate relativ limitată de descriere a fenomenului/procesului analizat, și nu permite realizarea unor aprecieri calitative, însă reprezintă punctul de plecare al analizei statistice;
2. Indicatori derivați – se obțin prin prelucrarea indicatorilor primari (absoluți) și fac posibilă analiza aspectelor calitative ale fenomenelor și proceselor analizate (de exemplu: mărimi relative, mărimi medii, indicatori ai variației, indici, indicatori ai corelației, etc.).

1.2.3. Statistica oficială în România

Statistica oficială în România se desfășoară prin serviciile de statistică oficială și este organizată și coordonată de Institutul Național de Statistică, organ de specialitate al administrației publice centrale, cu personalitate juridică, în subordinea Guvernului, finanțat de la bugetul de stat. Prin activitățile desfășurate, Institutul Național de Statistică urmărește să asigure

² Pentru elaborarea și utilizarea corectă a indicatorilor statistici este esențială îndeplinirea unor cerințe de principiu, generale. În acest sens, Yule (1945) precizează condițiile care ar trebui să le îndeplinească un astfel de indicator și anume:

- să fie definit în mod obiectiv, independent de dorința utilizatorului;
- să depindă determinarea sa de toate valorile individuale înregistrate;
- să aibă o semnificație concretă, ușor de înțeles chiar și de nespecialiști;
- să fie simplu și rapid de calculat;
- să fie puțin sensibil la fluctuațiile de selecție (să nu prezinte valori puternic diferite, dacă se calculează pe baza mai multor eșantioane, de același volum, extrase prin același procedeu din aceeași colectivitate);
- să se preteze la calcule algebrice (să poată fi utilizat în operații de comparare a mai multor serii statistice sau în operații de agregare/dezagregare).

dezvoltarea unor mecanisme și instrumente adaptate desfășurării activităților statistice în funcție de priorități, să consolideze structura și funcționalitatea Sistemului Statistic Național și să îmbunătățească permanent capacitatea sa de a răspunde în timp util la creșterea nevoilor de date și informații statistice ale utilizatorilor.

În realizarea obiectului său de activitate, stabilit potrivit legii, care se întemeiază pe principiul autonomiei, confidențialității, transparenței, relevanței, proporționalității, deontologiei statistice și raportului cost/eficiență, Institutul Național de Statistică are următoarele **atribuții**:

- a) elaborează sistemul de indicatori statistici, metodologiile de calcul, tehnologiile și standardele specifice de obținere a indicatorilor;
- b) organizează și conduce cercetările statistice referitoare la fenomenele și procesele economico-sociale, prin recensăminte, anchete statistice totale sau prin sondaje;
- c) proiectează, tipărește și difuzează formularele cercetărilor statistice și instrucțiunile de completare aferente;
- d) culege, prelucrează și stochează date și informații în vederea asigurării bazelor de date statistice;
- e) coordonează elaborarea clasificărilor și a nomenclatoarelor unitare de interes național din domeniul său de activitate;
- f) proiectează, realizează și exploatează sistemul informatic statistic;
- g) elaborează studii și analize pentru caracterizarea evoluției economico-sociale;
- h) desfășoară activități de cercetare științifică în domeniul statisticii;
- i) informează opinia publică, precum și autoritățile publice cu privire la evoluția economică și socială a țării și furnizează utilizatorilor interesați indicatorii statistici obținuți;
- j) organizează și conduce unitățile subordonate, stabilește atribuțiile și normele de funcționare ale acestora;
- k) colaborează cu ministerele și cu alte organe de specialitate ale administrației publice centrale, precum și cu serviciile publice, pentru compatibilizarea sistemului statisticii oficiale cu celelalte sisteme informaționale; sprijină aceste organisme pentru asigurarea pregătirii personalului cu atribuții în activitatea de statistică;
- l) avizează metodologiile cercetărilor statistice organizate de ministere, instituții centrale și de alte servicii publice, în vederea asigurării oportunității cercetărilor și corectitudinii metodelor statistice utilizate;
- m) organizează perfecționarea pregătirii profesionale în domeniul statisticii a personalului institutului, din unitățile sale subordonate și din celelalte servicii de statistică oficială și acționează pentru formarea unei culturi statistice la nivel național;

- n) reprezintă România în relațiile internaționale în domeniul statistic și cooperează cu organizații similare din alte țări, cu organismele de specialitate ale Organizației Națiunilor Unite și agențiile sale, cu alte organisme internaționale;
- o) asigură compatibilitatea sistemului statistic național cu sistemele statistice utilizate de organismele Organizației Națiunilor Unite, ale Uniunii Europene și de alte organisme internaționale, conform obligațiilor asumate.

La nivelul Uniunii Europene funcționează *Oficiul de Statistică al Comunității Europene (Eurostat)*, cu sediul în Luxemburg. Activitatea Eurostat este structurată pe nouă teme (Statistică generală și regională; Economie și finanțe; Populație și condiții sociale; Industrie, comerț și servicii; Agricultură și pescuit; Comerț exterior; Transport; Mediu și energie; Știință și tehnologie) și șase teme/direcții speciale (Indicatorii structurali, Euro-Indicatorii, Anuarul statistic, Dezvoltarea durabilă, Finanțele publice, Indicii armonizați ai prețurilor de consum). Eurostat asigură Uniunii Europene date comparabile și armonizate necesare pentru definirea, implementarea și analiza politicilor comunitare, iar informațiile statistice oferite de Eurostat servesc și mediului de afaceri, organizațiilor profesionale, mediului academic, organizațiilor neguvernamentale și presei.

1.2.4. Diseminarea informației statistice

Diseminarea este un proces prin care informațiile statistice sunt puse la dispoziția utilizatorilor din diferite domenii – instituțiile administrației publice centrale și locale, mass-media, cercetare, persoane fizice etc.

În cadrul sau sub coordonarea Institutului Național de Statistică se realizează cercetările statistice de anvergură națională (recensăminte, anchete, sisteme de rapoarte statistice, calculul sistemului de indicatori economici și sociali). Principalele informații statistice, rezultate ale cercetărilor statistice și date administrative, sunt puse la dispoziția utilizatorilor sub forma unor publicații statistice, realizate sub formă tipărită sau electronică (CD ROM) sau on-line (<http://www.insse.ro>).

Procesul de diseminare se desfășoară în baza respectării principiilor transparenței și confidențialității datelor statistice, în conformitate cu prevederile legale în vigoare privind organizarea și funcționarea statisticii oficiale în România și, mai ales, cu cele ale Codului de practici al statisticilor europene.

Cea mai importantă și complexă publicație editată de Institutul Național de Statistică este Anuarul Statistic al României, care cuprinde date statistice referitoare la situația economică și socială din România (populația, forța de muncă, veniturile, cheltuielile și consumul populației, protecția

socială, sănătate, învățământ, cultură, sport, cercetare, conturi naționale, investiții și imobilizări corporale, prețuri, rezultate și performanțe ale întreprinderilor, agricultură și silvicultură, industrie și construcții, transporturi și telecomunicații, comerț exterior, interior și servicii, turism, finanțe, justiție, alegeri), la nivel național și teritorial (regional și județean), precum și date din statistica internațională. Datele se referă, în principal, la situația din anul precedent celui în care se publică anuarul (T-1), precum și la evoluția principalilor indicatori în ultimii 5-6 ani.

Din lucrările de referință fac parte și Anuarul de comerț internațional, Starea socială și economică a României, Sistemul Conturilor Naționale, Ancheta structurală în agricultură, Rezultate și performanțe ale întreprinderilor.

Publicațiile periodice cuprind, printre altele, Buletinul statistic lunar, Buletinul statistic de comerț internațional, Buletinul statistic de industrie, Buletinul statistic de prețuri, Forța de muncă în România - Ocupare și șomaj, iar informațiile operative sunt publicate, sub formă de serii statistice specializate, din industrie, turism, mediu, agricultură, servicii, etc.

Prin intermediul componentei de e-commerce, disponibilă pe website, prin care se poate accesa baza de date TEMPO – online, utilizatorii de informații statistice pot solicita informații din bazele de date ale INS (baza de date privind populația, baza de date de comerț exterior, baza de date în profil teritorial, Registrul statistic al întreprinderilor și Registrul unităților administrativ-teritoriale), precum și alte date din fondul național de date statistice, gestionat de INS.

Informațiile statistice colectate și prelucrate de Eurostat, în colaborare cu oficiile de statistică ale statelor membre ale Uniunii Europene, sunt publicate în Anuarul statistic (Eurostat Yearbook), sub forma unor publicații statistice complexe (Statistical books) sau breviare statistice (Pocketbooks), tabele (Detailed tables), precum și în cadrul unei serii de analize statistice (Statistics in focus). Toate publicațiile sunt disponibile în format electronic pe site-ul Eurostat (<http://ec.europa.eu/eurostat>), care cuprinde și o bogată bază de date (Database – Data navigation tree).

Să ne reamintim

- Termenul de "statistică" este unul familiar, pe care îl auzim sau folosim în fiecare zi. Este asociat adesea cu calcule, cu un rezultat exprimat procentual sau cu un tabel de date. Cu toate acestea, termenul trebuie privit într-o accepțiune mai largă.
- Metoda statistică este constituită dintr-un ansamblu de operații, tehnici, procedee și metode de cercetare statistică. Specificitatea metodei statistice constă în faptul că se bazează pe observarea unui număr mare de cazuri și utilizează instrumente derivate din teoria probabilităților, dezvoltate în cadrul statisticii matematice.

- Datele statistice exprimă valori ale unor caracteristici cantitative ale unităților statistice, precum și ale unor grupe ale colectivității sau ale colectivității în ansamblul ei.
- Indicatorii statistici reprezintă expresia numerică ce exprimă manifestările unor fenomene economice sau sociale, caracterizând nivelul și structura, modificările în timp și variația spațială, precum și interdependențele dintre ele.
- Statistica oficială în România se desfășoară prin serviciile de statistică oficială și este organizată și coordonată de Institutul Național de Statistică.



1.3. Rezumat

Termenul statistică poate avea mai multe sensuri: date sau colecție de date, referitoare la un anumit fenomen sau domeniu. Într-o altă accepțiune, statistica înseamnă activitatea de producere a datelor statistice, de culegere, prelucrare și prezentare a informației statistice sub diverse forme (diseminare), la dispoziția utilizatorilor. Mai mult, termenul de statistică se utilizează și pentru a desemna metodele statistice aplicate în cunoașterea fenomenelor de masă, sau teoria statistică, implicând expunerea sistematică a conceptelor și metodelor de cercetare statistică.



1.4. Test de autoevaluare a cunoștințelor

1) Institutul Național de Statistică are următoarele atribuții:

- proiectează, realizează și exploatează sistemul informatic statistic;
- elaborează metodologiile de calcul, tehnologiile și standardele specifice de obținere a indicatorilor, la nivelul Uniunii Europene;
- elaborează studii și analize pentru caracterizarea evoluției economico-sociale;
- desfășoară activități de cercetare științifică în domeniul statisticii;
- colaborează cu ministerele și cu alte organe de specialitate ale administrației publice centrale, precum și cu serviciile publice, pentru compatibilizarea sistemului statisticii oficiale cu celelalte sisteme informaționale; sprijină aceste organisme pentru asigurarea pregătirii personalului cu atribuții în activitatea de statistică;

Care dintre variante nu este corectă?

2) Institutul Național de Statistică are următoarele atribuții:

- a. elaborează sistemul de indicatori statistici, metodologiile de calcul, tehnologiile și standardele specifice de obținere a indicatorilor;
- b. organizează și conduce cercetările statistice referitoare la fenomenele și procesele economico-sociale, prin recensăminte, anchete statistice totale sau prin sondaje;
- c. proiectează, tipărește și difuzează formularele cercetărilor statistice și instrucțiunile de completare aferente;
- d. culege, prelucrează și stochează date și informații în vederea asigurării bazelor de date statistice;
- e. toate atribuțiile enumerate mai sus.

Alegeți varianta corectă.

3) Procesul de diseminare se desfășoară în baza respectării principiilor:

- a. transparenței;
- b. transparenței și confidențialității;
- c. universalității.

Alegeți varianta corectă.



1.5. Bibliografie recomandată

- a. Caragea Nicoleta, *Statistica- Concepte si metode de analiza a datelor*, Editura MUSTANG, București, 2015
- b. Voineagu, V. , *Statistică –baze teoretice și aplicații*, Editura Economică, București, 2007
- c. Georgescu-Roegen N., *Metoda statistică*, Editura Expert, București, 1998.
- d. Andrei T., Stancu S., Pele D.T., *Statistică. Teorie și aplicații* (ediția a II-a), Editura Economică, București, 2003.
- e. Biji M., Biji E.M., Lilea E., Anghelache C., *Tratat de statistică*, Editura Economică, București, 2002.

Unitatea de învățare 2.

SURSELE DE DATE STATISTICE



2.1. Introducere

În cadrul acestui capitol, studenților le sunt prezentate sursele de date pe care le utilizează statistica oficială din România.

Durata medie de parcurgere a acestui capitol este de aproximativ o oră.



2.2. Conținut

Datele statistice utilizate pentru analiza unui fenomen economic sau social pot fi obținute din două categorii de surse: statistice și administrative.

2.2.1. Surse statistice

Sursele statistice sunt cercetări complexe, realizate conform unui program care include următoarele etape:

- observarea statistică;
- prelucrarea primară a rezultatelor observării;
- constituirea unor baze de microdate (date înregistrate la nivelul unităților statistice);
- construirea de serii statistice și prezentarea lor sub formă de tabele și grafice;
- calculul unor indicatori cu caracter descriptiv (de nivel, structură, variație, dinamică);
- analiza legăturilor dintre fenomene în vederea evidențierii cauzelor și a explicării fenomenelor;
- testarea valabilității și relevanței parametrilor calculați, în vederea determinării gradului de încredere cu care trebuie tratate concluziile cercetării.

Observarea statistică se realizează prin înregistrarea datelor referitoare la valorile pe care le iau caracteristicile înscrise în programul cercetării la toate unitățile care compun colectivitatea

statistică. Se utilizează, în principal, trei modalități de observare statistică: recensămintele, anchetele sau sondajele și rapoartele statistice.

Recensămintele sunt cercetări statistice de mare amploare, care presupun înregistrarea, pentru toate unitățile colectivității, a valorilor unui set de caracteristici menit să asigure caracterizarea complexă a fenomenului studiat.

Principalul obiectiv al recensământului este de a asigura datele statistice pentru caracterizarea schimbărilor produse în numărul și structura populației, a gospodăriilor și locuințelor, a resurselor și forței de muncă, de a asigura o bază riguroasă pentru calculele specifice în domeniul demografic și al forței de muncă. Totodată, recensământul asigură informațiile necesare pentru constituirea unei baze de sondaj necesare proiectării și realizării cercetărilor selective în gospodăriile populației în vederea caracterizării nivelului de trai și a condițiilor de locuit, ocupării forței de muncă, a șomajului și a altor aspecte ale vieții sociale.

La recensământ se înregistrează cetățenii români cu domiciliul în țară, precum și persoanele de altă cetățenie sau fără cetățenie care au domiciliul sau au reședință în România. Cetățenii au obligația de a furniza, pe proprie răspundere, informații corecte și complete. De asemenea, la recensământ se înregistrează clădirile de locuit, precum și locuințele situate în orice clădire, indiferent de forma de proprietate, de pe teritoriul României.

În general, recensămintele se realizează periodic, la intervale mai mari sau mai mici de timp, dar pot fi și ocazionale, organizate pentru acoperirea unei nevoi speciale de informații.

Cele mai cunoscute recensăminte sunt cele ale populației și locuințelor, realizate la intervale relativ mari (de aproximativ 10 ani). În ultimii 50 ani, în România au fost organizate astfel de recensăminte în 1956, 1966, 1977, 1992, 2002 și 2011. Recensămintele populației se realizează în scopul cunoașterii cât mai exacte a numărului și distribuției teritoriale a populației, a structurilor socio-culturale și economice, precum și a fondului locativ și a condițiilor de locuit ale populației. Pentru organizarea activităților de înregistrare și pentru verificarea exhaustivității înregistrării persoanelor au fost utilizate la ultimul recensământ și informații din Registrul de Evidență Informatizată a Persoanei. De asemenea, recensământul a adus un plus de informații, prin includerea în programul de înregistrare a unor indicatori noi privind caracteristicile demografice, de migrație, economice și de clădiri, astfel încât programul de prezentare și publicare a rezultatelor să răspundă cerințelor de informare ale utilizatorilor interni și internaționali.

Prelucrarea informațiilor constă, în principal, în gruparea populației, a gospodăriilor și a locuințelor după caracteristicile cuprinse în programul recensământului și calculul unor indicatori medii. Rezultatele prelucrării sunt publicate în format tipografic și electronic.

Se organizează, de asemenea, recensăminte industriale și recensăminte agricole.

Cercetările statistice selective (anchetele sau sondajele statistice) sunt cercetări parțiale. Acestea presupun observarea unui număr restrâns de unități ale colectivității, selectate după reguli stabilite în conformitate cu exigențele teoriei selecției, prelucrarea informațiilor colectate prin calculul unor indicatori de nivel, structură, variație, corelație etc. (denumiți estimatori sau statistici) și inferența statistică, adică extinderea rezultatelor prelucrării la colectivitatea totală, prin estimarea indicatorilor aferenți populației totale (denumiți parametri). Cercetările statistice selective sunt organizate periodic, la intervale mai mici (lunar, trimestrial, anual) sau mai mari (la doi, trei sau cinci ani), precum și ocazional, atunci când apar nevoi speciale de informație într-un anumit domeniu.

În România, cercetarea statistică în domeniul economic și social se bazează pe câteva anchete de mare amploare realizate în unități economice, în instituții și în gospodării.

realizate în unități economice

- ancheta structurală în întreprinderi;
- cercetările selective privind prețurile și tarifele mărfurilor și serviciilor care intră în consumul populației, prețurile de producție ale produselor și serviciilor industriale și prețurile principalelor produse vândute de producătorii particulari în piețele agroalimentare;
- anchetele de conjunctură;
- ancheta lunară privind numărul de salariați și veniturile salariale
- ancheta anuală asupra câștigurilor salariale și costului forței de muncă;
- ancheta anuală privind activitatea de cercetare-dezvoltare

realizate în gospodăriile populației

- ancheta trimestrială asupra forței de muncă
- ancheta anuală a bugetelor de familie
- ancheta anuală asupra condițiilor de viață
- ancheta privind sănătatea populației
- ancheta privind utilizarea timpului
- ancheta privind educația adulților

Aceste anchete au drept obiectiv culegerea unui set amplu de informații de tip cantitativ-numeric și calitativ privind activitatea economică și condițiile de viață, realizându-se pe eșantioane de mare volum, capabile să asigure relevanța statistică a rezultatelor estimării diferiților parametri, pe ansamblul populației și pe principalele grupe ale acesteia.

Se realizează, de asemenea, anchete modul, care au scopul de a răspunde unor necesități imediate de informații statistice. În ultimii ani, Institutul Național de Statistică a realizat mai multe astfel de anchete, de exemplu: ancheta privind tranziția de la școală la muncă, activitățile copiilor, ocuparea persoanelor cu disabilități și educația permanentă, sănătatea și siguranța muncii, ca module complementare la ancheta asupra forței de muncă.

Un tip special de cercetări selective, înrudite cu cele statistice, îl reprezintă anchetele sociologice și cele de opinie, care colectează informații calitative și se realizează, în general pe eșantioane de volum mult mai mic decât cele statistice. Asigurarea relevanței concluziilor cercetării impune însă și în aceste anchete respectarea unor reguli de selecție a unităților observate, de prelucrare a informațiilor colectate și de extindere a rezultatelor fundamentate pe teoria selecției și statistica matematică.

Activitatea de marketing se bazează, de asemenea, pe realizarea unor anchete referitoare la poziția pe piață, la preferințele clienților și piața de desfacere potențială pentru un produs nou³.

Un domeniu important al cercetării statistice selective este controlul calității și fiabilității producției, care presupune măsurarea uneia sau mai multor caracteristici tehnice pentru un număr restrâns de piese, semifabricate sau piese finite în vederea estimării parametrilor de calitate ai unui întreg lot de produse.

Rapoartele statistice constituie o formă de cercetare statistică totală, care constă în culegerea de către organisme specializate în lucrări de statistică (institute, oficii, direcții teritoriale sau de ramură etc.) a unor informații statistice sub formă de rapoarte, prin completarea unor formulare sau chestionare de către unitățile implicate în realizarea unei anumite activități. Astfel de rapoarte statistice, cuprinzând informații cu privire la diferite caracteristici cantitative și calitative ale activității desfășurate sunt completate, de exemplu, de către unitățile de învățământ, de cultură, de sănătate și asistență socială, de unitățile de transport, poștă și telecomunicații, de primării și organisme de specialitate ale consiliilor județene etc. De asemenea, se consideră rapoarte statistice

³ În practica economică, în special în cercetări de marketing, există și anchete realizate pe eșantioane a căror selecție nu respectă cu rigurozitate regulile statistice ale cercetării selective, fiind formate dintr-un număr mai mult sau mai puțin restrâns de unități (persoane, gospodării, unități economice sau instituții) care au răspuns la chestionarele adresate unui număr mare de unități (actuali sau potențiali clienți). Rezultatele prelucrării informațiilor colectate prin astfel de anchete au o valoare orientativă, fără pretenția de rigurozitate a celor obținute din anchete statistice.

datele furnizate din bilanțurile contabile completate de operatorii economici, unitățile social-culturale, unitățile administrației publice și organizațiile neguvernamentale.

2.2.2. Surse administrative

Sursele administrative constau în baze de date sau alte colecții de informații statistice existente la nivelul diferitelor instituții publice (ministere, oficii sau autorități naționale etc.) sau private constituite în cursul activității acestora, ca o componentă intrinsecă a activității și ca instrument de fundamentare și susținere a acesteia. La Ministerul Finanțelor există, de exemplu, date din Registrul Fiscal, din bilanțurile contabile ale unităților economice, date privind impozitele plătite de persoanele fizice și asociațiile familiale, precum și datele privind execuția bugetului de stat, a bugetelor locale și a bugetului asigurărilor sociale de stat; la Oficiul Național al Registrului Comerțului – date din Registrul Comerțului, la Banca Națională a României – date privind balanța de plăți externe, bilanțurile monetare ale băncilor, creditul extern și masa monetară; la Ministerul Justiției și Inspectoratul General al Poliției – date din statistica judiciară. Tot în categoria surselor administrative pot fi încadrate bazele de date ale unităților economice. Informațiile din aceste baze pot fi utilizate pentru analize statistice ale activității unităților respective.

Să ne reamintim

- Statistica oficială din România utilizează surse statistice și surse administrative de date.
- Sursele statistice sunt cercetări complexe, care pot fi selective sau exhaustive.
- Sursele administrative constau în baze de date sau alte colecții de informații statistice existente la nivelul diferitelor instituții publice (ministere, oficii sau autorități naționale etc.) sau private constituite în cursul activității acestora, ca o componentă intrinsecă a activității și ca instrument de fundamentare și susținere a acesteia.



2.3. Rezumat

Cercetările statistice selective sunt cercetări parțiale. Acestea presupun observarea unui număr restrâns de unități ale colectivității, selectate după reguli stabilite în conformitate cu exigențele teoriei selecției, prelucrarea informațiilor colectate prin calculul unor indicatori de nivel, structură, variație, corelație etc. (denumiți estimatori sau statistici) și inferența statistică, adică extinderea

rezultatelor prelucrării la colectivitatea totală, prin estimarea indicatorilor aferenți populației totale (denumiți parametri).



2.4. Test de autoevaluare a cunoștințelor

1) Sursele statistice sunt cercetări complexe, realizate conform unui program care include următoarele etape:

- a. analiza în timp a fenomenului studiat;
- b. observarea statistică, prelucrarea primară a rezultatelor observării, constituirea unor baze de microdate și construirea de serii statistice și prezentarea lor sub formă de tabele și grafice;
- c. calculul unor indicatori cu caracter descriptiv (de nivel, structură, variație, dinamică);
- d. analiza legăturilor dintre fenomene în vederea evidențierii cauzelor și a explicării fenomenelor și testarea valabilității și relevanței parametrilor calculați, în vederea determinării gradului de încredere cu care trebuie tratate concluziile cercetării.

Care dintre variante nu este corectă?

2) Cercetările statistice selective:

- a. sunt cercetări statistice totale;
- b. se mai numesc sondaje sau cercetări selective;
- c. presupun observarea unui număr restrâns de unități ale colectivității, selectate după reguli stabilite în conformitate cu exigențele teoriei selecției;
- d. sunt organizate periodic, la intervale mai mici (lunar, trimestrial, anual) sau mai mari (la doi, trei sau cinci ani), precum și ocazional, atunci când apar nevoi speciale de informație într-un anumit domeniu;
- e. sunt cercetări statistice parțiale.

Care dintre variante nu este corectă?

3) Rapoartele statistice:

- a. constituie o formă de cercetare statistică totală;
- b. presupun culegerea de către organisme specializate - în lucrări de statistică - a unor informații statistice;

- c. cuprind informații cu privire la diferite caracteristici cantitative și calitative ale activității desfășurate;
- d. presupun completarea unor formulare sau chestionare de către unitățile implicate în realizarea unei anumite activități;
- e. toate variantele de mai sus.

Alegeți varianta corectă.



2.5. Bibliografie recomandată

- a. Caragea Nicoleta, *Statistica- Concepte si metode de analiza a datelor*, Editura MUSTANG, București, 2015
- b. Voineagu, V. , *Statistică –baze teoretice și aplicații*, Editura Economică, București, 2007
- c. Georgescu-Roegen N., *Metoda statistică*, Editura Expert, București, 1998.
- d. Andrei T., Stancu S., Pele D.T., *Statistică. Teorie și aplicații* (ediția a II-a), Editura Economică, București, 2003.
- e. Biji M., Biji E.M., Lilea E., Anghelache C., *Tratat de statistică*, Editura Economică, București, 2002.

Unitatea de învățare 3.

SISTEMATIZAREA INFORMAȚIEI STATISTICE



3.1. Introducere

În cadrul acestui capitol sunt prezentate metodele de sistematizare și vizualizare a datelor statistice.

Durata medie de parcurgere a acestui capitol este de aproximativ o oră.



3.2. Conținut

În foarte multe situații, în special în cele în care ne propunem studii practice, datele de care dispunem sunt numeroase, iar operarea lor este dificilă sau uneori imposibil de realizat. De aceea, specialiștii în statistică au elaborat o serie de metode și procedee de sistematizare a datelor pentru ca acestea să fie prezentate într-o formă rezumată.

Sistematizarea și prezentarea datelor statistice reprezintă etape ale procesului de prelucrare statistică și vizează obținerea unei imagini a fenomenului studiat, într-o manieră în care mai ușor de analizat și interpretat.

Sistematizarea presupune ordonarea datelor după anumite caracteristici, iar rezultatul procesului de sistematizare se prezintă sub forma unor serii de date, tabele sau grafice.

3.2.1. Procedee de sistematizare

Centralizarea datelor, în sensul cel mai larg, constă în punerea în comun a tuturor informațiilor colectate prin programul cercetării statistice a fenomenului studiat și în determinarea indicatorilor agregați pentru toate caracteristicile ce pot fi însumate direct sau care admit folosirea unor coeficienți de echivalență (prețul unitar, consumul de timp de muncă etc.).

Centralizarea datelor poate fi simplă și pe grupe. Centralizarea simplă înseamnă agregarea valorilor individuale ale caracteristicilor pentru toate unitățile colectivității care permit însumarea din punctul de vedere al conținutului în scopul obținerii indicatorilor totalizatori. În urma centralizării simple se obțin informații privind volumul unităților (simple sau complexe) care compun colectivitatea statistică și valoarea totalizată a caracteristicilor care au fost cuprinse în

programul observării, și care sunt exprimate în unități de măsură însumabile (tone, hectare, milioane lei, număr persoane etc.).

Centralizarea simplă se folosește frecvent, când se urmărește obținerea unor informații generale înaintea terminării prelucrării datelor. Astfel, de exemplu, în cazul recensământului populației se comunică destul de repede numărul populației, în timp ce prelucrarea și calculul indicatorilor derivați și finalizează abia după o perioadă mai mare de timp.

Gruparea este o operațiune complexă, a cărei realizare presupune cunoașterea fenomenului analizat și luarea în considerare a particularităților informației statistice primare. Nu există metode mecanice, universal valabile, de stabilire a numărului de grupe sau a mărimii și limitelor intervalelor.

3.2.2. Gruparea datelor. Tipuri de grupări. Etapele grupării

Gruparea este un proces iterativ, al cărui rezultat final trebuie să răspundă exigențelor analizei statistice. De pildă, în stabilirea numărului de grupe trebuie să se aibă în vedere, pe de o parte, că acesta trebuie să fie suficient de mare încât să poată oferi informații cât mai analitice cu privire la structura fenomenului analizat și, pe de altă parte, că nu trebuie să ducă la o segregare prea mare a colectivității, astfel încât informația referitoare la fiecare grupă să fie relevantă și semnificativă din punct de vedere statistic.

Există totuși câteva metode care facilitează stabilirea paramerilor grupării, cu ajutorul cărora se poate realiza mecanic o primă grupare, care poate constitui o bază de pornire în construirea unei variante finale.

Principalele motive pentru care se realizează gruparea datelor sunt următoarele:

- asigurarea unei structuri care să facă posibilă o gestionare sistematică a acestora;
- cunoașterea structurii colectivității la un moment dat, precum și a modificărilor structurale ale colectivității într-o anumită perioadă de timp;
- cunoașterea modului de manifestare a fenomenelor, prin analiza legăturilor dintre variabilele înregistrate.

Grupările statistice își găsesc o largă aplicare în toate domeniile vieții economice și sociale. Varietatea grupărilor statistice utilizate în practică impune separarea lor după mai multe criterii.

După numărul caracteristicilor folosite, grupările pot fi:

- *Grupări simple*, care constau în repartizarea unităților colectivității pe clase în funcție de variația unei singure caracteristici. Grupările simple se pot realiza pentru toate caracteristicile din programul observării sau numai pentru acelea care urmează să fie

prelucrate în continuare. Pe baza grupărilor simple se poate stabili care este frecvența corespunzătoare fiecărei grupe, precum și valorile centralizate pentru celelalte caracteristici care interesează. Prin gruparea simplă se poate stabili structura colectivității în funcție de caracteristicile de grupare precum și legătura între acesta și celelalte caracteristici pentru care se centralizează, pe grupe, valori individuale.

- *Grupările combinate* sunt acele grupări statistice în care separarea datelor se face cel puțin după două caracteristici. În acest caz, se separă mai întâi unitățile pe grupe după prima caracteristică și apoi, fiecare grupă după cea de-a doua caracteristică. Caracteristicile în funcție de care se face o grupare combinată trebuie să fie interdependente, respectiv să existe o legătură de natură cauzală, logică etc. De exemplu, gruparea firmelor după cifra de afaceri și după numărul de salariați; gruparea angajaților din cadrul unei firme după vechimea în muncă, nivelul de instruire și mărimea salariului.

După conținutul caracteristicii, grupările statistice se clasifică astfel:

- *Grupările cronologice* se obțin dacă datele privind colectivitatea studiată se clasifică după o caracteristică de timp.
- *Grupările teritoriale* rezultă atunci când colectivitatea se ordonează pe grupe după o caracteristică de spațiu (țară, regiune, județ, etc).

După tipul caracteristicii, grupările pot fi:

- Grupări după o caracteristică calitativă - sunt cele mai omogene grupări, deoarece în aceeași grupă se cuprind numai unități care prezintă aceeași formă de manifestare. De exemplu, gruparea salariaților după ocupație, sau gruparea populației după starea civilă.
- Grupări după o caracteristică cantitativă. De exemplu, gruparea firmelor după cifra de afaceri sau după numărul de salariați, gruparea gospodăriilor după venit etc.

După numărul valorilor înregistrate pentru caracteristica de grupare, grupările statistice se clasifică astfel:

- *Grupări pe variante* - se aplică în cazurile în care numărul valorilor înregistrate pentru caracteristica de grupare este redus, fiecare valoare înregistrată constituie o grupă. De exemplu, gruparea gospodăriilor după numărul persoanelor care o compun, sau gruparea populației active pe medii de rezidență (observație: gruparea după o caracteristică calitativă este întotdeauna o grupare de variante).
- *Grupări pe intervale* se folosesc în cazurile în care gradul de variație al caracteristicii este mare, deci când gruparea pe variante conduce la obținerea unui număr prea mare de clase (de exemplu, gruparea salariaților după mărimea veniturilor).

3.2.3. Prezentarea datelor statistice

În vederea aplicării metodelor de calcul și de interpretare statistică, rezultatele sistematizării datelor se prezintă sub formă de tabele, grafice sau serii statistice.

3.2.3.1. Reprezentarea datelor statistice sub formă de tabele

Tabelele se utilizează în toate etapele cercetării statistice: în procesul colectării datelor, pe parcursul prelucrării, la analiza și interpretarea informațiilor statistice, la prezentarea și/sau publicarea indicatorilor statistici.

Tabelele statistice reprezintă rețele de rânduri și coloane, care servesc la prezentarea sistematizată a informațiilor statistice.

Elementele de bază ale tabelelor sunt:

- titlul tabelului
- unitatea (unitățile de măsură) ale variabilelor cuprinse în tabel
- sursa datelor
- notele explicative.

Tabelele statistice sunt extrem de variate și se clasifică în primul rând după etapa de cercetare statistică în care se întocmesc. Astfel, în *etapa observării* se întocmesc tabele descriptive sau enumerative folosite pentru înregistrarea datelor primare. Se mai numesc și tabele simple. În etapa prelucrării se întocmesc tabele de prelucrare ca instrumente intermediare pentru centralizarea datelor sau pentru aplicarea unui algoritm de calcul al indicatorilor derivați.

Tabelele simple sunt cele în care se prezintă indicatorii statistici ordonați cronologic, teritorial sau structural.

Tabelul pe grupe se folosește când se aplică gruparea simplă cu valori centralizate pentru frecvențele corespunzătoare grupelor și pentru valorile caracteristicilor dependente de variația caracteristicii principale de grupare.

În practică, în special în cazul prezentării datelor care provin din cercetări statistice mai complexe (recensămintele, de exemplu), se apelează la sistematizarea datelor în tabele mutidimensionale, denumite *hipercuburi*.

3.2.3.2. Reprezentarea grafică a datelor statistice – diagramele

Oricât de bine ar fi gândite tabelele statistice, utilizarea lor repetată într-o aceeași secțiune a lucrării este obositoare și determină o receptare dificilă a mesajului informativ, uneori suprapunerea sau asocierea defectuoasă a datelor din diferite tabele.

Reprezentările grafice reprezintă o modalitate expresivă de prezentare a informațiilor statistice, constituind o metodă rapidă de vizualizare a legăturilor dintre fenomenele economice și sociale analizate. Graficele pun în evidență caracteristicile acestor fenomene: structuri, corelații, tendințe, regularități, dar și anomalii, erori sau omisiuni în datele primare.

Elementele de bază ale graficelor sunt: titlul graficului, scara de reprezentare, rețeaua graficului, legenda, sursa datelor, notele explicative, graficul propriu-zis.

Principalele tipuri de diagrame

Pentru reprezentarea grafică a datelor statistice, Microsoft Office Excel oferă posibilitatea creării mai multor tipuri de diagrame, astfel încât datele să fie prezentate în cel mai semnificativ mod. Când veți crea o diagramă, aveți posibilitatea să selectați unul dintre următoarele tipuri:

- diagrame coloană
- diagrame liniare
- diagrame de structură radială
- diagrame bară
- diagrame zonă
- diagrame prin puncte
- diagrame bursiere
- diagrame suprafață
- diagrame de structură inelară
- diagrame cu bule
- diagrame radar

Prezentăm în continuare câteva exemple ale celor mai utilizate diagrame.

□ Diagrama în formă de coloană

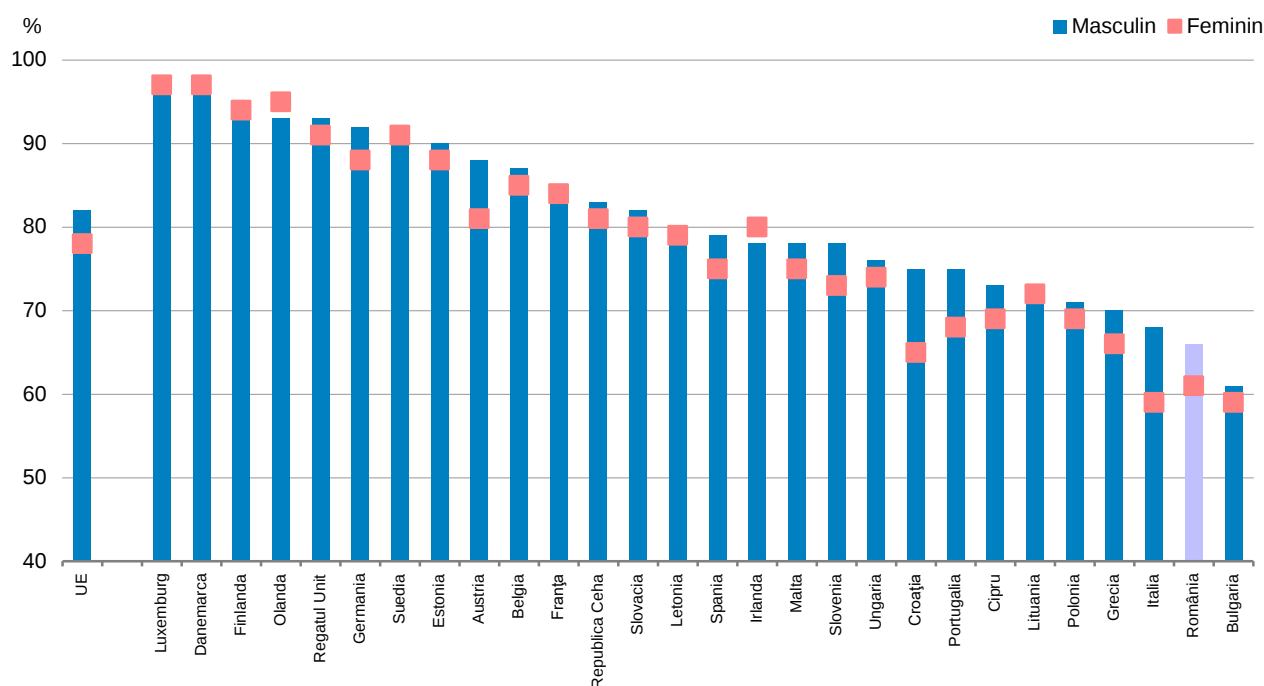
O diagramă în formă de coloană (bare verticale) ilustrează comparații între niveluri diferite ale aceluiași indicator statistic în perioade de timp, sau spații diferite. Datele vor fi aranjate în coloane sau rânduri într-o foaie de lucru, putând fi astfel reprezentate într-o diagramă coloană. În diagramele coloană, categoriile sunt de obicei organizate pe axa orizontală iar valorile pe axa verticală.

În general, acest tip de reprezentare grafică se utilizează în cazul datelor care prezintă o anumită ordine a variantelor, cum ar fi seriile cronologice precum și pentru reprezentarea comparațiilor între elementele aceleiași populații statistice.

Se poate utiliza un tip de diagramă *coloană grupată* când variabilele înregistrează mai multe categorii, care reprezintă:

- Intervale de valori (de exemplu, număr de elemente)
- Aranjamente specifice de scale (de exemplu, total de acord, de acord, neutru, dezacord, dezacord total)
- Nume care nu sunt într-o anumită ordine (de exemplu, nume de elemente, nume geografice sau nume de persoane)
- Variante ale variabilei înregistrate în perioade diferite de timp.

Figura 1. Ponderea persoanelor de 16-74 ani din statele membre ale UE, care au utilizat computerul în ultimele 12 luni, pe sexe, în anul 2015 (%)

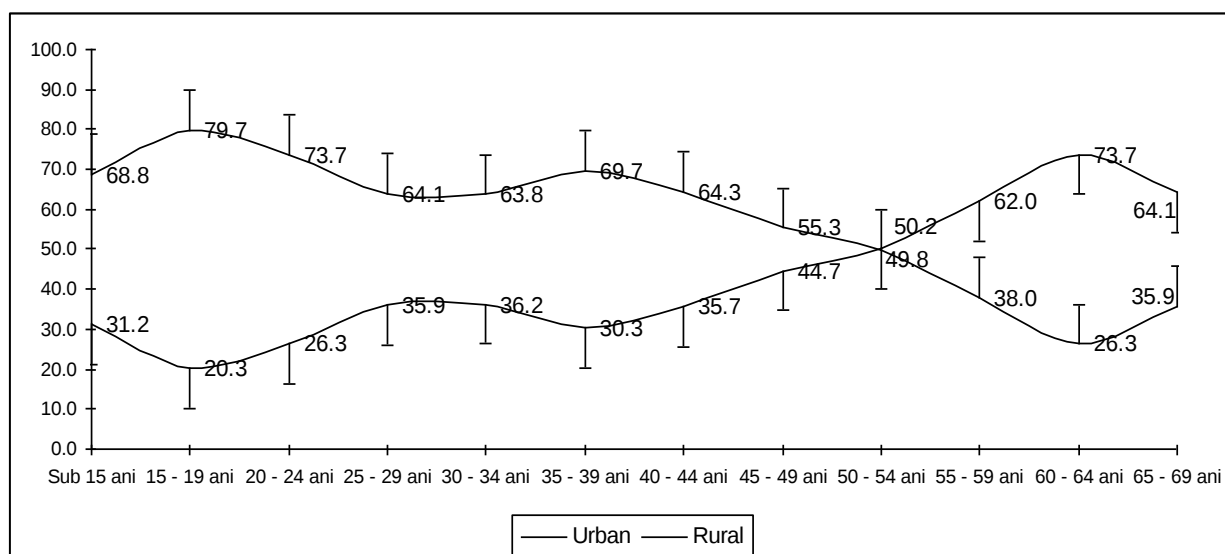


Sursa: EUROSTAT, <http://ec.europa.eu/eurostat/data/database>

□ Diagrama liniară

Datele care sunt aranjate în coloane sau rânduri pot fi reprezentate într-o diagramă linie.

Figura 2. Structura pe grupe de vârstă și medii de rezidență a populației plecată în străinătate pentru 12 luni și peste, în anul 2016



Sursa: Prelucrări pe baza datelor anchetei forței de muncă în gospodării

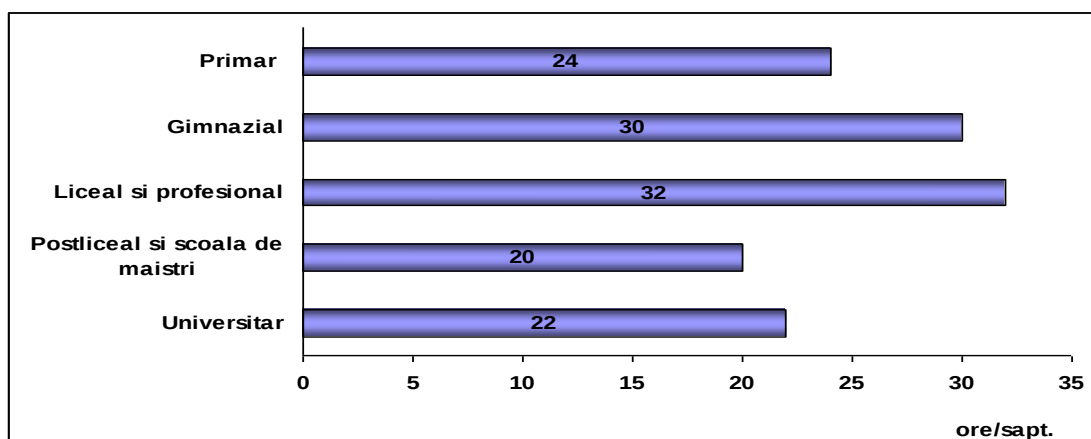
Diagramele linie pot afișa date continue în timp, în funcție de o scală comună și, de aceea, sunt ideale pentru afișarea tendințelor datelor la intervale egale. Într-o diagramă linie, datele din categorii sunt distribuite egal pe axa orizontală și toate datele valori sunt distribuite egal pe axa verticală. Ar trebui utilizată o diagramă liniară dacă etichetele categoriilor sunt text și reprezintă valori spațiate egal cum ar fi luni, trimestre sau ani fiscali.

Aceasta se aplică în special dacă există mai multe serii pentru o singură serie, luați în considerare utilizarea unei diagrame de categorii. De asemenea, ar trebui să utilizați o diagramă liniară dacă aveți câteva etichete numerice spațiate egal, în special ani.

□ Diagrama în formă de bară orizontală

Asemeni diagramei în formă de coloană, graficul în formă de bară orizontală se utilizează pentru ilustrarea comparațiilor între elementele unei populații statistice. Variantele sunt organizate pe ordonată iar valorile acestora pe abscisă.

Figura 3. Numărul maxim de ore de studiu în instituțiile de învățământ, pe niveluri de educație



Sursa: Planurile-cadru ale Ministerului Educației, Cercetării și Tineretului

□ Diagramele de structură radială

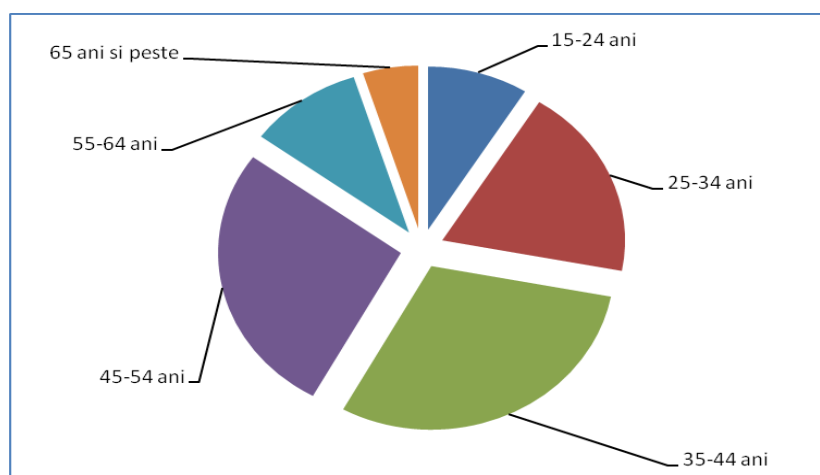
Diagrama radială, denumită și cerc de structură sau diagramă circulară se utilizează atunci când se dorește reprezentarea grafică a structurii populației statistice, după o anumită caracteristică (Fig. 5).

Acest tip de grafic se utilizează atunci când:

- valorile individuale prezintă variații semnificative;
- numărul de variante („felii”) nu este foarte mare (între 3 și 10).

O diagramă circulară afișează dimensiunea proporțională a elementelor care alcătuiesc o serie de date față de suma elementelor. Diagrama afișează întotdeauna doar o serie de date și este utilă atunci când se dorește evidențierea unui element semnificativ.

Figura 4. Structura pe grupe de vârstă a populației ocupate, în anul 2016



Sursa: ancheta asupra forței de muncă (AMIGO), INS

Utilizarea unei diagrame de structură radială este recomandată atunci când:

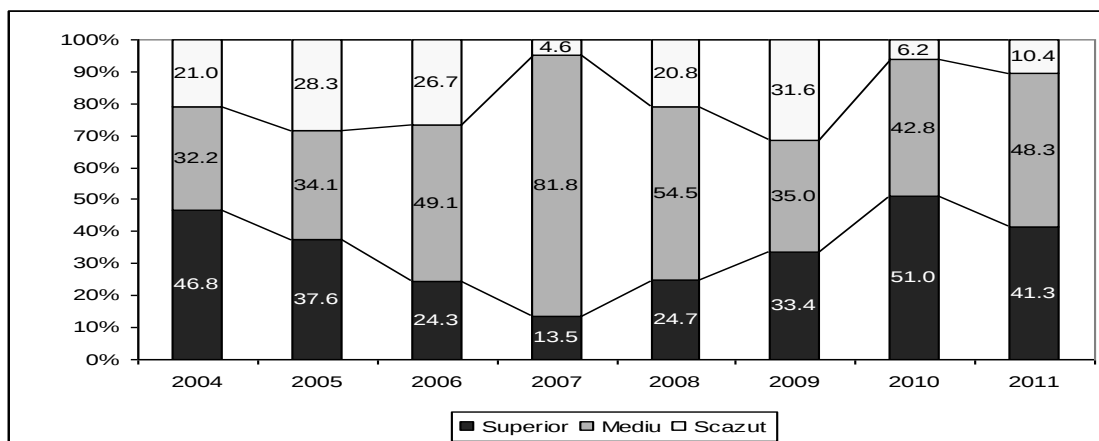
- Nu aveți decât o serie de date pe care doriți să o reprezentați.

- Nici una dintre valorile pe care doriți să le reprezentați nu este negativă.
- Aproape nici una dintre valorile pe care doriți să le reprezentați nu este valoare zero.
- Nu aveți mai mult de șapte categorii.
- Categoriile reprezintă părți din structura radială întreagă.

Datele care sunt aranjate în coloane sau rânduri într-o singură foaie de lucru pot fi reprezentate, de asemenea, și într-o diagramă de structură inelară. Asemenea unei diagrame de structură radială, o diagramă de structură inelară arată relația părților cu întregul, dar poate conține mai multe serii de date.

□ Diagramele care prezintă dinamica structurii

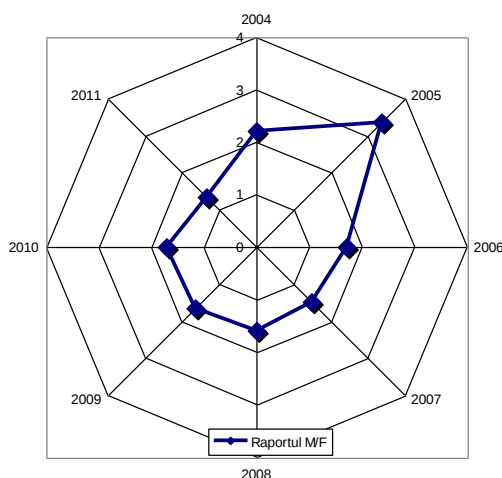
Figura 5. Structura populației plecată la studii în străinătate pentru 12 luni și peste, pe niveluri de educație, în perioada 2004-2011



Sursa: Prelucrări pe baza datelor AMIGO

□ Diagramele radar

Figura 6. Evoluția indicelui parității de gen pentru persoanele ocupate plecate la lucru în statele UE-27 pentru 6-12 luni, în perioada 2004-2011



Sursa: Prelucrări pe baza datelor AMIGO

□ Diagramele figurative și hărțile

Diagramele figurative sunt grafice în formă de coloană, în care coloanele se înlocuiesc cu imagini care ilustrează fenomenul analizat.

Figura 7. Numărul căsătoriilor din România, după grupa de vârstă a soților, în anul 2015



Sursa: Anuarul Statistic al României, INS, 2016

Reprezentările cu ajutorul hărților sunt utilizate pentru evidențierea proceselor economice și sociale care apar în profil teritorial. De exemplu, densitatea populației poate fi sugerată prin variații ale intensității culorii, în funcție de densitatea fiecărei regiuni.

Figura 8. Diagrama ilustrativă



La întocmirea tabelor și a graficelor se respectă următoarele reguli:

- tabelele și graficele trebuie să faciliteze perceperea rapidă și exactă a informațiilor prezentate;
- să cuprindă informațiile strict necesare caracterizării fenomenului analizat;
- să aibă un titlu scurt, clar, dar care sintetizează conținutul informativ al datelor prezentate;

- dacă se utilizează mai multe tabele sau/și grafice, acestea se numerotează;
- să se precizeze unitatea (unitățile de măsură) ale variabilelor;
- să fie însoțite de note explicative sau observații, care să explice metodologiile de colectare a datelor și de calcul al indicatorilor statistici etc.;
- să se indice sursa datelor.

În afara acestor reguli comune de elaborare a tabelelor și a graficelor, pot exista și anumite reguli specifice tabelelor, ca de exemplu:

- liniile și coloanele tabelului se numerotează pentru înlesnirea identificării datelor citate în text;
- în tabel nu se admit rubrici necompletate; acestea se marchează prin semne speciale, care se stabilesc convențional.

Să ne reamintim

Utilizarea unei diagrame de structură radială este recomandată atunci când:

- Nu aveți decât o serie de date pe care doriți să o reprezentați.
- Nici una dintre valorile pe care doriți să le reprezentați nu este negativă.
- Aproape nici una dintre valorile pe care doriți să le reprezentați nu este valoare zero.
- Nu aveți mai mult de șapte categorii.
- Categoriile reprezintă părți din structura radială întreagă.

Datele care sunt aranjate în coloane sau rânduri într-o singură foaie de lucru pot fi reprezentate, de asemenea, și într-o diagramă de structură inelară. Asemenea unei diagrame de structură radială, o diagramă de structură inelară arată relația părților cu întregul, dar poate conține mai multe serii de date.



3.3. Rezumat

Sistematizarea și prezentarea datelor statistice reprezintă etape ale procesului de prelucrare statistică și vizează obținerea unei imagini a fenomenului studiat, într-o manieră în care mai ușor de analizat și interpretat.

În vederea aplicării metodelor de calcul și de interpretare statistică, rezultatele sistematizării datelor se prezintă sub formă de tabele, grafice sau serii statistice.



3.4. Test de autoevaluare a cunoștințelor

1) Sistematizarea presupune ordonarea datelor după anumite caracteristici, iar rezultatul procesului de sistematizare se prezintă sub forma de:

- a. serii de date;
- b. funcții statistice;
- c. tabele;
- d. grafice;
- e. a)+c)+d)

Alegeți varianta corectă.

2) Centralizarea datelor reprezintă:

- a. înlănțuirea logică a etapelor sistematizării datelor statistice;
- b. centralizare pe grupe;
- c. agregarea valorilor individuale ale caracteristicilor pentru toate unitățile colectivității care permit însumarea din punctul de vedere al conținutului în scopul obținerii indicatorilor totalizatori;
- d. procesul prin care se obțin informații privind colectivitatea statistică;
- e. obținerea unor informații specifice.

Alegeți varianta corectă.

3) Utilizarea unei diagrame de structură radială este recomandată atunci când:

- a. Nu aveți decât o serie de date pe care doriți să o reprezentați.
- b. Nici una dintre valorile pe care doriți să le reprezentați nu este negativă.
- c. Aproape nici una dintre valorile pe care doriți să le reprezentați nu este valoare zero.
- d. Nu aveți mai mult de șapte categorii.
- e. Categoriile reprezintă părți din structura radială întreagă.
- f. Toate variantele de mai sus.

Ce variantă nu este corectă?



3.5. Bibliografie recomandată

- a. Caragea Nicoleta, *Statistica- Concepte si metode de analiza a datelor*, Editura MUSTANG, București, 2015
- b. Voineagu, V. , *Statistică –baze teoretice și aplicații*, Editura Economică, București, 2007
- c. Georgescu-Roegen N., *Metoda statistică*, Editura Expert, București, 1998.
- d. Andrei T., Stancu S., Pele D.T., *Statistică. Teorie și aplicații* (ediția a II-a), Editura Economică, București, 2003.
- e. Biji M., Biji E.M., Lilea E., Anghelache C., *Tratat de statistică*, Editura Economică, București, 2002.

Unitatea de învățare 4.

SERII DE DISTRIBUȚIE



4.1. Introducere

În cadrul acestui capitol, studenților le sunt prezentate metodele și procedeele de construire a seriilor de distribuție.

Durata medie de parcurgere a acestui capitol este de aproximativ o oră.



4.2. Conținut

4.2.1. *Prezentare generală*

Seriile de distribuție, denumite și repartiții empirice, sintetizează variația uneia sau mai multor caracteristici statistice în cadrul unei colectivități. Caracteristicile în funcție de care se construiesc seriile de distribuție sunt cantitative sau calitative. Există, de exemplu, serii de distribuție a salariaților după nivelul salariului (caracteristică de tip cantitativ), precum și după ramura de activitate, după ocupație sau statut profesional (caracteristici calitative).

4.2.2. *Tipuri de serii de distribuție*

Seriile de distribuție pot fi unidimensionale și multidimensionale.

Seriile unidimensionale reflectă variația unei singure caracteristici în cadrul colectivității studiate. Seriile multidimensionale reflectă variația mai multor caracteristici, ale căror variante/valori sunt înregistrate pentru fiecare unitate a colectivității. Astfel, analiza salariilor se poate realiza pe baza distribuției salariaților după o singură caracteristică (nivelul salariului) sau pe baza distribuției acestora după mai multe caracteristici (nivelul salariului, vechimea în activitate, nivelul de instruire, ocupație și altele).

Obiectul acestui capitol îl constituie seriile de distribuție unidimensionale, constituite pe baza caracteristicilor cantitative.

Seriile de distribuție sunt denumite și serii de frecvențe, întrucât cuprind frecvența cu care valorile/varianțele caracteristicii apar în colectivitatea studiată, respectiv numărul de unități cuprinse în fiecare grupă care formează seria. Seriile de frecvențe constituite pe baza caracteristicilor cantitative sunt de două feluri, în funcție de modul în care variază caracteristica de grupare:

- serii de distribuție pe variante, în cazul caracteristicilor discrete, care înregistrează puține valori;
- serii de distribuție pe intervale, în cazul caracteristicilor continue sau al celor discrete care înregistrează foarte multe valori, apropiindu-se astfel de cele continue.

Seriile de frecvențe au două componente.

- În cazul distribuției pe variante, prima componentă o constituie valorile caracteristicii (x_i), iar a doua frecvențele corespunzătoare valorilor (n_i):

Valorile/varianțele caracteristicii	Numărul unităților/frecvențele
x_i	n_i
x_1	n_1
x_2	n_2
...	...
x_k	n_k
Total	n

4.2.3. Construcția unei serii de distribuție

Construcția unei serii de frecvențe unidimensionale constă în gruparea unităților colectivității în funcție de o caracteristică.

Etapele grupării

Gruparea pe baza unei caracteristici cantitative quasi-continue (care înregistrează un număr mare de valori) presupune stabilirea:

- numărului de grupe,
- a mărimii intervalului de grupare, adică a înălțimii fiecărei grupe,

- și a limitelor intervalelor de grupare, adică a valorilor care delimitează fiecare grupă,
- numărarea unităților care intră în fiecare grupă, adică la care înregistrează valori cuprinse în fiecare dintre intervalele de grupare.

Stabilirea numărului de grupe

Una dintre metodele utilizate pentru determinarea numărului de grupe constă în aproximarea acestuia pe baza formulei:

$$g = \sqrt[3]{2 \times n}$$

unde g - reprezintă numărul aproximativ de grupe și n - numărul observațiilor (al unităților colectivității).

Numărul de grupe va fi egal cu numărul întreg care depășește valoarea rezultată din calcul, ceea ce înseamnă că numărul de grupe se stabilește prin rotunjirea *în plus* a rezultatului. Dacă numărul observațiilor este egal cu 70, numărul de grupe poate fi egal cu 6:

$$g = \sqrt[3]{2 \times n} = \sqrt[3]{2 \times 70} = 5,19 \cong 6$$

La un număr de 160 de observații, numărul de grupe poate fi stabilit la 7, iar la 800 de observații - la 12:

$$g = \sqrt[3]{2 \times 160} = 6,84 \cong 7$$

$$g = \sqrt[3]{2 \times 800} = 11,70 \cong 12$$

Conform acestui algoritm, numărul de grupe care corespunde unui anumit număr de observații poate fi determinat pe baza tabelului următor:

Corespondența între numărul de observații și numărul de grupe

<i>Numărul de observații/unități</i>	<i>Numărul de grupe</i>
<i>n</i>	<i>g</i>
14 – 32	4
33 – 62	5
63 – 108	6
109 – 171	7
172 – 255	8
256 – 364	9
365 – 500	10
501 – 665	11
666 – 864	12

Mărimea intervalului de grupare

Mărimea intervalului de grupare (înălțimea grupelor) poate fi aproximată împărțind diferența dintre valoarea maximă și cea minimă la numărul grupelor, conform relației de calcul:

$$h = \frac{X_{\max} - X_{\min}}{g}$$

Mărimea efectivă a intervalului se stabilește prin rotunjirea în plus a rezultatului raportului.

De exemplu, dacă cercetarea se referă la cifra de afaceri și s-a înregistrat o valoare maximă de 2850 mii lei și una minimă de 1290 mii lei, iar numărul grupelor a fost stabilit la 5 (corespunzător unui număr de 50 unități observate), mărimea intervalului de grupare poate fi stabilită la 315 sau 320 mii lei:

$$h = \frac{X_{\max} - X_{\min}}{g} = \frac{2850 - 1290}{5} = 312 \text{ mii lei.}$$

Acest algoritm presupune efectuarea unei grupări cu *intervale egale*. În cazul în care unele intervale nu cuprind valorile nici unei unități sau dacă numărul de unități este insuficient pentru a fi semnificativ statistic, se pot comasa două sau mai multe intervale. Rezultatul va fi o grupare cu *intervale inegale*, așa cum sunt cele mai multe grupări care stau la baza formării distribuțiilor din domeniul economiei.

Limitele intervalelor

Specificarea limitelor intervalelor de grupare trebuie astfel realizată încât să fie îndeplinite următoarele două condiții:

- intervalele constituite să acopere toate valorile rezultate din observare;
- o anumită valoare să poată fi atribuită unei singure grupe.

Practic, limita inferioară a primului interval (intervalul corespunzător celor mai mici valori) poate fi stabilită adoptând o valoare arbitrară, egală sau mai mică decât valoarea minimă, dar nu mult mai mică decât aceasta. Această limită poate fi determinată pe baza formulei:

$$x_1^{\inf} = \frac{(g \times h) - (x_{\max} - x_{\min})}{2}$$

Limita superioară a primului interval se calculează însumând înălțimea intervalului la limita sa inferioară:

$$x_1^{\sup} = x_1^{\inf} + h$$

Limita inferioară a intervalului următor este egală cu limita superioară a primului interval, iar limita superioară se determină adăugând la aceasta înălțimea intervalului.

În mod similar se stabilesc și limitele celorlalte intervale.

Astfel:

$$x_2^{\text{inf}} = x_1^{\text{sup}} \text{ și } x_2^{\text{sup}} = x_2^{\text{inf}} + h$$

$$x_3^{\text{inf}} = x_2^{\text{sup}} \text{ și } x_3^{\text{sup}} = x_3^{\text{inf}} + h$$

.....

$$x_k^{\text{inf}} = x_{k-1}^{\text{sup}} \text{ și } x_k^{\text{sup}} = x_k^{\text{inf}} + h$$

Este posibil însă ca unele unități statistice să înregistreze valori egale cu limitele superioare ale intervalelor (și implicit cu limitele inferioare ale următoarelor intervale), ceea ce ar putea determina includerea lor în ambele grupe. Pentru evitarea riscului unor asemenea duble înregistrări, se practică diferențierea limitei inferioare a unui interval de limita superioară a intervalului precedent, prin adăugarea unei unități (sau a unei fracțiuni - zecimi, sutimi etc., în funcție de ordinul de mărime a valorilor caracteristicii) la limita inferioară a intervalului sau prin scăderea unei unități (fracțiuni) din limita superioară a intervalului precedent. În aceste condiții, limita inferioară a intervalului i se va calcula astfel:

$$x_i^{\text{inf}} = x_{i-1}^{\text{sup}} + 1$$

și limita superioară se va calcula după relația:

$$x_i^{\text{sup}} = x_{i-1}^{\text{sup}} + h$$

Intervalele pentru care sunt precizate ambele limite se numesc *intervale închise*. În practica grupării se obișnuiește însă ca limita inferioară a primului interval și limita superioară a ultimului interval să nu fie precizate. Această soluție este determinată de necesitatea comasării unora dintre intervalele aflate în cele două extreme ale distribuției, având în vedere că în general frecvențele sunt foarte scăzute la extremele distribuției. Cele două intervale extreme sunt astfel intervale deschise.

Exemplu: A fost efectuată o cercetare privind mărimea (măsurată pe baza numărului de salariați) a 80 de firme industriale din orașul M. Datele referitoare la numărul de salariați înregistrat în cursul observării sunt următoarele:

166	162	121	126	128	85	136	158	135	127
-----	-----	-----	-----	-----	----	-----	-----	-----	-----

141	142	92	148	177	80	156	188	205	144
155	230	100	129	160	159	105	150	110	98
182	102	128	198	115	122	124	163	130	133
132	150	75	206	149	170	112	142	119	151
134	224	135	236	126	175	215	130	121	128
190	156	108	143	218	172	180	120	169	129
123	156	142	127	133	146	139	140	138	138

Analiza mărimii întreprinderilor necesită gruparea acestora după numărul de salariați.

Numărul de grupe este egal cu 6:

$$g = \sqrt[3]{2 \times 80} = 5,43$$

iar mărimea intervalului de grupare este 30 salariați:

$$h = \frac{236 - 75}{6} = 26,83$$

Limita inferioară a primului interval poate fi stabilită pornind de la:

$$x_1^{\text{inf}} = 75 - \frac{6 \times 30 - (236 - 75)}{2} = 65,5$$

Limita inferioară poate fi stabilită la 65 salariați, la 60 sau la 70 salariați. Dacă este stabilită la 70 salariați, limita superioară va fi egală cu:

$$x_1^{\text{sup}} = 70 + 30 = 100$$

iar limitele celorlalte grupe vor fi:

$$x_2^{\text{inf}} = 100 \quad x_2^{\text{sup}} = 100 + 30 = 130$$

$$x_3^{\text{inf}} = 130 \quad x_3^{\text{sup}} = 130 + 30 = 160$$

$$x_4^{\text{inf}} = 160 \quad x_4^{\text{sup}} = 160 + 30 = 190$$

$$x_5^{\text{inf}} = 190 \quad x_5^{\text{sup}} = 190 + 30 = 220$$

$$x_6^{\text{inf}} = 220 \quad x_6^{\text{sup}} = 220 + 30 = 250$$

Dacă se are în vedere și necesitatea diferențierii limitelor inferioare de limitele superioare ale intervalelor precedente, precum și numărul unităților la care numărul de salariați se înscrie în intervalele astfel constituite, seria de frecvențe va înregistra următoarele valori:

Distribuția firmelor după numărul de salariați

<i>Grupe după numărul de salariați</i>	<i>Numărul firmelor</i>
70 – 100	6
101 – 130	24
131 – 160	30
161 – 190	12
191 – 220	5
221 - 250	3

Frecvențele absolute se determină prin numărarea firmelor care au un număr de salariați corespunzător grupelor stabilite.

Să ne reamintim

- Seriile de distribuție, denumite și repartiții empirice, sintetizează variația uneia sau mai multor caracteristici statistice în cadrul unei colectivități. Caracteristicile în funcție de care se construiesc seriile de distribuție sunt cantitative sau calitative. Există, de exemplu, serii de distribuție a salariaților după nivelul salariului (caracteristică de tip cantitativ), precum și după ramura de activitate, după ocupație sau statut profesional (caracteristici calitative).
- Seriile de distribuție pot fi unidimensionale și multidimensionale.



4.3. Rezumat

O serie de distribuție se mai numește și serie de frecvențe. Seriile de distribuție pot fi: pe variante sau pe intervale. Construcția unei serii de frecvențe unidimensionale constă în gruparea unităților colectivității în funcție de o caracteristică.



4.4. Test de autoevaluare a cunoștințelor

1) Serii de distribuție sunt denumite și serii de frecvențe, deoarece:

- a. sunt determinate pe baza frecvențelor cumulate;
- b. cuprind frecvența cu care valorile/varianțele caracteristicii apar în colectivitatea studiată,
- c. cuprind numărul de unități cuprinse în fiecare grupă care formează seria
- d. b+c)
- e. rezultă din calcule de frecvență.

Alegeți varianta corectă.

2) Elementele componente ale unei serii de distribuție pe variante sunt următoarele:

- a. valorile cuprinse în fiecare grupă (x_i) și frecvențele corespunzătoare valorilor (n_i);
- b. frecvențele corespunzătoare valorilor (x_i) și valorile caracteristicii (n_i);
- c. frecvențele relative și frecvențele absolute;
- d. valorile caracteristicii (x_i) și frecvențele corespunzătoare valorilor (n_i);
- e. frecvențele cumulate crescător și frecvențele cumulate descrescător.

Alegeți variant corectă.

3) În cazul seriilor de distribuție pe intervale, elementele componente sunt următoarele:

- a. prima componentă cuprinde intervalele de grupare, iar a doua - frecvențele corespunzătoare grupelor;
- b. valorile cuprinse în fiecare grupă (x_i) și frecvențele corespunzătoare valorilor (n_i);
- c. frecvențele corespunzătoare valorilor (x_i) și valorile caracteristicii (n_i);
- d. valorile caracteristicii (x_i) și frecvențele corespunzătoare valorilor (n_i);
- e. frecvențele relative și frecvențele absolute;

Alegeți variant corectă.

4) Fie distribuția studenților după notele la examenul de statistică:

Note x_i	Numărul studenților	n_i
6	5	
7	8	
8	14	
9	8	
10	5	
Total	40	

Care este valoarea frecvenței relative corespunzătoare notei 8:

- a. 40%
- b. 35%
- c. 59%
- d. 44%
- e. Nu se poate determina.

Alegeți varianta corectă.

5) Se cunosc indicatorii frecvenței pentru distribuția firmelor după numărul de salariați

Grupe de firme după numărul de salariați	n_i	n'_i (%)	$N_i^{\wedge}$	$N_i'^{\wedge}$ (%)	N_i	N_i' (%)
Până la 50	5	3,57	5	3,57	140	100,00
51 - 100	12	8,57	17	12,14	135	96,43
101 - 150	35	25,00	52	37,14	123	87,86
151 - 200	28	20,00	80	57,14	88	62,86
201 - 250	30	21,43	110	78,57	60	42,86
251 - 300	24	17,14	134	95,71	30	21,43
Peste 300	6	4,29	140	100,00	6	4,29
Total	140	100	-	-	-	-

Frecvența cumulată crescător, exprimată sub formă de coeficient, corespunzătoare celui de-al treilea interval este egală cu:

- a. 35 persoane
- b. 25%
- c. 52%
- d. 37,14%
- e. Nu se poate determina.

Alegeți varianta corectă.

6) Conform Anuarului Statistic al României, distribuția cercetătorilor din România, după vârstă se prezintă astfel:

Grupe de vârstă	Număr de cercetători n_i
până la 25 ani	753
25-34 ani	5700
35-44 ani	6455
45-54 ani	8080
55-64 ani	4206
65 ani și peste	774

Frecvența cumulată crescător, corespunzătoare ultimului interval este de:

- a. 25968 persoane
- b. 45678 persoane
- c. 234589 persoane
- d. 34890 persoane
- e. 774 persoane.

Alegeți varianta corectă.



4.5. Bibliografie recomandată

- a. Caragea Nicoleta, *Statistica- Concepte si metode de analiza a datelor*, Editura MUSTANG, București, 2015
- b. Voineagu, V. , *Statistică –baze teoretice și aplicații*, Editura Economică, București, 2007
- c. Georgescu-Roegen N., *Metoda statistică*, Editura Expert, București, 1998.
- d. Andrei T., Stancu S., Pele D.T., *Statistică. Teorie și aplicații* (ediția a II-a), Editura Economică, București, 2003.
- e. Biji M., Biji E.M., Lilea E., Anghelache C., *Tratat de statistică*, Editura Economică, București, 2002.

Unitatea de învățare 5.

INDICATORII FRECVENȚEI



5.1. Introducere

În cadrul acestui capitol, studenților le sunt prezentate metodele și procedeele de construire a seriilor de distribuție.

Durata medie de parcurgere a acestui capitol este de aproximativ o oră.



5.2. Conținut

5.2.1. Frecvențe absolute și relative

Frecvența caracterizează structura unei colectivități/populații în raport cu o anumită caracteristică de grupare. Seriile de distribuție sunt construite cu frecvențe absolute sau relative.

Frecvențele absolute desemnează numărul de unități la care se înregistrează același nivel al caracteristicii sau niveluri cuprinse într-un interval/grupă.

Suma frecvențelor absolute este egală cu numărul unităților care compun colectivitatea analizată:

$$\sum_{i=1}^n n_i = n$$

Distribuția de frecvențe absolute este o formă de rezumare tabelară a datelor statistice pentru construcția căreia stabilim un număr de variante și apoi înscriem pentru fiecare dintre ele frecvențele absolute corespunzătoare, pe baza datelor statistice de care dispunem.

Frecvențele relative (denumite și ponderi) măsoară partea pe care o dețin în volumul total al colectivității unitățile la care se înregistrează același nivel al caracteristicii sau niveluri cuprinse într-un interval/grupă.

Frecvențele relative sunt exprimate procentual și se calculează prin raportarea frecvențelor absolute aferente fiecărui interval sau fiecărei variante/valori la suma frecvențelor absolute:

$$n'_i = \frac{n_i}{\sum_{i=1}^n n_i} \times 100$$

Suma frecvențelor relative este egală cu 100%:

$$\sum_{i=1}^n n'_i = 100\%$$

În calcule se utilizează și frecvențele relative exprimate sub formă de coeficienți:

$$n''_i = \frac{n_i}{\sum_{i=1}^n n_i}$$

a căror sumă este egală cu unitatea:

$$\sum_{i=1}^n n''_i = 1.$$

Distribuția de frecvențe relative este o formă de rezumare a datelor statistice care se obține prin plasarea într-un tabel a variantelor și a frecvențelor relative corespunzătoare.

5.2.2. Frecvențe cumulate

Un alt indicator important pentru caracterizarea structurii îl reprezintă *frecvențele cumulate*. Acestea se calculează prin însumarea succesivă a frecvențelor grupelor:

$$N_i^{\wedge} = \sum_{i=1}^n n_i$$

Frecvențele cumulate pornind de la frecvențele aferente valorilor mici ale variabilei, se numesc *frecvențe cumulate crescător* ($N_i^{\wedge}$). Ele evidențiază numărul unităților care înregistrează valori ale caracteristicii mai mici sau egale cu limita superioară a grupei i (x_i^{sup} , în cazul grupării pe intervale), respectiv mai mici sau egale cu valoarea x_i (în cazul grupării pe variante/valori).

Astfel, prima frecvență cumulată crescător va fi întodeauna egală cu prima frecvență (absolută sau relativă):

$$N_1^{\wedge} = n_1$$

Următoarele $k-1$ frecvențe cumulate se vor calcula astfel:

$$N_2^{\wedge} = n_1 + n_2$$

$$N_3^{\wedge} = n_1 + n_2 + n_3$$

.....

$$N_k^{\wedge} = n_1 + n_2 + n_3 + \dots + n_k$$

Se poate observa faptul că ultima frecvență cumulată crescător este egală cu suma unităților colectivității:

$$N_k = n$$

Frecvențele cumulate descrescător se calculează prin însumarea frecvenței grupe i cu frecvențele tuturor grupelor cu valori superioare. Calculul se poate face pornind de la frecvențele corespunzătoare valorilor mari ale variabilei. În acest caz, frecvența cumulată descrescător, corespunzătoare primului interval/grupă, este egală cu suma unităților colectivității:

$$N_1 = n$$

Frecvențele cumulate (crescător sau descrescător) se calculează și pentru frecvențele relative.

Exemplu: În seria de distribuție a firmelor după numărul de salariați, frecvența absolută a primei grupe este egală cu 18 ($n_1 = 18$ firme). Frecvența relativă a acestei grupe este:

$$n_1' = \frac{18}{140} \times 100 = 12,9\% \text{ respectiv } n_1'' = \frac{18}{140} = 0,1286.$$

Indicatorii de frecvență ai distribuției firmelor după numărul de salariați

<i>Grupe după numărul de salariați</i> $x_i^{\text{inf}} - x_i^{\text{sup}}$	<i>Frecvențe absolute</i> <i>(numărul firmelor)</i> n_i	<i>Frecvențe relative</i> <i>(pondere în total, %)</i> n_i'	<i>Frecvențe relative</i> n_i''
Până la 100	18	12,9	0,129
101 - 120	30	21,4	0,214
121 - 140	35	25,0	0,250
141 - 160	24	17,2	0,172
161 - 180	16	11,4	0,114
181 - 200	10	7,1	0,071
Peste 200	7	5,0	0,050
Total	140	100,0	1,000

În cazul distribuției firmelor după numărul de salariați, frecvențele absolute cumulate crescător se calculează în felul următor:

$$N_1 = 18,$$

$$N_2 = 18 + 30 = 48,$$

$$N_3 = 18 + 30 + 35 = 83,$$

$$N_4 = 18 + 30 + 35 + 24 = 107,$$

$$N_5 = 18 + 30 + 35 + 24 + 16 = 123,$$

$$N_6 = 18 + 30 + 35 + 24 + 16 + 10 = 133,$$

$$N_7 = 18 + 30 + 35 + 24 + 16 + 10 + 7 = 140,$$

iar cele cumulate descrescător:

$$N_7 = 7,$$

$$N_6 = 7 + 10 = 17,$$

$$N_5 = 7 + 10 + 16 = 33,$$

$$N_4 = 7 + 10 + 16 + 24 = 57,$$

$$N_3 = 7 + 10 + 16 + 24 + 35 = 92,$$

$$N_2 = 7 + 10 + 16 + 24 + 35 + 30 = 122,$$

$$N_1 = 7 + 10 + 16 + 24 + 35 + 30 + 18 = 140.$$

Frecvențele relative cumulate crescător și descrescător se calculează conform aceluiași algoritm:

$$N_1' = 12,9$$

$$N_2' = 12,9 + 21,4 = 34,3$$

$$N_3' = 12,9 + 21,4 + 25,0 = 59,3 \text{ ș.a.m.d.}$$

Frecvențele cumulate ale distribuției firmelor după numărul de salariați

<i>Grupe după numărul de salariați</i>	<i>Frecvențe absolute cumulate</i>		<i>Frecvențe relative cumulate (%)</i>	
	<i>crescător</i>	<i>descrescător</i>	<i>crescător</i>	<i>descrescător</i>
$x_i^{\text{inf}} - x_i^{\text{sup}}$	$N_i \uparrow$	$N_i \downarrow$	$N_i' \uparrow$	$N_i' \downarrow$
Până la 100	18	140	12,9	100,0
101 - 120	48	122	34,3	87,1
121 - 140	83	92	59,3	65,7
141 - 160	107	57	76,5	40,7

161 – 180	123	33	87,9	23,5
181 – 200	133	17	95,0	12,1
Peste 200	140	7	100,0	5,0

Să ne reamintim

- *Frecvențele absolute* desemnează numărul de unități la care se înregistrează același nivel al caracteristicii sau niveluri cuprinse într-un interval/grupă.
- *Frecvențele relative* (denumite și ponderi) măsoară partea pe care o dețin în volumul total al colectivității unitățile la care se înregistrează același nivel al caracteristicii sau niveluri cuprinse într-un interval/grupă.
- Un alt indicator important pentru caracterizarea structurii îl reprezintă *frecvențele cumulate*. Acestea se calculează prin însumarea succesivă a frecvențelor grupelor.



5.3. Rezumat

Indicatorii frecvenței au o importanță deosebită pentru analiza seriilor de distribuție. Frecvența caracterizează structura unei colectivități/populații în raport cu o anumită caracteristică de grupare. Seriile de distribuție sunt construite cu frecvențe absolute sau relative.



5.4. Test de autoevaluare a cunoștințelor

1) Fie distribuția studenților după notele la examenul de statistică:

Note x_i	Numărul studenților n_i
6	5
7	8
8	14

9	8
10	5
Total	40

Care este valoarea frecvenței relative corespunzătoare notei 8:

- a. 40%
- b. 35%
- c. 59%
- d. 44%
- e. Nu se poate determina.

Alegeți varianta corectă.

2) Se cunosc indicatorii frecvenței pentru distribuția firmelor după numărul de salariați

Grupe de firme după numărul de salariați	n_i	n'_i (%)	$N_i^{\wedge}$	$N_i'^{\wedge}$ (%)	N_i	N_i' (%)
Până la 50	5	3,57	5	3,57	140	100,00
51 - 100	12	8,57	17	12,14	135	96,43
101 - 150	35	25,00	52	37,14	123	87,86
151 - 200	28	20,00	80	57,14	88	62,86
201 - 250	30	21,43	110	78,57	60	42,86
251 - 300	24	17,14	134	95,71	30	21,43
Peste 300	6	4,29	140	100,00	6	4,29
Total	140	100	-	-	-	-

Frecvența cumulată crescător, exprimată sub formă de coeficient, corespunzătoare celui de-al treilea interval este egală cu:

- a. 35 persoane
- b. 25%
- c. 52%
- d. 37,14%
- e. Nu se poate determina.

Alegeți varianta corectă.

3) Se cunosc frecvențele relative pentru distribuția firmelor după cifra de afaceri:

Grupe de firme după cifra de	n'_i
------------------------------	--------

afaceri (mii lei)	(%)
Până la 100	13,57
101 - 150	15,00
151 - 200	10,00
201 - 250	31,43
251 - 300	17,14
Peste 300	4,29
Total	100

Frecvența absolută, corespunzătoare celui de-al doilea interval este egală cu:

- a. 35 persoane
- b. 25%
- c. 52%
- d. 37,14%
- e. Nu se poate determina.

Alegeți varianta corectă.

- 4) Conform Anuarului Statistic al României, distribuția cercetătorilor din România, după vârstă se prezintă astfel:

Grupe de vârstă	Număr de cercetători n_i
până la 25 ani	753
25-34 ani	5700
35-44 ani	6455
45-54 ani	8080
55-64 ani	4206
65 ani și peste	774

Frecvența cumulată crescător, corespunzătoare ultimului interval este de:

- a. 25968 persoane
- b. 45678 persoane
- c. 234589 persoane
- d. 34890 persoane
- e. 774 persoane.

Alegeți varianta corectă.

- 5) Conform Anuarului Statistic al României, investițiile brute ale unităților locale active din industrie, construcții, comerț și alte servicii, pe clase de mărime se prezintă astfel:

Grupe de unități după mărime (mil lei prețuri curente)	Număr de unități n_i
0-9	40260
10-49	27064
50-249	28672
250 și peste	45401
Total	

Frecvența relativă, exprimată sub formă de coeficient, corespunzătoare intervalului 10-49 mil. lei prețuri curente este de:

- a. 27064
- b. 100
- c. 12,8
- d. 10-49
- e. 0,128

Alegeți varianta corectă.

- 6) Frecvența cumulată descrescător, corespunzătoare primului interval/grupă, este egală cu:

- a. suma frecvențelor înregistrate de primele k grupe din serie;
- b. suma unităților colectivității;
- c. frecvența primului interval/grupă;
- d. frecvența intervalelor deschise;
- e. suma frecvențelor înregistrate de ultimele k-1 grupe din serie.

Alegeți varianta corectă.



5.5. Bibliografie recomandată

- a. Caragea Nicoleta, *Statistica- Concepte si metode de analiza a datelor*, Editura MUSTANG, București, 2015
- b. Voineagu, V. , *Statistică –baze teoretice și aplicații*, Editura Economică, București, 2007

- c. Georgescu-Roegen N., *Metoda statistică*, Editura Expert, București, 1998.
- d. Andrei T., Stancu S., Pele D.T., *Statistică. Teorie și aplicații* (ediția a II-a), Editura Economică, București, 2003.
- e. Biji M., Biji E.M., Lilea E., Anghelache C., *Tratat de statistică*, Editura Economică, București, 2002.

5.6. Teme de control

Realizați un referat de maximum 5 pagini în care să tratați sursele de date și sistematizarea datelor statistice, utilizând datele și metadatele din baza de date TEMPO a Institutului Național de Statistică.

- Fiecare student își va alege o serie de date de frecvențe din baza de date TEMPO
- Stabilește sursa de date
- Definește metadatele
- Construiește un grafic al seriei de distribuție.

Unitatea de învățare 6.

INDICATORII TENDINȚEI CENTRALE. MEDIA



6.1. Introducere

În cadrul acestui capitol, studenților le sunt prezentate cunoștințe generale privind indicatorii tendinței centrale. Detalierea modului de calcul va fi prezentată pentru fiecare indicator în parte, pe parcursul mai multor capitole.

Durata medie de parcurgere a acestui capitol este de aproximativ o oră.



6.2. Conținut

Seriile de distribuție construite pe baza datelor obținute în urma observării statistice oferă informații referitoare la structura fenomenului economic sau social care formează obiectul analizei statistice. Aceasta necesită însă și utilizarea unor indicatori cu caracter descriptiv, mențiți să asigure descrierea numerică sintetică a diferitelor aspecte ale fenomenului analizat. Indicatorii rezultă în urma unor operațiuni relativ simple de prelucrare a datelor înregistrate în cursul observării statistice sau a celor care formează seriile de distribuție.

În analiza seriilor de distribuție se utilizează următoarele grupe de indicatori:

- indicatori de nivel sau indicatori de poziție;
- indicatori ai variației sau dispersiei;
- indicatori ai formei de repartiție;
- indicatori ai concentrării.

Fiecare dintre indicatorii cuprinși în aceste grupe pune în evidență o anumită caracteristică a distribuției. Rolul lor în analiza seriilor de distribuție este divers, ansamblul lor servind la caracterizarea complexă a fenomenului evidențiat de seria de distribuție. Indicatorii numerici descriptivi oferă valori precise și determinate în mod obiectiv, valori care pot fi ușor folosite, interpretate și comparate una cu alta. Pe scurt, aceștia permit o analiză mai atentă a datelor, față de impresia generală pe care o oferă prezentarea datelor sub formă de serii, tabele și grafice.

Indicatorii cu ajutorul cărora se analizează tendința centrală a fenomenelor de masă sintetizează valorile individuale într-un singur nivel, reprezentativ pentru întreaga colectivitate.

În categoria indicatorilor de nivel, denumiți și indicatori de poziție, sunt cuprinși indicatorii tendinței centrale și cuantilele.

Indicatorii cu ajutorul cărora se analizează tendința centrală a fenomenelor de masă sintetizează valorile individuale într-un singur nivel, reprezentativ pentru întreaga colectivitate. Principalii indicatori ai tendinței centrale sunt: media, mediana și modul.

6.2.1. Media

Media este principalul indicator al tendinței centrale, cel mai frecvent utilizat în analiza statistică. Media este un indicator sintetic determinat pe baza tuturor datelor repartiției. Ea exprimă nivelul care ar fi fost înregistrat de caracteristică în cazul în care toate valorile ar fi fost egale și reprezintă toate variantele caracteristicii, putând să fie utilizată în evaluări în locul fiecăreia dintre ele.

Media statistică este o valoare ce sintetizează într-o singură expresie numerică toate valorile din serie.

Necesitățile diverse ale analizei statistice și natura diferită a caracteristicilor înregistrate au impus utilizarea practică a unor forme variate de medii, dintre care mai frecvent întâlnite sunt: aritmetică, armonică, geometrică, pătratică, cubică și parabolică; de asemenea, media cronologică. Mediile se calculează utilizând toate valorile individuale și se exprimă în unitatea de măsură a caracteristicii considerate.

Se consideră că media este reprezentativă dacă răspunde condițiilor formulate de statisticianul G.U. Yule⁴, sintetizate astfel:

- Media este definită obiectiv (precis definită, prin definiție proprie sau formulă de calcul);
- La calculul mediei se utilizează toate valorile individuale înregistrate;
- Media are o semnificație concretă, ușor de observat;
- Se caracterizează prin simplitatea și rapiditatea calculului;
- Se caracterizează prin stabilitate (să fie puțin sensibilă la fluctuațiile de eșantionare, în cazul utilizării mai multor eșantioane de același volum extrase din aceeași colectivitate);
- Media poate fi utilizată în calcule algebrice ulterioare.

Formula de calcul al mediei, ca și al altor indicatori de nivel și ai variației, diferă în funcție de informația disponibilă, de modul de organizare a acesteia.

⁴ George Udny Yule (1871-1951) – statistician englez

- Dacă informația disponibilă este sistematizată și prezentată sub forma unor serii simple formate dintr-un număr de valori egal cu numărul unităților care compun colectivitatea ($x_1, x_2, x_3, \dots, x_n$) sau date primare rezultate din observare, se calculează media simplă.
- Dacă informația disponibilă este sistematizată și prezentată sub forma unor serii de distribuție/de frecvențe, construite pe variante/valori sau pe intervale, se calculează media ponderată⁵.

Este important de subliniat că, atunci când sunt disponibile și datele primare (valorile aferente fiecărei unități din colectivitate) pe baza cărora a fost construită o serie de distribuție pe intervale, este de preferat calculul mediei simple celei ponderate, calculate pe baza seriei de frecvențe pe intervale. Construirea seriei de distribuție pe intervale presupune o anumită pierdere din informația asigurată de datele primare, iar calculul mediei (ca și a altor indicatori) pe baza distribuției pe intervale – un grad mai scăzut de precizie a rezultatului.

6.2.2. Media aritmetică

Media aritmetică este media cea mai cunoscută și mai utilizată. Pentru desemnarea ei se utilizează de obicei numai termenul de medie.

6.2.2.1. Media aritmetică simplă

Calculul mediei aritmetice simplă se realizează prin însumarea valorilor înregistrate de variabilă la toate unitățile caracteristicii și raportarea sumei la numărul de unități care compun colectivitatea studiată.

Formula de calcul pentru media aritmetică simplă este:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

6.2.2.2. Media aritmetică ponderată

Media aritmetică ponderată se calculează după o formulă care presupune însumarea produsului dintre valorile caracteristicii și frecvențele aferente și raportarea sumei la suma frecvențelor:

⁵ Media ponderată se utilizează și atunci când este necesar calculul mediei unor valori cărora li se atribuie în mod normativ, exogen, greutate specifică (importanță) diferită.

$$\bar{x} = \frac{\sum_{i=1}^k x_i n_i}{\sum_{i=1}^k n_i},$$

unde n_i reprezintă frecvențele absolute.

În cazul în care seria de distribuție cuprinde frecvențe relative, formula mediei aritmetice devine:

$$\bar{x} = \frac{\sum_{i=1}^k x_i n'_i}{100},$$

în cazul frecvențelor exprimate procentual, și

$$\bar{x} = \sum_{i=1}^k x_i n''_i,$$

în cazul frecvențelor exprimate sub formă de coeficienți.

Valorile x_i utilizate în calculul mediei unei serii de distribuție pe intervale sunt *centrele intervalelor de grupare*.

6.2.2.3. Centrele intervalelor

Calculul mediei seriilor de frecvențe presupune determinarea centrului fiecărui interval de grupare.

Algoritmul pentru determinarea centrelor intervalelor este următorul:

- I. Se calculează mărimea intervalelor de grupare;
- II. Se calculează centrul fiecărui interval de grupare.

Mărimea sau înălțimea unui interval se determină prin diferența dintre limita sa superioară și limita superioară a intervalului precedent:

$$h_i = x_i^{\text{sup}} - x_{i-1}^{\text{sup}}$$

pentru $i = 2, 3, \dots, k$.

Dacă limita inferioară a intervalului este egală cu limita superioară a intervalului precedent, $x_i^{\text{inf}} = x_{i-1}^{\text{sup}}$, atunci mărimea intervalului este egală cu diferența dintre limita sa superioară și cea inferioară:

$$h_i = x_i^{\text{sup}} - x_i^{\text{inf}}$$

În cazul intervalelor deschise, a căror mărime nu poate fi calculată după metoda prezentată mai sus, se utilizează mărimea intervalelor vecine, adoptându-se în mod convențional ipoteza că intervalele sunt egale.

Astfel:

$$h_1 = h_2$$

$$h_k = h_{k-1}$$

După determinarea mărimii tuturor intervalelor seriei de distribuție, se pot calcula centrele acestora. Centrul x_i al unui interval se determină ca semisumă a limitei superioare a intervalului respectiv și a limitei superioare a intervalului precedent:

$$x_i = \frac{x_i^{\text{sup}} + x_{i-1}^{\text{sup}}}{2}$$

pentru $i = 2, 3, \dots, k$.

În cazul în care limita inferioară a intervalului este egală cu limita superioară a intervalului precedent, $x_i^{\text{inf}} = x_{i-1}^{\text{sup}}$, centrul intervalului se poate calcula ca semisumă a limitelor intervalului:

$$x_i = \frac{x_i^{\text{inf}} + x_i^{\text{sup}}}{2}$$

În cazul intervalelor deschise, centrul celor două intervale extreme poate fi determinat astfel:

- Centrul primului interval

$$x_1 = x_1^{\text{sup}} - \frac{h_1}{2}$$

- Centrul ultimului interval

$$x_k = x_{k-1}^{\text{sup}} + \frac{h_k}{2}$$

Sau, dacă $x_i^{\text{inf}} = x_{i-1}^{\text{sup}}$

atunci:

$$x_k = x_k^{\text{inf}} + \frac{h_k}{2}$$

Este important de subliniat că, atunci când sunt disponibile și datele primare (valorile aferente fiecărei unități din colectivitate) pe baza cărora a fost construită o serie de distribuție pe intervale, este de preferat calculul mediei simple celei ponderate, calculate pe baza seriei de frecvențe pe intervale. Construirea seriei de distribuție pe intervale presupune o anumită pierdere din informația asigurată de datele primare, iar calculul mediei (ca și a altor indicatori) pe baza distribuției pe intervale – un grad mai scăzut de precizie a rezultatului.



6.3. Rezumat

Media face parte din indicatorii tendinței centrale, fiind cel mai important indicator din statistica descriptivă.

Formula de calcul al mediei, ca și al altor indicatori de nivel și ai variației, diferă în funcție de informația disponibilă, de modul de organizare a acesteia.

- Dacă informația disponibilă este sistematizată și prezentată sub forma unor serii simple formate dintr-un număr de valori egal cu numărul unităților care compun colectivitatea ($x_1, x_2, x_3, \dots, x_n$) sau date primare rezultate din observare, se calculează media simplă.
- Dacă informația disponibilă este sistematizată și prezentată sub forma unor serii de distribuție/de frecvențe, construite pe variante/valori sau pe intervale, se calculează media ponderată.



6.4. Test de autoevaluare a cunoștințelor

1) Se consideră că media este reprezentativă dacă răspunde condițiilor formulate de statisticianul G.U. Yule:

- a. Media are o semnificație concretă, ușor de observat;
- b. Media se caracterizează prin simplitatea și rapiditatea calculului;
- c. Media este precis definită;
- d. La calculul mediei se utilizează numai o parte din valorile individuale înregistrate;
- e. Media poate fi utilizată în calcule algebrice ulterioare.

Care dintre variante nu este corectă?

2) Media aritmetică simplă este utilizată în cazul:

- a. Seriilor de distribuție pe variante;
- b. Seriilor de distribuție pe intervale;
- c. Seriilor simple;
- d. Atât în cazul seriilor simple cât și pentru seriile de distribuție;
- e. a)+b)+d).

Alegeți varianta corectă.

3) Se cunosc date privind numărul de salariați angajați în cele cinci departamente ale firmei X.

(persoane)

D1	D2	D3	D4	D5
23	32	6	67	22

Media numărului de salariați, calculată pe baza formulei $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ este egală cu:

- a. 25;
- b. 30;
- c. 32;
- d. 22;
- e. 40.

Alegeți varianta corectă.

4) Se cunosc date privind situația rezultatelor unei grupe de studenți la examenul de economie:

x_i (puncte de notă)	n_i (numărul studenților)
5	2
6	7
7	11
8	25
9	10
10	5
Total	60

Media obținută de cei 60 de studenți la examenul de economie, calculată pe baza formulei

$$\bar{x} = \frac{\sum_{i=1}^k x_i n_i}{\sum_{i=1}^k n_i} \text{ este egală cu:}$$

- a. 7,7;
- b. 7,8;
- c. 8,9;
- d. 9,3;
- e. 7,9.

Alegeți varianta corectă.

5) Conform estimărilor Băncii Mondiale, distribuția produsului intern brut pe locuitor în unele state membre ale Uniunii Europene se prezintă astfel :

	Regatul Unit	România	Slovenia	Spania	Ungaria
PIB/locuitor (dolari)	35,4	12,2	27,6	31,9	19,3
Numărul locuitorilor (mil. persoane)	61,60	21,39	2,12	46,06	10,07

Media numărului de salariați este egală cu:

- a. 29,4;
- b. 30,1;
- c. 35,4;
- d. 31,9;
- e. 29,3.

Alegeți varianta corectă.

6) Conform datelor publicate în Raportul PNUD privind dezvoltarea umană, distribuția țărilor după numărul utilizatorilor de Internet ce revin la 1000 locuitori se prezintă astfel:

Grupe după numărul de utilizatori Internet la 1000 locuitori	Numărul de țări
Până la 10	33
10 - 50	45
50 - 100	29
100 – 200	20
200 – 400	21
400 – 500	11
Peste 500	18
Total	177

Media utilizatorilor de internet la nivel mondial este egală cu:

- a. 156,4
- b. 88,9
- c. 45,0
- d. 176,9
- e. 123,7.

Alegeți varianta corectă.



6.5. Bibliografie recomandată

- a. Caragea Nicoleta, *Statistica- Concepte si metode de analiza a datelor*, Editura MUSTANG, București, 2015
- b. Voineagu, V. , *Statistică –baze teoretice și aplicații*, Editura Economică, București, 2007
- c. Georgescu-Roegen N., *Metoda statistică*, Editura Expert, București, 1998.
- d. Andrei T., Stancu S., Pele D.T., *Statistică. Teorie și aplicații* (ediția a II-a), Editura Economică, București, 2003.
- e. Biji M., Biji E.M., Lilea E., Anghelache C., *Tratat de statistică*, Editura Economică, București, 2002.

Unitatea de învățare 7.

ALTE TIPURI DE MEDIE. PROPRIETĂȚILE MEDIEI



7.1. Introducere

În cadrul acestui capitol, studenților le sunt prezentate toate tipurile de medii cu care operează statistica și proprietățile mediei.

Durata medie de parcurgere a acestui capitol este de aproximativ o oră.



7.2. Conținut

7.2.1. Media armonică

Media armonică este un tip de medie care se utilizează în cazuri speciale; de exemplu, la calculul productivității medii a muncii, atunci când aceasta este măsurată cu ajutorul consumului specific de muncă, sau la estimarea indicilor sintetici ai prețurilor, atunci când aceștia sunt calculați ca medie ponderată a indicilor individuali ai prețurilor (conform metodei de ponderare Laspeyres).

Media armonică simplă se calculează după o formulă care presupune inversarea valorilor variabilei:

$$\bar{x}_h = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

iar cea ponderată după formula:

$$\bar{x}_h = \frac{\sum_{i=1}^n n_i}{\sum_{i=1}^n \frac{1}{x_i} n_i}$$

Exemple:

1. Se cunosc notele obținute de către cinci studenți la examenul de statistică:
10, 5, 8, 6, 9

Aplicând, în scop demonstrativ⁶, formula mediei armonice simple la calculul mediei notelor rezultă:

$$\bar{x}_h = \frac{5}{\frac{1}{10} + \frac{1}{5} + \frac{1}{8} + \frac{1}{6} + \frac{1}{9}} = \frac{5}{0,703} = 7,11.$$

2. Distribuția celor 25 de studenți după notele obținute la examenul de statistică se prezintă astfel:

x_i (puncte de notă)	n_i (numărul de studenți)
6	2
7	5
8	7
9	10
10	1
Total	25

Media armonică a notelor va fi:

$$\bar{x}_h = \frac{2 + 5 + 7 + 10 + 1}{\frac{1}{6} \times 2 + \frac{1}{7} \times 5 + \frac{1}{8} \times 7 + \frac{1}{9} \times 10 + \frac{1}{10} \times 1} = \frac{25}{3,1337} = 7,98.$$

Media armonică este mai mică decât media aritmetică, dacă valorile caracteristicii sunt pozitive:

$$\bar{x}_h < \bar{x}.$$

7.2.2. Media pătratică

Media pătratică este, de asemenea, o medie utilizată în situații speciale, și anume atunci când se dorește să se acorde o importanță mare existenței unor valori mari, precum și atunci când media aritmetică nu poate fi calculată din cauză că valorile sunt pozitive și negative și se anulează reciproc, așa cum se întâmplă în cazul mediei abaterilor față de medie. Calculul mediei pătratice presupune utilizarea pătratelor valorilor.

Formula de calcul al mediei pătratice este:

⁶ Media notelor se calculează, în mod tradițional, cu ajutorul mediei aritmetice. Media armonică, cea pătratică și cea geometrică sunt, așa cum s-a arătat, medii care se aplică în cazuri speciale. Calculul mediei notelor cu ajutorul celor trei medii speciale are numai un scop demonstrativ, și anume pentru a exemplifica modul de aplicare a formulelor de calcul al acestor medii și a evidenția diferențele dintre rezultatele obținute prin calculul mai multor tipuri de medii pe baza aceluiași set de date.

- pentru media simplă

$$\bar{x}_p = \sqrt{\frac{\sum_{i=1}^n x_i^2}{n}}$$

- pentru media ponderată

$$\bar{x}_p = \sqrt{\frac{\sum_{i=1}^n x_i^2 n_i}{\sum_{i=1}^n n_i}}$$

Exemple:

1. Media pătratică a notelor celor cinci studenți din exemplul 1 se calculează după cum urmează:

$$\bar{x}_p = \sqrt{\frac{10^2 + 5^2 + 8^2 + 6^2 + 9^2}{5}} = \sqrt{\frac{306}{5}} = 7,82$$

2. Media pătratică a notelor celor 25 studenți din exemplul 2, astfel:

$$\bar{x}_p = \sqrt{\frac{6^2 \times 2 + 7^2 \times 5 + 8^2 \times 7 + 9^2 \times 10 + 10^2 \times 1}{2 + 5 + 7 + 10 + 1}} = \sqrt{\frac{1675}{25}} = 8,19.$$

Media pătratică este întotdeauna mai mare decât media aritmetică:

$$\bar{x}_p > \bar{x}$$

7.2.3. Media geometrică

Media geometrică reprezintă o medie utilizată atunci când termenii seriei nu se află în relație de aditivare, ci de multiplicare, așa cum se întâmplă în cazul seriilor de indici de creștere. Calculul mediei geometrice presupune utilizarea produsului valorilor seriei, conform formulelor:

- pentru media simplă

$$\bar{x}_g = \sqrt[n]{\prod_{i=1}^n x_i}$$

- pentru media ponderată

$$\bar{x}_g = \sqrt[n]{\prod_{i=1}^n x_i^{n_i}}$$

Exemple:

Aplicând, cu caracter demonstrativ, formulele de mai sus pe datele din exemplele considerate anterior, media geometrică simplă și ponderată se calculează în felul următor:

$$\bar{x}_g = \sqrt[5]{10 \times 5 \times 8 \times 6 \times 9} = \sqrt[5]{21600} = 7,36$$

$$\bar{x}_g = \sqrt[2+5+7+10+1]{6^2 \times 7^5 \times 8^7 \times 9^{10} \times 10^1} = 8,05.$$

7.2.4. Proprietățile mediei

1. Media aritmetică este cuprinsă între valoarea maximă și cea minimă a caracteristicii:

$$x_{\min} < \bar{x} < x_{\max}$$

2. Suma abaterilor valorilor individuale față de media lor este nulă, adică:

- pentru serii simple

$$\sum_{i=1}^k (x_i - \bar{x}) = 0$$

- pentru serii de frecvențe

$$\sum_{i=1}^k (x_i - \bar{x})n_i = 0$$

3. Media calculată pe baza variantelor caracteristicii mărite sau micșorate cu o constantă a se modifică în același sens și cu aceeași mărime față de media seriei inițiale:

$$\bar{x}' = \frac{\sum_{i=1}^n (x_i \pm a)}{n} = \frac{\sum_{i=1}^n x_i}{n} \pm a = \bar{x} \pm a$$

4. Media calculată din variantele caracteristicii înmulțite sau simplificate cu o constantă c se înmulțește sau se simplifică cu aceeași constantă:

$$\bar{x}'' = \frac{\sum_{i=1}^n cx_i}{n} = c\bar{x}$$

$$\bar{x}'' = \frac{\sum_{i=1}^n \frac{x_i}{c}}{n} = \frac{\bar{x}}{c}$$

Având în vedere proprietățile (3) și (4), media aritmetică simplă poate fi calculată după formula:

$$\bar{x} = \frac{\sum_{i=1}^n \frac{x_i - a}{c}}{n} c + a$$

denumită formula de calcul simplificat al mediei aritmetice. În cazul mediei aritmetice ponderate, formula de calcul simplificat devine:

$$\bar{x} = \frac{\sum_{i=1}^k \frac{x_i - a}{c} n_i}{\sum_{i=1}^k n_i} c + a$$

5. Dacă frecvențele unei serie de distribuție se amplifică sau se simplifică cu o constantă p , media calculată cu noile frecvențe este egală cu media seriei inițiale, astfel:

$$\bar{x}''' = \frac{\sum_{i=1}^k x_i (pn_i)}{\sum_{i=1}^k pn_i} = \frac{p \sum_{i=1}^k x_i n_i}{p \sum_{i=1}^k n_i} = \bar{x}$$

$$\bar{x}''' = \frac{\sum_{i=1}^k x_i \frac{n_i}{p}}{\sum_{i=1}^k \frac{n_i}{p}} = \frac{\frac{1}{p} \sum_{i=1}^k x_i n_i}{\frac{1}{p} \sum_{i=1}^k n_i} = \bar{x}$$

Exemple:

1. Se cunosc date privind distribuția firmelor din orașul M după valoarea totală a activelor, prezentată în tabelul de mai jos:

Distribuția firmelor după valoarea totală a activelor

Grupe de firme după valoarea totală a activelor (mii lei)	n_i	x_i (centrul intervalului)	$x_i n_i$
$x_i^{\text{inf}} - x_i^{\text{sup}}$			
Până la 100	20	75	1500
101 - 150	23	125	2875
151 - 200	55	175	9625
201 - 250	79	225	17775
251 - 300	78	275	21450
Peste 300	35	325	11375
Total	290	-	64600

Media activelor totale, calculată după formula mediei aritmetice ponderate, este egală cu 222,8 mii lei.

Proprietatea 1. Media aritmetică este cuprinsă între valoarea maximă și cea minimă a caracteristicii:

$$X_{\min} < \bar{X} < X_{\max}$$

$$75 < 222,8 < 325$$

Proprietatea 2. Suma abaterilor valorilor individuale față de media lor este nulă, adică:

- pentru serii de frecvențe

$$\sum_{i=1}^6 (x_i - \bar{x})n_i = 0$$

Conform datelor prezentate în tabelul de mai jos:

Grupe de firme după valoarea totală a activelor (mii lei)				
$X_i^{\inf} - X_i^{\sup}$	n_i	x_i	$x_i - \bar{x}$	$(x_i - \bar{x})n_i$
Până la 100	20	75	-147,76	-2955,17
101 - 150	23	125	-97,76	-2248,45
151 - 200	55	175	-47,76	-2626,72
201 - 250	79	225	2,24	177,07
251 - 300	78	275	52,24	4074,83
Peste 300	35	325	102,24	3578,45
Total	290	-	-	0

2. Se cunosc notele obținute de către cinci studenți la examenul de statistică: 10, 5, 8, 6, 9.

Proprietatea 3. Media calculată pe baza variantelor caracteristicii mărite sau micșorate cu o constantă (de ex. $a=10$) se modifică în același sens și cu aceeași mărime față de media seriei inițiale:

$$\bar{x}' = \frac{\sum_{i=1}^6 (x_i - a)}{n} = \frac{\sum_{i=1}^6 (x_i - 10)}{n} = \frac{-12}{5} = -2,4 = \bar{x} - 10 = 7,6 - 10 = -2,4$$

$$\bar{x}' = \frac{\sum_{i=1}^6 (x_i + a)}{n} = \frac{\sum_{i=1}^6 (x_i + 10)}{n} = \frac{88}{5} = 17,6 = \bar{x} + 10 = 7,6 + 10 = 17,6$$

Conform datelor prezentate în tabelul de mai jos:

x_i	$x_i - a$	$x_i + a$
10	0	20
5	-5	15
8	-2	18
6	-4	16
9	-1	19
38	-12	88

Proprietatea 4. Media calculată din variantele caracteristicii multiplicată sau simplificată cu o constantă (de ex. $c=3$) se multiplică sau se simplifică cu aceeași constantă:

$$\bar{x}'' = \frac{\sum_{i=1}^6 c \cdot x_i}{n} = \frac{\sum_{i=1}^6 3 \cdot x_i}{n} = \frac{114}{5} = 22,8 = 3 \cdot \bar{x} = 3 \cdot 7,6 = 22,8$$

$$\bar{x}'' = \frac{\sum_{i=1}^n \frac{x_i}{c}}{n} = \frac{12,66}{5} = 2,53 = \frac{\bar{x}}{c} = \frac{7,6}{3} = 2,53$$

Conform datelor prezentate în tabelul de mai jos:

x_i	cx_i	x_i/c
10	30	3,33
5	15	1,66
8	24	2,66
6	18	2,00
9	27	3,00
38	114	12,66

Să ne reamintim

- *Media armonică* este un tip de medie care se utilizează în cazuri speciale;
- Calculul mediei pătratice presupune utilizarea pătratelor valorilor.

- *Media geometrică* reprezintă o medie utilizată atunci când termenii seriei nu se află în relație de aditivare, ci de multiplicare, așa cum se întâmplă în cazul seriilor de indici de creștere.



7.3. Rezumat

Deși media aritmetică este unul dintre cei mai frecvent utilizați indicatori în statistica descriptivă, există și alte tipuri de medie care se utilizează în funcție de natura și obiectivele analizei statistice. De exemplu, la calculul productivității medii a muncii, atunci când aceasta este măsurată cu ajutorul consumului specific de muncă, sau la estimarea indicilor sintetici ai prețurilor, atunci când aceștia sunt calculați ca medie ponderată a indicilor individuali ai prețurilor (conform metodei de ponderare Laspeyres).

Media pătratică este, de asemenea, o medie utilizată în situații speciale, și anume atunci când se dorește să se acorde o importanță mare existenței unor valori mari, precum și atunci când media aritmetică nu poate fi calculată din cauză că valorile sunt pozitive și negative și se anulează reciproc, așa cum se întâmplă în cazul mediei abaterilor față de medie.

Între tipurile de medii studiate în cadrul acestui capitol amintim: media pătratică, media armonică și media geometrică.



7.4. Test de autoevaluare a cunoștințelor

1) Media armonică se utilizează în cazuri speciale, de exemplu:

- la calculul productivității medii a muncii, atunci când aceasta este măsurată cu ajutorul consumului specific de muncă;
- la calculul mediei notelor în învățământul primar;
- la estimarea indicilor sintetici ai prețurilor, atunci când aceștia sunt calculați ca medie ponderată a indicilor individuali ai prețurilor (conform metodei de ponderare Laspeyres);
- la estimarea produsului intern brut;
- a)+c).

Alegeți varianta corectă.

2) Media pătratică este utilizată în situații speciale, și anume atunci când:

- f. media aritmetică nu poate fi calculată din cauză că valorile sunt pozitive și negative și se anulează reciproc;
- g. se dorește să se acorde o importanță mare existenței unor valori mari în serie;
- a. se dorește să se calculeze productivitatea medie a muncii;
- b. media aritmetică nu poate fi calculată din cauză că valorile sunt pozitive;
- c. a)+b)

Alegeți varianta corectă.

3) Media are următoarele proprietăți:

- a. este cuprinsă între valoarea maximă și cea minimă a caracteristicii;
- b. Suma abaterilor valorilor individuale față de media lor este nulă;
- c. Media calculată pe baza variantelor caracteristicii mărite sau micșorate cu o constantă se modifică în același sens și cu aceeași mărime față de media seriei inițiale;
- d. Media calculată din variantele caracteristicii înmulțite sau simplificate cu o constantă se înmulțește sau se simplifică cu aceeași constantă;
- e. Toate variantele de mai sus.

Alegeți varianta corectă.

4) Distribuția notelor obținute la examenul de matematică pentru o grupă formată din 50 de studenți este următoarea:

X_i (puncte de notă)	n_i (numărul de studenți)
6	8
7	15
8	13
9	10
10	4
Total	50

Media pătratică la examenul de matematică, determinată conform formulei de calcul

$$\bar{x}_p = \sqrt{\frac{\sum_{i=1}^k x_i^2 n_i}{\sum_{i=1}^k n_i}}, \text{ este egală cu:}$$

- a. 7,23
- b. 7,82
- c. 7,89
- d. 8,11
- e. 8,90

Alegeți varianta corectă.

- 5) Distribuția notelor obținute la examenul de statistică pentru o grupă formată din 25 de studenți este următoarea:

X_i (puncte de notă)	n_i (numărul de studenți)
8	13
9	10
10	4
Total	25

Media geometrică, determinată conform formulei de calcul $\bar{X}_g = \sqrt[\sum_{i=1}^k n_i]{\prod_{i=1}^k x_i^{n_i}}$, este egală cu:

- a. 7,23
- b. 8,60
- c. 7,89
- d. 8,11
- e. 9,90

Alegeți varianta corectă.

- 6) Se cunosc date privind distribuția firmelor din orașul M după valoarea vânzărilor:

Grupe de firme după valoarea vânzărilor (mii lei)	n_i	X_i (centrul intervalului)
Până la 100	10	50
101 - 150	23	125
151 - 200	35	175
201 - 300	29	250
Peste 300	15	350

Total	112	-
--------------	------------	----------

Media armonică este egală cu:

- a. 50,8 mil. lei
- b. 1150,8 mii lei
- c. 150,8 mil. lei
- d. 150,8 lei
- e. 150,8 mii lei.

Alegeți varianta corectă.

7) Se cunosc date privind distribuția firmelor din orașul M după valoarea totală a activelor:

Grupe de firme după valoarea totală a activelor (mii lei)	n_i	X_i (centrul intervalului)
Până la 100	20	75
101 - 150	23	125
151 - 200	55	175
201 - 250	79	225
251 - 300	78	275
Peste 300	35	325
Total	290	-

Media armonică a valorii activelor totale este egală cu:

- f. 192,4 mii lei;
- g. 237,8 mii lei;
- h. 222,8 mil. lei;
- i. 192,4 mil. lei;
- j. 224,9 mii lei.

Alegeți varianta corectă.

8) Se cunosc date privind cifra de afaceri a șase firme din orașul M:

Cifra de afaceri (mii lei)	
1	134
2	150

3	567
4	201
5	251
6	450
total	

Media armonică a cifrei de afaceri este egală cu:

- a. 192,4 mii lei;
- b. 221,6 mii lei;
- c. 222,8 mil. lei;
- d. 222,4 mil. lei;
- e. 224,9 mii lei.

Alegeți varianta corectă.

9) Se cunosc date privind vânzările a șase magazine din orașul P.

(mii lei)

M1	M2	M3	M4	M5	M6
23	36	44	67	59	22

Media pătratică a vânzărilor este egală cu:

- f. 45,12 mii lei
- g. 30,51 mii lei
- h. 45,82 mii lei
- i. 22,8 mii lei
- j. 60,0 mii lei.

Alegeți varianta corectă.



7.5. Bibliografie recomandată

- f. Caragea Nicoleta, *Statistica- Concepte si metode de analiza a datelor*, Editura MUSTANG, București, 2015
- g. Voineagu, V. , *Statistică –baze teoretice și aplicații*, Editura Economică, București, 2007
- h. Georgescu-Roegen N., *Metoda statistică*, Editura Expert, București, 1998.

- i. Andrei T., Stancu S., Pele D.T., *Statistică. Teorie și aplicații* (ediția a II-a), Editura Economică, București, 2003.
- j. Biji M., Biji E.M., Lilea E., Anghelache C., *Tratat de statistică*, Editura Economică, București, 2002.

Unitatea de învățare 8.

MEDIANA



8.1. Introducere

În cadrul acestui capitol, studenților le este prezentată mediana, ca principal indicator al tendinței centrale, alături de medie. De asemenea, va fi prezentat algoritmul de calcul al medianei, atât pentru seriile simple, cât și pentru seriile de date grupate.

Durata medie de parcurgere a acestui capitol este de aproximativ o oră.



8.2. Conținut

Mediana este un indicator mediu de poziție care face parte din categoria cuantilelor, fiind egală cu acea valoare a caracteristicii ce împarte colectivitatea ordonată crescător sau descrescător în două părți egale, astfel încât numărul valorilor mai mari decât ea este egal cu numărul valorilor mai mici⁷.

Determinarea medianei presupune stabilirea locului acesteia, adică a numărului de ordine/rangului unității mediane (a unității care înregistrează valoarea mediană).

Mediana se exprimă în aceleași unități de măsură ca și variabilele.

8.2.1. Mediana seriilor simple

Cazul 1. Serie impară

În cazul unei serii simple formate dintr-un număr impar de termeni, mediana este egală cu termenul central, adică:

$$Me = x_i$$

Locul medianei fiind determinat după formula:

⁷ Altfel spus, mediana este valoarea înregistrată de unitatea situată în mijlocul seriei ordonate crescător sau descrescător, marcând împărțirea colectivității în două părți egale ca număr de unități.

$$i = \frac{n+1}{2}$$

Unde n reprezintă numărul de termeni care compun seria.

Cazul 2. Serie pară

Dacă seria este formată dintr-un număr par de termeni, mediana se calculează ca medie aritmetică simplă a celor doi termeni centrali:

$$Me = \frac{x_i + x_{i+1}}{2}$$

Corespunzător rangului medianei:

$$i = \frac{n}{2}$$

Exemple

1. Se cunosc date privind volumul vânzărilor de frigidere înregistrate în primul trimestru al anului t, de către șapte firme care comercializează aparatură electrocasnică:

(bucăți)

F1	F2	F3	F4	F5	F6	F7
17	15	12	14	16	18	13

Pentru a determina mediana, se ordonează crescător sau descrescător termenii seriei:

F1	F2	F3	F4	F5	F6	F7
12	13	14	15	16	17	18

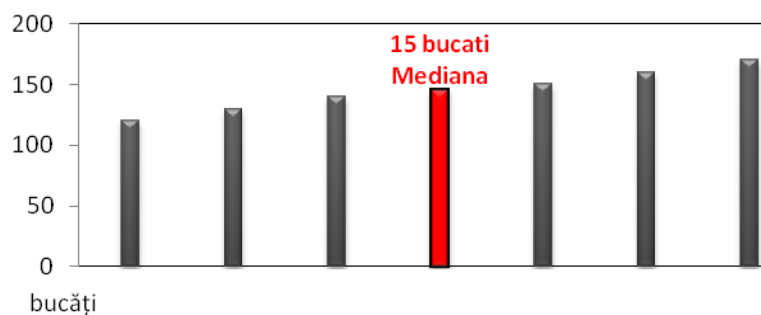
Fiind o serie formată dintr-un număr impar de termeni ($n=7$), mediana este egală cu termenul central, adică: $Me = x_i$

$$\text{Unde : } i = \frac{n+1}{2} = \frac{7+1}{2} = 4$$

Astfel, valoarea medianei va fi de 15 bucăți:

$$Me = x_4 = 15$$

Mediana volumului vânzărilor



2. Se cunosc date privind cifra de afaceri înregistrată în anul t de opt firme din orașul M:

(mii lei)

F1	F2	F3	F4	F5	F6	F7	F8
170	150	120	140	160	180	130	110

Pentru a determina mediana, se ordonează crescător sau descrescător termenii seriei:

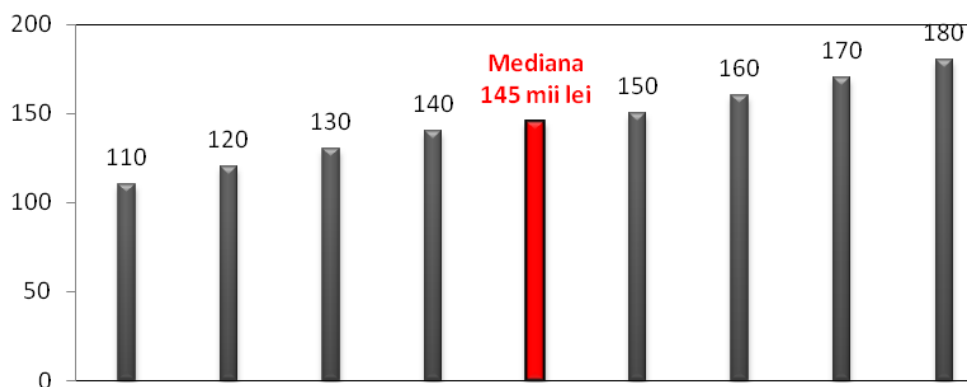
F1	F2	F3	F4	F5	F6	F7	F8
110	120	130	140	150	160	170	180

Fiind o serie formată dintr-un număr par de termeni ($n=8$), mediana este egală cu media aritmetică simplă a celor doi termeni centrali, adică:

$$Me = \frac{x_i + x_{i+1}}{2}$$

$$\text{Corespunzător rangului medianei: } i = \frac{n}{2} = \frac{8}{2} = 4$$

Mediana cifrei de afaceri



În acest caz, mediana este egală cu 145 mii lei:

$$Me = \frac{x_4 + x_5}{2} = \frac{140 + 150}{2} = 145$$

8.2.2. Mediana seriilor de frecvențe

Într-o serie de frecvențe, locul medianei /rangul unității mediane se determină după formula:

$$U_{Me} = \frac{\sum_{i=1}^k n_i + 1}{2}$$

Se stabilește care este valoarea mediană, în cazul seriilor de distribuție pe variante, sau intervalul median, adică grupa în care se află valoarea mediană, în cazul distribuțiilor pe intervale.

Identificarea valorii mediane sau a intervalului median poate fi realizată în două modalități:

- Se calculează frecvențele cumulate crescător ($N_i^{\wedge}$) și se compară cu rangul unității mediane (U_{Me}). Prima frecvență cumulată care depășește rangul medianei, $N_i^{\wedge} > U_{Me}$, este cea care indică varianta mediană / intervalul median.
- Se calculează frecvențele cumulate crescător ($N_i^{\wedge}$) și cele cumulate descrescător (N_i). Grupa pentru care cele două frecvențe sunt egale ($N_i^{\wedge} = N_i$) sau pentru care diferența dintre ele este minimă ($N_i^{\wedge} - N_i = \min.$) reprezintă varianta sau intervalul median.

În cazul distribuțiilor pe intervale, după identificarea intervalului median, se calculează nivelul medianei, cu ajutorul unei formule de interpolare:

$$Me = x_{Me-1}^{sup} + h_{Me} \frac{\frac{\sum_{i=1}^k n_i + 1}{2} - N_{Me-1}}{n_{Me}}$$

unde:

x_{Me-1}^{sup} reprezintă limita superioară a intervalului precedent celui median;

h_{Me} - mărimea intervalului median;

$N_{Me-1}^{\wedge}$ - frecvența cumulată crescător corespunzătoare intervalului precedent celui median;

$N_{Me}^{\wedge}$ - frecvența intervalului median.

Exemple:

1. Produsul intern brut pe locuitor în unele state membre ale Uniunii Europene se prezintă astfel:

	Regatul Unit	România	Slovenia	Spania	Ungaria
PIB/locuitor (euro)	35,4	12,2	27,6	31,9	19,3
Numărul locuitorilor (mil. persoane)	61,60	21,39	2,12	46,06	10,07

Mediana PIB/locuitor se determină astfel:

- Se determină locul mediane distribuției țărilor după valoarea PIB/locuitor:

$$U_{Me} = \frac{141,24 + 1}{2} = 71,12$$

- Se aranjează crescător sau descrescător termenii seriei (PIB/locuitor) și se calculează frecvențele cumulate crescător (sau descrescător):

Algoritmul de calcul al mediane PIB/locuitor

	PIB/locuitor (euro)	n (mil. pers.)	N [^]
România	12,2	21,39	21,39
Ungaria	19,3	10,07	31,46
Slovenia	27,6	2,12	33,58
Spania	31,9	46,06	79,64
Regatul Unit	35,4	61,6	141,24
	-	141,24	-

- Prima frecvență cumulată crescător care depășește valoarea rangului mediane (79,64 mil. persoane) corespunde mediane (Me = 31,9 euro – Spania).

2. Distribuția țărilor după numărul utilizatorilor de Internet ce revin la 1000 locuitori se prezintă astfel:

Distribuția țărilor după numărul utilizatorilor de Internet ce revin la 1000 locuitori

Grupe după numărul de utilizatori Internet (la 1000 locuitori)	Numărul de țări
Până la 10	33
10 - 50	45
50 - 100	29
100 – 200	20
200 – 400	21
400 – 500	11
Peste 500	18
Total	177

Mediana utilizatorilor de Internet ce revin la 1000 locuitori se determină după următorul algoritm:

- Se determină rangul mediane distribuției țărilor după numărul utilizatorilor de Internet ce revin la 1000 locuitori:

$$UMe = \frac{177 + 1}{2} = 89$$

- Se determină intervalul median (care conține mediana), astfel:
 - o Se calculează frecvențele cumulate crescător
 - o Se compară rangul mediane cu frecvențele cumulate crescător. Prima frecvență mai mare decât locul mediane, corespunde intervalului median
 - o Se determină mediana conform formulei de interpolare.

Algoritmul de calcul al mediane numărului de utilizatori Internet

Utilizatori Internet (la 1000 locuitori)	număr țări n (frecvențe absolute)	N[^] (frecvențe cumulate)
Până la 10	33	33
10 - 50	45	78
Interval median → 50 - 100	29	107
100 – 200	20	127
200 – 400	21	148
400 – 500	11	159
Peste 500	18	177
Total	177	-

Din tabelul de mai sus, se observă că prima frecvență cumulată crescător - mai mare decât locul medianei (89) - este cea de a treia (107); aceasta corespunde intervalului (50-100), denumit interval median.

Mediana se calculează după formula:

$$Me = x_{Me-1}^{sup} + h_{Me} \times \frac{\sum_{i=1}^n n_i + 1}{2} - N_{Me-1} = 50 + 50 \times \frac{89 - 78}{29} = 68,96$$

utilizatori Internet (la 1000 locuitori).

Să ne reamintim

- *Mediana este un indicator mediu de poziție care face parte din categoria quantilelor, fiind egală cu acea valoare a caracteristicii ce împarte colectivitatea ordonată crescător sau descrescător în două părți egale, astfel încât numărul valorilor mai mari decât ea este egal cu numărul valorilor mai mici .*



8.3. Rezumat

În categoria indicatorilor de nivel, denumiți și indicatori de poziție, sunt cuprinși indicatorii tendinței centrale și quantilele.

Indicatorii cu ajutorul cărora se analizează tendința centrală a fenomenelor de masă sintetizează valorile individuale într-un singur nivel, reprezentativ pentru întreaga colectivitate. Principalii indicatori ai tendinței centrale sunt: media, mediana și modul.

Mediana presupune un algoritm de calcul diferit pentru situațiile în care seriile de date sunt simple sau de frecvențe.



8.4. Test de autoevaluare a cunoștințelor

1) Identificarea valorii mediane sau a intervalului median poate fi realizată în două modalități:

- a. Se calculează frecvențele cumulate crescător ($N_i^{\wedge}$) și se compară cu rangul unității mediane (UMe). Prima frecvență cumulată care depășește rangul medianei, $N_i^{\wedge} > UMe$, este cea care indică varianta mediană / intervalul median;

- b. Se calculează frecvențele cumulate descrescător (N_i) și se compară cu rangul unității mediane (U_{Me});
- c. Se calculează primele două frecvențe cumulate crescător. Dacă această valoare depășește rangul medianei, $N_i^{\wedge} > U_{Me}$, indică varianta mediană / intervalul median;
- d. Se calculează frecvențele cumulate crescător ($N_i^{\wedge}$) și cele cumulate descrescător (N_i). Grupa pentru care cele două frecvențe sunt egale ($N_i^{\wedge} = N_i$) sau pentru care diferența dintre ele este minimă ($N_i^{\wedge} - N_i = \min.$) reprezintă varianta sau intervalul median;
- e. a+d)

Alegeți varianta corectă.

2) Etapele algoritmului de calcul pentru determinarea medianei unei serii de distribuție pe intervale este următorul:

- a. Se calculează frecvențele cumulate crescător și se determină mediana conform formulei de interpolare;
- b. Se determină mediana conform formulei de interpolare și se determină rangul medianei;
- c. Se determină rangul medianei, se calculează frecvențele cumulate crescător, se compară rangul medianei cu frecvențele cumulate crescător, astfel încât prima frecvență mai mare decât locul medianei, corespunde intervalului median, se determină mediana conform formulei de interpolare;
- d. Se calculează frecvențele cumulate crescător, se compară rangul medianei cu frecvențele cumulate crescător, astfel încât prima frecvență mai mare decât locul medianei, corespunde intervalului median, se determină mediana conform formulei de interpolare;
- e. Se compară rangul medianei cu frecvențele cumulate crescător, astfel încât prima frecvență mai mare decât locul medianei, corespunde intervalului median.

Alegeți varianta corectă.

3) Mediana seriei:

45 89 56 67 55 33 12 47 33 89 este egală cu:

- a. 51
- b. 56
- c. 44
- d. 33
- e. 89

Alegeți varianta corectă.

4) Mediana seriei:

245 890 256 367 155 233 312 247 133 289 este egală cu:

- a. 151,0
- b. 194,5
- c. 251,5
- d. 155,0
- e. 233,5

Alegeți varianta corectă.

5) Mediana seriei:

24 90 25 36 87 este egală cu:

- a. 36
- b. 90
- c. 24
- d. 87
- e. 25

Alegeți varianta corectă.

6) Mediana seriei:

367 155 233 312 247 133 289 este egală cu:

- a. 247
- b. 367
- c. 289
- d. 233
- e. 155

Alegeți varianta corectă.

7) Se cunoaște distribuția firmelor din orașul M după cifra de afaceri:

Grupe de firme după cifra de afaceri (mii lei)	n (număr firme)	N [^] (frecvențe cumulate)
Până la 10	25	25
10 - 50	45	70
50 - 100	29	99
100 – 200	65	164
200 – 400	21	185
400 – 500	11	196
Peste 500	18	214
Total	214	-

Mediana cifrei de afaceri se află în intervalul:

- f. 10-50

- g. 100-200
- h. Până la 10
- a. Peste 500
- b. 200-400

Alegeți varianta corectă.

8) Se cunoaște distribuția firmelor din orașul M după numărul de salariați:

Grupe de firme după numărul de salariați (persoane)	n (număr firme)	N [^] (frecvențe cumulate)
Până la 10	40	40
10 - 50	45	85
50 - 100	70	155
100 – 200	65	220
Peste 200	18	238
Total	238	

Mediana cifrei de afaceri se află în intervalul:

- a. 10-50
- b. 100-200
- c. 50-100
- d. Peste 200
- e. 200-400

Alegeți varianta corectă.

9) Se cunoaște distribuția firmelor din orașul M după cifra de afaceri:

Grupe de firme după cifra de afaceri (mii lei)	n (număr firme)	N [^] (frecvențe cumulate)
Până la 10	25	25
10 - 50	45	70
50 - 100	29	99
100 – 200	65	164
200 – 400	21	185
400 – 500	11	196
Peste 500	18	214
Total	214	

Mediana cifrei de afaceri, calculată după formula de interpolare

$$Me = x_{Me-1}^{sup} + h_{Me} \frac{\frac{\sum_{i=1}^k n_i + 1}{2} - N_{Me-1}}{n_{Me}}$$

este egală cu:

- a. 113,07 mii lei
- b. 56,80 mii lei
- c. 120,90 mii lei
- d. 568,05 mii lei
- e. 123,90 mii lei.

Alegeți varianta corectă.

10) Se cunoaște distribuția firmelor din județul S după volumul producției anuale de lapte:

Grupe de firme după volumul producției de lapte (mii hl)	n (număr firme)	N [^] (frecvențe cumulate)
Până la 2	10	10
2 - 5	15	25
5 - 10	21	46
10 – 20	13	59
20 – 40	10	69
40 – 50	11	80
Peste 50	2	82
Total	82	-

Mediana producției de lapte, calculată după formula de interpolare

$$Me = x_{Me-1}^{sup} + h_{Me} \frac{\frac{\sum_{i=1}^k n_i + 1}{2} - N_{Me-1}}{n_{Me}}, \text{ este egală cu:}$$

- a. 6,8 mii hl
- b. 8,9 mii hl
- c. 20,0 mii hl
- d. 8,0 mii hl
- e. 3,9 mii hl.

Alegeți varianta corectă.



8.5. Bibliografie recomandată

- k. Caragea Nicoleta, *Statistica- Concepte si metode de analiza a datelor*, Editura MUSTANG, București, 2015
- l. Voineagu, V. , *Statistică –baze teoretice și aplicații*, Editura Economică, București, 2007
- m. Georgescu-Roegen N., *Metoda statistică*, Editura Expert, București, 1998.
- n. Andrei T., Stancu S., Pele D.T., *Statistică. Teorie și aplicații* (ediția a II-a), Editura Economică, București, 2003.
- o. Biji M., Biji E.M., Lilea E., Anghelache C., *Tratat de statistică*, Editura Economică, București, 2002.

Unitatea de învățare 9.

MODUL



9.1. Introducere

În cadrul acestui capitol, studenților le este prezentat modul, ca principal indicator al tendinței centrale, alături de medie și mediană. De asemenea, va fi prezentat algoritmul de calcul al modului, atât pentru seriile simple, cât și pentru seriile de date grupate.

Durata medie de parcurgere a acestui capitol este de aproximativ o oră.



9.2. Conținut

Modul (valoarea modală sau dominantă a seriei) unei serii reprezintă varianta cu frecvența cea mai mare.

Modul unei distribuții pe variante este ușor de identificat pe baza frecvențelor absolute sau relative. Există în practică și serii de distribuție multimodale. În astfel de situații se determină mai multe valori modale, dar ele nu pot fi sintetizate pentru a se obține o singură valoare modală pentru întreaga colectivitate.

9.2.1. Modul seriilor simple

În cazul unei serii simple, modul este acea valoare din serie care se repetă cel mai des. Dacă nu există valori care se repetă, seria nu prezintă mod.

9.2.2. Modul seriilor de distribuție

Modul unei distribuții pe variante este ușor de identificat pe baza frecvențelor absolute sau relative. Cea mai mare frecvență a unei variante indică faptul că acea variantă este valoarea modală a seriei.

Într-o serie de distribuție pe intervale, se identifică mai întâi intervalul modal, adică intervalul cu frecvența cea mai mare, grupa în care se află cele mai multe unități ale colectivității. În cadrul acestui interval, valoarea modală se aproximează, ca și mediana, pe baza unei formule de interpolare:

$$Mo = x_{Mo-1}^{sup} + h_{Mo} \frac{\Delta_1}{\Delta_1 + \Delta_2}$$

unde:

x_{Mo-1}^{sup} reprezintă limita superioară a intervalului care îl precede pe cel modal;

h_{Mo} - mărimea intervalului modal.

Δ_1 - diferența dintre frecvența intervalului modal n_{Mo} și cea a intervalului precedent n_{Mo-1}

$$\Delta_1 = n_{Mo} - n_{Mo-1}$$

Δ_2 - diferența dintre frecvența intervalului modal n_{Mo} și cea a intervalului următor n_{Mo+1}

$$\Delta_2 = n_{Mo} - n_{Mo+1}$$

Exemple

S-a efectuat o cercetare privind mărimea (măsurată pe baza cifrei de afaceri) a 32 de firme, având ca activitate principală comercializarea cu amănuntul a mărfurilor alimentare. Datele referitoare la cifra de afaceri înregistrată în cursul observării sunt următoarele (în mii lei):

166	162	121	126
141	142	92	150
155	230	100	129
182	102	128	198
132	150	75	206
134	224	166	236
190	156	108	143
150	156	142	127

Valoarea modală a cifrei de afaceri este 150 mii lei, deoarece este valoarea cu cea mai mare frecvență de apariție (se regăsește de trei ori).

1. S-au înregistrat următoarele niveluri ale populației ocupate, pe grupe de vârstă:

Distribuția populației ocupate, după vârstă

Grupe de vârstă (ani)	n Populația ocupată (mii persoane)
Până la 20	893
20-25	1177
25-30	1288
30-40	1434
40-50	1245
Peste 50	789
Total	

Modul distribuției populației ocupate pe grupe de vârstă se determină astfel:

- se identifică mai întâi intervalul modal, adică intervalul cu frecvența cea mai mare, grupa în care se află cele mai multe unități ale colectivității. Acesta corespunde celui de al patrulea interval (30-40 ani).
- în cadrul acestui interval, valoarea modală se aproximează pe baza formulei de interpolare, fiind egală cu 34 ani:

$$Mo = x_{Mo-1}^{sup} + h_{Mo} \frac{\Delta_1}{\Delta_1 + \Delta_2} = 30 + 10 \times \frac{146}{146 + 189} = 30 + 4,35 = 34,35$$

Unde:

$$\Delta_1 = n_{Mo} - n_{Mo-1} = 1434 - 1288 = 146$$

$$\Delta_2 = n_{Mo} - n_{Mo+1} = 1434 - 1245 = 189$$

Modul, la fel ca și media și mediana, se exprimă în aceleași unități de măsură ca și variabilele. Metodologia de calcul a modului pentru seriile de distribuție de frecvențe relative este similară cu cea prezentată mai sus.

În practică există serii de distribuție care prezintă două sau mai multe valori modale. Acestea nu sunt aditive și, în consecință, nu pot fi agregate pentru determinarea unei valori modale globale.

Este important se sublinieze, de asemenea, că fiecare dintre indicatorii tendinței centrale (media, mediana, modul) sunt calculați cu precizie numai pe baza seriilor brute, care cuprind valorile înregistrate de toate unitățile colectivității. Indicatorii calculați pe baza seriilor de distribuție pe intervale prezintă un grad mai scăzut de precizie, cu atât mai scăzut cu cât intervalele de grupare sunt mai mari și numărul grupelor mai mic. În aceste condiții, se recomandă utilizarea acestora din

urmă numai în situațiile în care datele primare/neprelucrate nu sunt disponibile sau dacă volumul colectivității este mare și nu este posibilă folosirea unui calculator de capacitate corespunzătoare.

Să ne reamintim

- Modul (valoarea modală sau dominantă seriei) unei serii reprezintă varianta cu frecvența cea mai mare.
- Modul unei distribuții pe variante este ușor de identificat pe baza frecvențelor absolute sau relative. Există în practică și serii de distribuție multimodale. În astfel de situații se determină mai multe valori modale, dar ele nu pot fi sintetizate pentru a se obține o singură valoare modală pentru întreaga colectivitate.



9.3. Rezumat

Indicatorii cu ajutorul cărora se analizează tendința centrală a fenomenelor de masă sintetizează valorile individuale într-un singur nivel, reprezentativ pentru întreaga colectivitate. Principalii indicatori ai tendinței centrale sunt: media, mediana și modul.

Modul presupune un algoritm de calcul diferit pentru situațiile în care seriile de date sunt simple sau de frecvențe.



9.4. Test de autoevaluare a cunoștințelor

1) Se cunoaște distribuția salariaților după nivelul salariului:

Grupe de salarii (lei)	n (număr de salariați)
500-1000	10
1001-1500	15
1501-2000	21
2001-3000	13
3001-4000	10
Peste 4000	11
Total	80

Salariul modal este cuprins în grupa:

- a. 2001-3000
- b. 1001-1500
- c. 1501-2000
- d. Seria nu prezintă valoare modală
- e. Peste 4000

Alegeți varianta corectă.

2) Se cunosc date privind numărul cheltuielilor lunare pentru întreținerea locuinței, efectuate de 10 gospodării:

gospodăria	Cheltuieli lunare (lei lunar)
1	300
2	150
3	210
4	130
5	410
6	1100
7	345
8	130
9	455
10	290

Modul seriei este egal cu:

- a. 300 lei lunar
- b. 1100 lei lunar
- c. 130 lei lunar
- d. Seria nu prezintă mod
- e. 400 lei lunar

Alegeți varianta corectă.

3) S-a efectuat o cercetare privind mărimea gospodăriilor ($n=30$) din localitatea X (măsurată pe baza numărului de persoane ce compun gospodăria):

5	2	4	1	2	10	3	2	5	3
4	5	6	4	8	4	6	5	5	4
3	10	7	1	5	1	2	5	1	5

Valoarea dominantă a seriei este:

- a. Seria nu prezintă o valoare dominantă
- b. 4 persoane
- c. 2 persoane
- d. 5 persoane
- e. 8 persoane

Alegeți varianta corectă.

4) Se cunoaște distribuția salariaților după nivelul salariului:

Grupe de salarii (lei)	n (număr de salariați)
500-1000	10
1001-1500	15
1501-2000	21
2001-3000	13
3001-4000	10
Peste 4000	11
Total	80

Salariul modal, calculat pe baza formulei de interpolare

$$Mo = x_{Mo-1}^{sup} + h_{Mo} \frac{n_{Mo} - n_{Mo-1}}{(n_{Mo} - n_{Mo-1}) + (n_{Mo} - n_{Mo+1})}$$

este egal cu:

- a. 1714,3 lei
- b. 1712,4 lei
- c. 1713,3 lei
- d. 1765,2 lei

e. 1772,3 lei

Alegeți varianta corectă.

5) Conform Anuarului Statistic al României s-a înregistrat următoarea distribuție a municipiilor și orașelor după mărime (evaluată pe baza numărului locuitorilor):

Grupe de municipii și orașe după numărul locuitorilor	Pondere față de numărul total de municipii și orașe, %
Sub 3000	1,6
3000-4999	3,9
5000-9999	31,4
10000-19999	29,2
20000-49999	19,2
50000-99999	6,7
100000-199999	4,5
200000-999999	3,2
1000000 și peste	0,3

Modul distribuției municipiilor și orașelor după mărime este cuprins în grupa:

- a. 3000-4999 locuitori
- b. 10000-19999 locuitori
- c. Peste 1000000 locuitori
- d. 50000-199999 locuitori
- e. Nici una dintre variantele de mai sus.

Alegeți varianta corectă.



9.5. Bibliografie recomandată

- p. Caragea Nicoleta, *Statistica- Concepte si metode de analiza a datelor*, Editura MUSTANG, București, 2015
- q. Voineagu, V. , *Statistică –baze teoretice și aplicații*, Editura Economică, București, 2007
- r. Georgescu-Roegen N., *Metoda statistică*, Editura Expert, București, 1998.
- s. Andrei T., Stancu S., Pele D.T., *Statistică. Teorie și aplicații* (ediția a II-a), Editura Economică, București, 2003.
- t. Biji M., Biji E.M., Lilea E., Anghelache C., *Tratat de statistică*, Editura Economică, București, 2002.

9.6. Teme de control

Tema 2 – Capitolul 6-9: „Realizați un referat de maximum 5 pagini în care să calculați indicatorii tendinței centrale ai unei serii de distribuție”;

Unitatea de învățare 10.

INDICATORII DE NIVEL. QUANTILELE



10.1. Introducere

În cadrul acestui capitol, studenților le sunt prezentați indicatorii de nivel. De asemenea, va fi prezentat algoritmul de calcul al quantilelor.

Durata medie de parcurgere a acestui capitol este de aproximativ o oră.



10.2. Conținut

Ca și mediana, quantilele sunt valori ale caracteristicii pe baza cărora colectivitatea - ordonată crescător sau descrescător în funcție de valorile caracteristicii - este împărțită în grupe cu frecvențe egale:

- quartilele, care împart colectivitatea în patru părți egale, fiecare cuprinzând un sfert din numărul de unități ale colectivității;
- quintilele, care împart colectivitatea în cinci părți egale, fiecare cuprinzând câte 20% din numărul total de unități;
- decilele, în raport cu care colectivitatea este împărțită în zece părți egale, a câte 10% din numărul total de unități;
- semidecilele, care servesc la împărțirea colectivității în 20 de părți egale, a câte 5% din numărul total de unități.

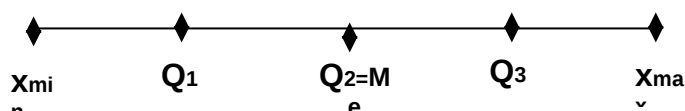
În practica statistică, sunt denumite quartile, quintile, decile sau semidecile și cele patru, cinci, zece sau douăzeci de grupe care cuprind fiecare câte 25%, 20%, 10% sau 5% din numărul total de unități.

Decilele și, mai ales, centilele se calculează doar pentru seriile cu număr mare de observații. Decilele și centilele sunt indicatori utilizați îndeosebi pentru seriile pronunțat asimetrice.

10.2.1. Quartilele - Definiție

Quartilele sunt acele valori ale caracteristicii care împart seria valorilor ordonate în patru părți egale.

- quartila inferioară (Q_1) este acea valoare care separă 25 % din valorile mici de restul de 75 % din valori;
- quartila a doua (Q_2) împarte seria în două părți egale, deci coincide cu mediana ($Q_2 = Me$);
- quartila a treia (Q_3) separă primii 75 % din valori de restul de 25 % din valorile mai mari.



10.2.2. Algoritmul de calcul al quartilelor

Ca și mediana, se determină pe baza valorilor ordonate ale seriei statistice. Prima operație constă în localizarea celor trei quartile.

- Locul primei quartile Q_1 (quartila inferioară):

$$L_{Q_1} = \frac{\sum_{i=1}^n n_i + 1}{4}$$

- Locul celei de a treia quartile Q_3 (quartila superioară):

$$L_{Q_3} = \frac{3 \times (\sum_{i=1}^n n_i + 1)}{4}$$

Pentru identificarea intervalului median, se calculează frecvențele cumulate crescător, apoi:

- ☐ Se compară locul quartilei cu frecvențele cumulate crescător. Prima frecvență mai mare decât locul quartilei, corespunde intervalului quartilei respective;
- ☐ Se determină quartilele conform formulei de interpolare:

$$Q_1 = x_{Q_1-1}^{\sup} + h_{Q_1} \times \frac{\frac{(\sum_{i=1}^n n_i + 1)}{4} - N_{Q_1-1}}{n_{Q_1}}$$

$$Q_3 = x_{Q_3-1}^{\text{sup}} + h_{Q_3} \times \frac{\frac{3 \times (\sum_{i=1}^n n_i + 1)}{4} - N_{Q_3-1}}{n_{Q_3}}$$

Să ne reamintim

Problemă propusă:

Se cunosc date privind distribuția a 80 de persoane după veniturile lunare.

Distribuția persoanelor, după venitul lunar

	grupe de venit (lei)	ni	hi	xi	xini	Ni^
	– 1000	6	300	850	5100	6
interval quartila inferioară	1001 – 1300	24	300	1150	27600	30
interval median	1301 – 1600	30	300	1450	43500	60
interval quartila superioară	1601 – 1900	12	300	1750	21000	72
	1901 – 2200	5	300	2050	10250	77
	2201 -	3	300	2350	7050	80
	total	80	-	-	114500	

Să se determine media, mediana și quartilele venitului lunar și să se interpreteze semnificația rezultatelor.

Răspuns:

- quartila inferioară ($Q_1 = 1178.125$ lei); este acea valoare care separă 25 % din cele 80 observații – ordonate crescător - de restul de 75 % din valori;
- quartila a doua (1405 lei) coincide cu mediana ($Q_2 = Me$); este valoarea care împarte cele 80 de observații – ordonate crescător - în două părți egale, astfel încât, jumătate din veniturile lunare sunt mai mici decât 1405 lei, iar jumătate sunt mai mari.
- quartila superioară ($Q_3 = 1618.75$ lei); separa primele 75% din cele 80 observații – ordonate crescător - de restul de 25 % din valori.



10.3. Rezumat

În practica statistică, sunt denumite quartile, quintile, decile sau semidecile și cele patru, cinci, zece sau douăzeci de grupe care cuprind fiecare câte 25%, 20%, 10% sau 5% din numărul total de unități. Decilele și, mai ales, centilele se calculează doar pentru seriile cu număr mare de observații. Decilele și centilele sunt indicatori utilizați îndeosebi pentru seriile pronunțat asimetrice.



10.4. Test de autoevaluare a cunoștințelor

1) Se cunoaște distribuția salariaților după nivelul salariului:

Grupe de salarii (lei)	n (număr de salariați)
500-1000	10
1001-1500	15
1501-2000	21
2001-3000	13
3001-4000	10
Peste 4000	11
Total	80

Quartila inferioară este egală cu:

- a. 17,3 lei
- b. 171,4 lei
- c. 1713,3 lei
- d. 1765,2 lei
- e. Nicio variantă nu este corectă.

Alegeți varianta corectă.

2) Se cunoaște distribuția salariaților după nivelul salariului:

Grupe de salarii (lei)	n (număr de salariați)
500-1000	10
1001-1500	15
1501-2000	21
2001-3000	13
3001-4000	10
Peste 4000	11
Total	80

Quartila superioară este egală cu:

- a. 17,3 lei
- b. 171,4 lei
- c. 1713,3 lei
- d. 1765,2 lei
- e. Nicio variantă nu este corectă.

Alegeți varianta corectă.

3) Conform Anuarului Statistic al României, s-a înregistrat următoarea distribuție a municipiilor și orașelor după mărime (evaluată pe baza numărului locuitorilor):

Grupe de municipii și orașe după numărul locuitorilor	Pondere față de numărul total de municipii și orașe, %
Sub 3000	1,6
3000-4999	3,9
5000-9999	31,4
10000-19999	29,2
20000-49999	19,2
50000-99999	6,7
100000-199999	4,5
200000-999999	3,2
1000000 și peste	0,3

Quartila superioară a distribuției municipiilor și orașelor după mărime este cuprinsă în grupa:

- a. 3000-4999 locuitori
- b. 5000-9999 locuitori
- c. Peste 1000000 locuitori
- d. 10000-19999 locuitori
- e. Nici una dintre variantele de mai sus.

Alegeți varianta corectă.



10.5. Bibliografie recomandată

- u. Caragea Nicoleta, *Statistica- Concepte si metode de analiza a datelor*, Editura MUSTANG, București, 2015
- v. Voineagu, V. , *Statistică –baze teoretice și aplicații*, Editura Economică, București, 2007

- w. Georgescu-Roegen N., *Metoda statistică*, Editura Expert, București, 1998.
- x. Andrei T., Stancu S., Pele D.T., *Statistică. Teorie și aplicații* (ediția a II-a), Editura Economică, București, 2003.
- y. Biji M., Biji E.M., Lilea E., Anghelache C., *Tratat de statistică*, Editura Economică, București, 2002.

Unitatea de învățare 11.

INDICATORII SIMPLI AI VARIAȚIEI



11.1. Introducere

În cadrul acestui capitol, studenților le sunt prezentați indicatorii simpli ai variației: amplitudinea și abaterile individuale.

Durata medie de parcurgere a acestui capitol este de aproximativ o oră.



11.2. Conținut

Valorile unei serii statistice diferă unele față de altele, acestea fiind mai apropiate sau mai îndepărtate față de o valoare centrală de referință. Comparăția se face cu ajutorul mediei aritmetice, considerată un reper reprezentativ pentru toate caracteristicile seriei.

Indicatorii variației, denumiți și indicatori ai dispersiei sau ai împrăștierii, caracterizează modul în care valorile caracteristicii gravitează în jurul valorii centrale, gradul de împrăștiere a unităților colectivității în raport cu o anumită caracteristică, respectiv gradul de omogenitate al colectivității din punctul de vedere al acestei caracteristici.

În descrierea variației unei serii statistice se utilizează:

- indicatori simpli și
- indicatori sintetici ai variației.

Indicatorii simpli ai variației se calculează utilizând numai una sau două valori ale seriei, în timp ce calculul indicatorilor sintetici se bazează pe toate valorile.

Cei mai importanți **indicatori simpli** sunt:

- amplitudinea și
- abaterile individuale.

Indicatorii sintetici ai variației sunt:

- abaterea medie;
- dispersia;

- abaterea standard;
- coeficientul de variație.

Obiectul acestei secțiuni constă în prezentarea indicatorilor simpli ai variației.

11.2.1. Amplitudinea

Amplitudinea variației reprezintă distanța dintre valorile extreme ale seriei. Este un indicator ușor de calculat, iar semnificația lui este evidentă. Capacitatea lui de a caracteriza gradul de împrăștiere este însă foarte redusă din cauză că nu are în vedere și celelalte valori ale seriei.

Se calculează amplitudinea absolută și amplitudinea relativă.

Amplitudinea absolută a variației se determină ca diferență între valoarea maximă și cea minimă a caracteristicii:

$$A = x_{\max} - x_{\min}$$

și se exprimă în unitatea de măsură – naturală, natural-convențională sau valorică - a caracteristicii.

Amplitudinea absolută poate fi utilizată numai pentru a caracteriza variația a două sau mai multe serii formate din valorile aceleiași caracteristici, exprimate în aceeași unitate de măsură și referitoare la colectivități de dimensiuni apropiate. Se poate compara, de exemplu, amplitudinea absolută a nivelului pensiilor (exprimate în lei sau mii lei) realizate într-o anumită lună de pensionarii din două sau mai multe județe, dar nu se poate face o comparație similară cu privire la nivelul pensiilor realizate de pensionarii din două sau mai multe țări (exprimate în unități monetare diferite), așa cum nu se poate compara amplitudinea absolută a variației salariilor cu cea a variației productivității muncii, chiar dacă analiza se referă la aceeași colectivitate statistică și ambii indicatori sunt exprimați în mii sau milioane de lei.

Amplitudinea relativă a variației se calculează prin raportarea amplitudinii absolute la media caracteristicii:

$$A\% = \frac{A}{\bar{x}} \cdot 100$$

și se exprimă procentual.

Amplitudinea relativă poate fi utilizată, în anumite limite, pentru a compara variația a două caracteristici diferite exprimate în unități de măsură diferite sau variația aceleiași caracteristici exprimată în unități de măsură diferite. Dacă amplitudinea relativă a variației caracteristicii X este mai mare decât amplitudinea relativă a caracteristicii Y referitoare la aceeași colectivitate, se poate aprecia că gradul de împrăștiere (sau de inegalitate) a colectivității este mai mare din perspectiva

caracteristicii X decât din cea a caracteristicii Y. De exemplu, dacă amplitudinea relativă a distribuției veniturilor populației din România este mai mare decât cea a veniturilor din Germania, se poate formula ipoteza că inegalitatea veniturilor este mai mare în România decât în Germania.

11.2.1. Abaterile individuale

Abaterile individuale față de medie reprezintă diferențe dintre valorile caracteristicii și media valorilor, distanța la care se situează fiecare unitate a colectivității față de valoarea centrală a caracteristicii. Se pot calcula tot atâtea abateri individuale câte valori are seria statistică.

Abaterile individuale sunt absolute și relative.

Abaterile absolute se calculează ca diferență între valorile caracteristicii și media aritmetică și se exprimă în unitatea de măsură a caracteristicii:

$$d_i = x_i - \bar{x}$$

Abaterile relative se calculează ca raport între abaterile absolute și medie și sunt exprimate procentual:

$$d_i\% = \frac{d_i}{\bar{x}} \cdot 100$$

O parte a abaterilor individuale – absolute și relative – sunt pozitive, o altă parte sunt negative, suma lor fiind egală cu zero. Pentru analiză este importantă evidențierea **abaterii maxime**:

- abatere maximă pozitivă reprezintă distanța de la medie la cea mai mare valoare din serie:

$$d_{\max.(+)} = x_{\max} - \bar{x}$$

- abatere maximă negativă reprezintă distanța de la medie la cea mai mică valoare din serie:

$$d_{\max.(-)} = x_{\min} - \bar{x}$$

- Amplitudinea variației reprezintă distanța dintre valorile extreme ale seriei. Este un indicator ușor de calculat, iar semnificația lui este evidentă. Capacitatea lui de a caracteriza gradul de împrăștiere este însă foarte redusă din cauză că nu are în vedere și celelalte valori ale seriei.
- Abaterile individuale față de medie reprezintă diferențe dintre valorile caracteristicii și media valorilor, distanța la care se situează fiecare unitate a colectivității față de valoarea centrală a caracteristicii. Se pot calcula tot atâtea abateri individuale câte valori are seria statistică.



11.3. Rezumat

Indicatorii variației, denumiți și indicatori ai dispersiei sau ai împrăștierii, caracterizează modul în care valorile caracteristicii gravitează în jurul valorii centrale, gradul de împrăștiere a unităților colectivității în raport cu o anumită caracteristică, respectiv gradul de omogenitate al colectivității din punctul de vedere al acestei caracteristici.

În descrierea variației unei serii statistice se utilizează indicatori simpli și indicatori sintetici ai variației.



11.4. Test de autoevaluare a cunoștințelor

1) Dacă amplitudinea relativă a distribuției veniturilor populației din România este mai mare decât cea a veniturilor din Austria, se poate formula ipoteza că:

- nu există inegalitate a veniturilor în Austria
- inegalitatea veniturilor în România este mai mare decât în Austria
- inegalitatea veniturilor este mai mică în România decât în Austria
- inegalitatea veniturilor este aceeași în România și Austria
- nu există inegalitate a veniturilor în România.
- Alegeți varianta corectă.

2) Determinați abaterea individuală a exportului celei de a patra firme:

Firma	1	2	3	4	5	6
Valoarea exportului (mii. lei)	2600	6800	1450	2320	3970	825

- a. 2320,00 mii lei
- b. -674,16 mii lei
- c. 674,16 mii lei
- d. -1554,16 mii lei
- e. -394,16 mii lei.

Alegeți varianta corectă.

3) Abaterile individuale ale valorii mijloacelor fixe active înregistrate în cazul a 10 firme din orașul B sunt prezentate în tabelul de mai jos:

Valoarea mijloacelor fixe active (mil. lei)	Abaterile individuale	
x_i	absolute d_i	relative $d_i\%$
30	-17,2	-36,44
44	-3,2	-6,78
30	-17,2	-36,44
26	-21,2	-44,92
104	56,8	120,34
18	-29,2	-61,86
6	-41,2	-87,29
80	32,8	69,49
69	21,8	46,19
65	17,8	37,71

- a. abaterea individuală absolută maximă pozitivă este 56,8 mil. lei
- b. abaterea individuală absolută maximă negativă este -41,2 mil. lei
- c. abaterea individuală relativă maximă pozitivă este 120,34%
- d. abaterea individuală relativă maximă pozitivă este -87,29%
- e. toate variantele de mai sus.

Alegeți varianta corectă.

4) Determinați abaterea individuală maximă pozitivă a exportului:

Firma	1	2	3	4	5	6
Valoarea exportului (mii. lei)	2600	6800	1450	2320	3970	825

- a. 3805,84 mii lei
- b. 6800 mii lei
- c. 674,16 mii lei
- d. 1554,16 mii lei
- e. 394,16 mii lei.

Alegeți varianta corectă.

5) Abaterile individuale relative ale valorii vânzărilor lunare înregistrate în cazul a 5 firme din orașul B sunt prezentate în tabelul de mai jos:

Valoarea vânzărilor (mii lei) x_i	$d_i\%$
18	-62,18
6	-87,39
80	---
69	44,96
65	36,55

Abaterea relativă a valorii vânzărilor lunare înregistrate de cea de a treia firmă este de:

- a. 80,2%
- b. 34,7%
- c. 44,8%
- d. 66,9%
- e. 60,07%

Alegeți varianta corectă.

6) În cazul a 8 întreprinderi industriale s-au înregistrat următoarele date privind volumul producției:

Producția (mii lei lunar)	44	48	34	89	70	6	12	18	100	65
------------------------------	----	----	----	----	----	---	----	----	-----	----

Abaterea absolută a variației producției, calculată pentru firma care a înregistrat un volum al producției de 89 mii lei este egală cu:

- a. -16,75
- b. 28,25
- c. 9,25
- d. 4,65
- e. -54,75

Alegeți varianta corectă.

7) Determinați amplitudinea absolută a exportului:

Firma	1	2	3	4	5	6
Valoarea exportului (mii. lei)	26	68	14	23	39	82

- a. 38 mii lei
- b. 68 mii lei
- c. 67 mii lei
- d. 15 mii lei
- e. 39 mii lei.

Alegeți varianta corectă.

8) În cazul a 10 firme de transport s-au înregistrat următoarele date consumul lunar de combustibil:

Consumul de combustibil (tone)	4,4	4,8	3,4	8,9	7,0	6,0	1,2	1,8	10,0	6,5
--------------------------------	-----	-----	-----	-----	-----	-----	-----	-----	------	-----

Amplitudinea relativă a variației consumului de combustibil este egală cu:

- a. 162,96%
- b. 28,25%
- c. 123,98%
- d. 4,65%
- e. 54,75%

Alegeți varianta corectă.

9) Se cunosc următoarele date privind timpul de muncă a cinci din șase salariați ai unei firme (ore lunar):

160	156	---	46	108	120
-----	-----	-----	----	-----	-----

Determinați numărul de ore lucrate de al treilea salariat, dacă abaterea absolută este de 6 ore.

- a. 125
- b. 62
- c. 123
- d. 96
- e. 154

Alegeți varianta corectă.

10) Determinați abaterea individuală maximă negativă a importului:

Firma	1	2	3	4	5	6
Valoarea importului (lei)	42600	26800	11450	2320	23970	1825

- a. -16338,5
- b. -394,16
- c. 16338,5
- d. -16335,8
- e. 394,16.

Alegeți varianta corectă.



11.5. Bibliografie recomandată

- z. Caragea Nicoleta, *Statistica- Concepte si metode de analiza a datelor*, Editura MUSTANG, București, 2015
- aa. Voineagu, V. , *Statistică –baze teoretice și aplicații*, Editura Economică, București, 2007
- bb. Georgescu-Roegen N., *Metoda statistică*, Editura Expert, București, 1998.
- cc. Andrei T., Stancu S., Pele D.T., *Statistică. Teorie și aplicații* (ediția a II-a), Editura Economică, București, 2003.
- dd. Biji M., Biji E.M., Lilea E., Anghelache C., *Tratat de statistică*, Editura Economică, București, 2002.

Unitatea de învățare 12.

INDICATORII SINTETICI AI VARIAȚIEI



12.1. Introducere

În cadrul acestui capitol, studenților le sunt prezentați indicatorii sintetici ai variației. Definirea indicatorilor simpli ai variației: amplitudinea, abaterea individuale, abaterea medie, dispersia, abaterea standard, coeficientul de variație. Prezentarea algoritmului de calcul pentru determinarea indicatorilor sintetici ai variației unei serii de frecvențe. Deprinderea abilităților de interpretare a rezultatelor

Durata medie de parcurgere a acestui capitol este de aproximativ o oră.



12.2. Conținut

Indicatorii sintetici ai variației (împrăștierii) sintetizează într-o singură expresie numerică variația valorilor individuale față de tendința centrală a caracteristicilor urmărite, într-o populație statistică. Principalii indicatori sintetici cu care se caracterizează împrăștierea (varianța) termenilor seriei față de tendința lor centrală sunt: abaterea medie, dispersia, abaterea standard (sau abaterea medie pătratică) și coeficientul de variație. Diferența dintre indicatori derivă din modalitatea de rezolvare a unei probleme care apare la calculul mediei abaterilor, și anume faptul că suma abaterilor este nulă. La baza determinării indicatorilor sintetici ai variației stau abaterile individuale.

12.2.1. Abaterea medie

Abaterea medie, denumită și abatere absolută medie sau abatere medie liniară, se calculează ca o medie aritmetică simplă sau ponderată a abaterilor individuale în valoare absolută.

Formulele de calcul ale abaterii medii sunt următoarele:

- pentru serii simple, formate din valori individuale

$$d = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}$$

- pentru seriile de frecvențe absolute

$$d = \frac{\sum_{i=1}^k |x_i - \bar{x}| n_i}{\sum_{i=1}^k n_i}$$

Evident, calculul indicatorilor seriilor de distribuție de frecvențe se face numai pe baza uneia dintre cele trei formule, alegerea ei fiind în funcție de datele disponibile, adică de modul de exprimare a frecvențelor. Este important de subliniat, de asemenea, că atunci când sunt disponibile datele individuale este mai corectă utilizarea formulei destinate seriilor simple.

12.2.2. Dispersia

Dispersia, denumită și varianță, este media aritmetică simplă sau ponderată a pătratelor abaterilor individuale ale tuturor termenilor seriei față de media lor.

Formula de calcul a dispersiei

- a) pentru o serie simplă, formată din valori individuale:

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

- b) pentru o serie de frecvențe absolute:

$$\sigma^2 = \frac{\sum_{i=1}^k (x_i - \bar{x})^2 n_i}{\sum_{i=1}^k n_i}$$

Fiind definită și calculată ca medie a pătratelor unor valori, dispersiei nu i se poate atribui o unitate de măsură concretă, iar semnificația ei este dificil de intuit în termeni practici. Din acest motiv, dispersia nu este utilizată efectiv ca indicator menit să caracterizeze direct variația unei caracteristici, mai ales că dependența sa de unitatea de măsură și de ordinul de mărime a valorilor caracteristicii o fac improprie evaluărilor comparative privind gradul de împrăștiere a două sau mai multe serii de date. Dispersia reprezintă însă un parametru fundamental în statistica matematică, fiind larg utilizat în construcția și aplicarea unor importante metode de analiză statistică (de exemplu, în verificarea ipotezelor statistice, în evaluarea influenței factorilor, în extinderea rezultatelor cercetărilor selective).

12.2.3. Abaterea standard

Abaterea standard⁸, numită și abatere medie pătratică sau abatere tip, este o medie pătratică – simplă sau ponderată - a abaterilor individuale ale tuturor valorilor seriei față de media lor aritmetică. Indicatorul măsoară distanța medie dintre valorile seriei și medie cu ajutorul unuia dintre tipurile speciale de medie, și anume: media pătratică.

Formulele de calcul ale abaterii standard sunt următoarele:

a) pentru o serie simplă

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

b) pentru o serie de frecvențe absolute

$$\sigma = \sqrt{\frac{\sum_{i=1}^k (x_i - \bar{x})^2 n_i}{\sum_{i=1}^k n_i}}$$

Se remarcă faptul că abaterea standard derivă din dispersie (σ^2), care este de fapt pătratul abaterii standard. În aceste condiții, deși semnificația sa este mai dificil de intuit, indicatorul este mai frecvent utilizat în aplicații statistice decât abaterea medie. Dezavantajul constă în faptul că abaterea standard se exprimă în unitatea de măsură a variabilei pentru care se calculează, făcând imposibilă compararea variațiilor a două caracteristici exprimate în unități de măsură diferite.

Ca și abaterea medie, abaterea standard se exprimă în aceeași unitate de măsură ca și caracteristica a cărei variație o reprezintă. Ambele abateri au valori pozitive. Ele arată o variație cu atât mai mare cu cât sunt mai îndepărtate de zero.

Abaterea nulă, adică $\sigma = 0$ sau $d = 0$, indică absența variației, ceea ce înseamnă că toate valorile individuale care compun seria sunt egale.

Abaterea standard este întotdeauna mai mare decât abaterea medie:

$$\sigma > d$$

⁸ En. Standard deviation

12.2.4. Coeficientul de variație

Coeficientul de variație reprezintă abaterea standard în expresie relativă. Se calculează ca raport între abaterea standard și media aritmetică și este exprimat procentual:

$$v = \frac{\sigma}{\bar{X}} \cdot 100$$

Spre deosebire de ceilalți indicatori sintetici ai variației, care nu permit compararea gradului de împrăștiere a două sau mai multe serii diferite, formate din valori exprimate în unități de măsură diferite sau având ordine de mărime diferite, coeficientul de variație poate fi utilizat și pentru astfel de comparații. Dacă în cazul unei serii coeficientul de variație este mai mare decât în cazul altei serii, aceasta înseamnă că prima serie este caracterizată printr-un grad de împrăștiere mai mare decât a doua.

Coeficientul de variație ia valori pozitive, $v \geq 0$, fiind egal cu zero dacă toate valorile individuale sunt egale. Cu cât are o valoare mai apropiată de zero cu atât seria este mai omogenă, indicând lipsa variației, iar media este un indicator reprezentativ pentru seria respectivă. Se apreciază⁹ că un coeficient de variație de peste 35-40 % indică faptul că seria nu este omogenă, iar media nu este reprezentativă.

O variantă a coeficientului de variație derivă din abaterea medie:

$$v'_d = \frac{d}{\bar{X}} \cdot 100$$

Să ne reamintim

Indicatorii sintetici ai variației (împrăștierii) sintetizează într-o singură expresie numerică variația valorilor individuale față de tendința centrală a caracteristicilor urmărite, într-o populație statistică. Principalii indicatori sintetici cu care se caracterizează împrăștierea (varianța) termenilor seriei față de tendința lor centrală sunt: abaterea medie, dispersia, abaterea standard (sau abaterea medie pătratică) și coeficientul de variație. Diferența dintre indicatori derivă din modalitatea de rezolvare a unei probleme care apare la calculul mediei abaterilor, și anume faptul că suma abaterilor este nulă. La baza determinării indicatorilor sintetici ai variației stau abaterile individuale.

⁹ conform statisticii experimentale



12.3. Rezumat

- Abaterea medie, denumită și abatere absolută medie sau abatere medie liniară, se calculează ca o medie aritmetică simplă sau ponderată a abaterilor individuale în valoare absolută.
- Dispersia, denumită și varianță, este media aritmetică simplă sau ponderată a pătratelor abaterilor individuale ale tuturor termenilor seriei față de media lor.
- Abaterea standard, numită și abatere medie pătratică sau abatere tip, este o medie pătratică – simplă sau ponderată - a abaterilor individuale ale tuturor valorilor seriei față de media lor aritmetică.
- Coeficientul de variație este un indicator propus de Pearson cu scopul de a înlătura unitatea de măsură a variației, reprezentând abaterea standard în expresie relativă.



12.4. Test de autoevaluare a cunoștințelor

1) Principalii indicatori sintetici cu care se caracterizează împrăștierea termenilor seriei față de tendința lor centrală sunt:

- a. dispersia și coeficientul de variație
- b. abaterea medie, dispersia, abaterea medie pătratică (sau abaterea standard) și coeficientul de variație
- c. amplitudinea și abaterile individuale
- d. indicatorii medii
- e. indicatorii frecvenței

Alegeți varianta corectă.

2) La baza calculului indicatorilor sintetici ai variației unei serii de distribuție se regăsesc:

- a. abaterile individuale, media și frecvențele
- b. media
- c. frecvențele (absolute sau relative)
- d. toate valorile seriei
- e. frecvențele cumulate

Alegeți varianta corectă.

3) Abaterea medie:

- a. este denumită și abatere absolută medie
- b. se calculează ca o medie aritmetică simplă sau ponderată a abaterilor individuale în valoare absolută
- c. este denumită și abatere medie liniară
- d. a)+b)+c)
- e. este denumită și abatere standard

Alegeți varianta corectă.

4) Abaterea standard:

- a. este denumită și abatere medie pătratică
- b. este denumită și abatere tip
- c. este o medie pătratică – simplă sau ponderată - a abaterilor individuale ale tuturor valorilor seriei față de media lor aritmetică
- d. măsoară distanța medie dintre valorile seriei și medie cu ajutorul unuia dintre tipurile speciale de medie, și anume: media pătratică
- e. toate variantele de mai sus

Alegeți varianta corectă.

5) Dispersia:

- a. este denumită și varianță
- b. este media aritmetică simplă sau ponderată a pătratelor abaterilor individuale ale tuturor termenilor seriei față de media lor
- c. este pătratul abaterii standard
- d. este definită și calculată ca medie a pătratelor unor valori, din acest motiv dispersiei nu i se poate atribui o unitate de măsură concretă
- e. are o semnificație ușor de intuit în termeni practici

Care dintre variante nu este corectă?

6) Coeficientul de variație:

- a. se calculează ca raport între abaterea standard și media aritmetică și se exprimă procentual
- b. permite compararea gradului de împrăștiere a două sau mai multe serii diferite, formate din valori exprimate în unități de măsură diferite sau având ordine de mărime diferite
- c. ia valori pozitive, $v \geq 0$, fiind egal cu zero dacă toate valorile individuale sunt egale

- d. se apreciază că un coeficient de variație de peste 35-40 % indică faptul că seria nu este omogenă, iar media nu este reprezentativă
 - e. toate variantele de mai sus
- Alegeți varianta corectă.

7) Coeficientul de variație al unei serii de distribuție este egal cu 73%. În această situație:

- a. seria nu este omogenă
- b. gradul de împrăștiere al seriei este mai mare decât limita de reprezentativitate admisă de statistica experimentală
- c. media nu este reprezentativă
- d. seria prezintă un nivel înalt de variație
- e. toate variantele de mai sus

Alegeți varianta corectă.

8) Coeficientul de variație al unei serii de distribuție este egal cu 24%, media seriei fiind de 30 unități.

Dispersia are valoarea:

- a. 7,25
- b. 51,84
- c. nu se poate determina
- d. 24,65
- e. 44,94

Alegeți varianta corectă.

9) Coeficientul de variație al unei serii de distribuție este egal cu 12%, media seriei fiind de 400 unități. Abaterea standard are valoarea:

- a. 2304
- b. 450
- c. nu se poate determina
- d. 24,65
- e. 48

Alegeți varianta corectă.

1. În cazul unei serii de distribuție se cunosc următoarele informații: abaterea standard este egală cu 5,6 unități, iar media de 22 unități. Coeficientul de variație are valoarea:

- a. 23,04%
- b. 45,02%
- c. 25,45%
- d. 24,65

e. 48,02

Alegeți varianta corectă.



12.5. Bibliografie recomandată

- a. Caragea Nicoleta, *Statistica- Concepte si metode de analiza a datelor*, Editura MUSTANG, București, 2015
- b. Voineagu, V. , *Statistică –baze teoretice și aplicații*, Editura Economică, București, 2007
- c. Georgescu-Roegen N., *Metoda statistică*, Editura Expert, București, 1998.
- d. Andrei T., Stancu S., Pele D.T., *Statistică. Teorie și aplicații* (ediția a II-a), Editura Economică, București, 2003.
- e. Biji M., Biji E.M., Lilea E., Anghelache C., *Tratat de statistică*, Editura Economică, București, 2002.

12.6. Teme de control

Tema 3 – Capitolul 11-12: „Realizați un referat de maximum 5 pagini în care să calculați indicatorii simpli și sintetici ai unei serii de distribuție”.

Unitatea de învățare 13.

MĂRIMILE RELATIVE



13.1. Introducere

În cadrul acestui capitol, studenților le sunt prezentate mărimile relative:

- definirea indicatorilor statistici primari și derivați
- prezentarea principalelor tipuri de mărimi relative
- deprinderea utilizării corecte a mărimilor relative

Durata medie de parcurgere a acestui capitol este de aproximativ o oră.



13.2. Conținut

Mărimile relative constituie una dintre principalele categorii de indicatori statistici utilizați în analiza economică. Indicatorii statistici reprezintă expresii numerice ale unor fenomene, procese sau categorii economico-sociale, definite în timp, în spațiu și instituțional-organizatoric.

Statistica operează cu indicatori primari și derivați.

Indicatorii primari sunt mărimi absolute care exprimă nivelul caracteristicii cercetate, obținându-se prin înregistrare directă sau prin însumarea datelor individuale. Indicatorii primari se exprimă în unități de măsură, naturale sau monetare, specifice caracteristicii observate: bucăți, tone, kilograme, metri, metri pătrați, dolari, Euro, lei, mii lei, milioane lei etc. Un astfel de indicator este, de exemplu: cifra de afaceri (în milioane lei) înregistrată de o firmă având ca principală activitate comercializarea carburanților și cea rezultată prin însumare la nivelul întregii țări; volumul producției la nivelul unei firme sau la nivel național, numărul populației unei localități, a unui județ și a țării; numărul salariaților unei unități economice (societate comercială sau instituție publică), al celor dintr-o anumită ramură de activitate sau dintr-o unitate teritorială (localitate sau județ) și al tuturor salariaților din economia națională; veniturile și cheltuielile bugetelor de asigurări sociale (în miliarde lei); volumul exportului sau importului realizat la o anumită grupă de produse, cu o anumită țară sau cu toate țările (în milioane dolari sau Euro).

Indicatorii derivați se obțin în urma prelucrării datelor primare, prin aplicarea diferitelor metode și procedee de calcul statistic. Există o mare diversitate de indicatorii derivați, cu metode variate de calcul, de la operații simple de împărțire sau scădere la calcule bazate pe formule complexe. În categoria indicatorilor derivați intră mediile, quantilele, indicatorii variației și ai inegalității, indicatorii corelației, indicii etc. O parte a acestora au caracter absolut și se exprimă în unități de măsură naturale sau monetare specifice caracteristicilor care au stat la baza calculului lor. Altele de exprimă sub formă de coeficienți sau procente. Mărimile relative formează una dintre categoriile de indicatori derivați cele mai frecvent utilizate în analiză.

Mărimile relative rezultă din compararea a doi indicatori, prin raportarea unuia, numit termen de comparat, la altul numit bază de comparație. Rezultatul obținut este un număr întreg sau fracționar, care arată câte unități din indicatorul de comparat revin la o unitate din indicatorul bază de raportare. În cazul celor mai multe mărimi relative, rezultatul raportului capătă relevanță în analiză prin multiplicarea cu 10^k (respectiv cu 10, 100, 1000, 10000, 100000,, corespunzător lui $k = 1, 2, 3, 4, 5, \dots$).

Mărimile relative calculate ca raport a doi indicatori absoluți similari, exprimați în aceeași unitate de măsură, iau forma unor coeficienți rezultați direct din împărțire sau se exprimă sub formă de procente (%), dacă rezultatul raportului este înmulțit cu 100), promile (‰, dacă rezultatul este multiplicat cu 1000), prodecimile (în cazul înmulțirii cu 10000) sau procentimile (în cazul înmulțirii cu 100000).

Mărimile relative calculate ca raport a doi indicatori diferiți, exprimați în unități de măsură diferite, sunt exprimate în unități de măsură compuse din unitățile de măsură ale celor doi indicatori. Dacă asigurarea relevanței rezultatului impune multiplicarea raportului cu 100, 1000 etc., aceasta este reflectată în unitatea de măsură a indicatorului rezultat.

Principalele tipuri de mărimi relative sunt: mărimile relative de structură, mărimile relative de coordonare, mărimile relative de intensitate, mărimile relative ale dinamicii, mărimile relative ale planului.

13.2.1. Mărimi relative de structură

Mărimile relative de structură reprezintă raportul dintre un element sau un grup de elemente ale unei colectivități și întreaga colectivitate, calculat pe baza unei caracteristici de volum sau a frecvenței. Mărimea relativă de structură este raportul dintre o parte și întregul de care aparține această parte.

Mărimile relative de structură sunt exprimate, în principal, sub formă procentuală, indicând câte unități revin părții analizate la fiecare 100 de unități ale întregului. Dacă sunt exprimate sub formă

de coeficienți, acestea arată ce fracțiune din întreg revine părții analizate (0,25, adică 25 de sutimi un sfert; 0,10 sau o zecime; 0,33(3) sau o treime etc).

Dacă se calculează pe baza unei caracteristici de volum (volumul producției, al cheltuielilor de producție, al profitului, al importului sau exportului, al veniturilor etc), mărimea relativă de structură poartă denumirea de pondere sau greutate specifică. Aceasta exprimă raportul în care se află volumul aferent părții analizate (x_i) față de volumul însumat al tuturor părților care compun întregul:

$$g = \frac{x_i}{\sum_{i=1}^n x_i} \times 100$$

Observații:

- Suma ponderilor tuturor componentelor este egală cu 100%.
- Diferența dintre doi indicatori calculați sub formă procentuală este exprimată în puncte procentuale (p.p.).

Exemple:

1. Producția industrială realizată în anul t , pe principalele grupe de activități, a înregistrat următoarele valori, conform tabelului de mai jos:

Producția industrială în anul t

	<i>Q (mil.lei)</i>
Industria extractivă	55403,4
Industria prelucrătoare	769938,6
Energie electrică și termică, gaze și apă	141103,0
<i>Total</i>	966445,0

În anul t , ponderea industriei extractive în volumul total al producției industriale a fost de:

$$g_1 = \frac{x_i}{\sum_{i=1}^n x_i} \times 100 = \frac{55403,4}{966445} \times 100 = 5,7\%$$

Ponderile celorlalte două grupe de activități se calculează în același mod, astfel încât structura producției industriale are următoarea configurație:

Structura producției industriale în anul t

	g (%)
Industria extractivă	5,7
Industria prelucrătoare	79,7
Energie electrică și termică, gaze și apă	14,6
<i>Total</i>	100,0

Mărimile relative de structură calculate pentru seriile de distribuție de frecvențe, numite **frecvențe relative** sau **ponderi**, exprimă raportul în care se află numărul unităților din fiecare grupă (n_i) față de numărul unităților din întreaga colectivitate.

2. În anul t , s-a înregistrat următoarea distribuție a salariaților după nivelul salariilor:

Distribuția salariaților după nivelul salariului brut, în anul t

	Grupe de salarii brute realizate, lei	Pondere în numărul total de salariați, %
1	Până la 350	23,6
2	350-400	8,1
3	400-500	12,6
4	500-700	17,9
5	700-1000	16,1
6	1000-1500	12,5
7	1500-2000	4,6
8	2000-3000	2,7
9	3000-4000	1,0
10	4000-5000	0,4
11	Peste 5000	0,5
	Total	100,0

Dacă numărul total al salariaților a fost de 3814 mii în anul t , poate fi estimat și numărul celor cuprinși în prima grupă (în mii):

$$n_1 = \frac{n'_1}{100} \times N = \frac{23,6}{100} \times 3814 = 900$$

În mod similar se determină și numărul salariaților cuprinși în celelalte grupe. De exemplu, pentru grupa a unsprezecea:

$$n_{11} = \frac{n'_{11}}{100} \times N = \frac{0,5}{100} \times 3814 = 19,07$$

Așadar, numărul salariaților cu un salariu mai mare de 5000 este de 19 mii.

13.2.2. Mărimi relative de intensitate

Mărimile relative de coordonare se utilizează pentru a compara două niveluri ale aceluiași indicator, niveluri care privesc două grupe/părți (A și B) ale aceleiași populații statistice sau două populații (A și B) de același tip situate în spații diferite.

Calculul mărimilor relative de coordonare se poate face raportând nivelul înregistrat în cazul populației/grupe A la cel înregistrat în cazul populației/grupe B sau, invers, raportând nivelul aferent populației B la cel corespunzător populației/grupe A:

$$k_{A/B} = \frac{x_A}{x_B}$$

$$k_{B/A} = \frac{x_B}{x_A}$$

Rezultatul este supraunitar într-una dintre variantele de calcul și subunitar în cealaltă. Dacă $x_A > x_B$, raportul $k_{A/B}$ este supraunitar și arată de câte ori este mai mare nivelul corespunzător populației A, comparativ cu B. În acest caz, raportul invers $k_{B/A}$ este subunitar, indicând ce fracțiune reprezintă nivelul înregistrat de populația B în raport cu nivelul A. Indicatorul subunitar este mai relevant dacă este exprimat procentual.

Exemple:

1. În anul t, raportul dintre produsul intern brut pe locuitor al Italiei și cel al României (calculate în dolari la paritatea puterii de cumpărare) a fost de 3,3.

$$k_{Italia/Romania} = \frac{PIB_{Italia}}{PIB_{Romania}}$$

$$k_{Italia/Romania} = \frac{23300}{7100} = 3,3 \text{ indicând un nivel de bunăstare de 3,3 ori mai înalt în Italia decât în România.}$$

Raportul invers:

$$k_{Romania/Italia} = \frac{PIB_{Romania}}{PIB_{Italia}}$$

$$k_{Romania/Italia} = \frac{700}{23300} = 0,304 \text{ arată că nivelul de bunăstare înregistrat în România reprezintă}$$

numai 30,4% în raport cu cel atins de Italia.

13.2.3. Mărimi relative de coordonare

Mărimile relative de intensitate se determină prin raportarea a doi indicatori absoluți (x și y) de natură diferită, dar care se află într-o relație de interdependență. În calculul indicatorului relativ, rezultatul raportului este multiplicat cu 10^k , dacă această operație este necesară pentru asigurarea relevanței:

$$r = \frac{x}{y} \times 10^k$$

Indicatorii utilizați în calculul mărimilor relative de intensitate sunt exprimați, în general, în unități de măsură diferite, iar rezultatul este exprimat într-o unitate de măsură compusă.

Din categoria mărimilor relative de intensitate fac parte următorii indicatori:

- a) **productivitatea muncii (w)**, calculată ca raport între volumul fizic sau valoarea producției (Q) și consumul de muncă (T, dacă este exprimat în unități de timp de muncă; N, dacă este exprimat în număr de persoane ocupate/salariați):

$$w = \frac{Q}{N}$$

$$w = \frac{Q}{T}$$

sau ca raport între consumul de muncă (T) și volumul producției (Q):

$$w = \frac{T}{Q}$$

- b) **consumul specific (c)**, calculat ca raport între consumul total de materii prime, materiale și energie (C) și volumul producției (Q):

$$w = \frac{C}{Q}$$

- c) **randamentele în agricultură**, calculate ca raport între producția vegetală realizată și suprafețele cultivate cu diferite culturi (kilograme de grâu sau de porumb boabe pe hectar, de exemplu) sau între producția animală și numărul animalelor;
- d) **indicatorii demografici**: ratele natalității și mortalității, nupțialității și divorțialității, calculate ca raport (multiplicat cu 1000) între numărul copiilor născuți vii, al deceselor, al căsătoriilor și divorțurilor înregistrate în cursul unui an și numărul populației; rata mortalității infantile și juvenile, calculată ca raport (multiplicat cu 1000) între numărul

copiilor sub un an, respectiv sub 5 ani, decedați și numărul copiilor născuți vii în cursul unui an etc.

O altă categorie de mărimi relative de intensitate este formată din cele calculate pe baza a doi indicatori diferiți sub aspectul conținutului economic, dar exprimați în aceeași unitate de măsură.

Indicatorul relativ rezultat din calcul are forma unei rate procentuale:

- rata de activitate și rata de ocupare, calculate ca raport între populația activă, respectiv ocupată și populația în vârstă de muncă;
- rata șomajului, calculată ca raport între numărul șomerilor și numărul populației active;
- rata sărăciei, calculată ca raport între populația săracă și populația totală;
- rata rentabilității;
- rata dobânzii etc.

13.2.4. Mărimi relative ale dinamicii și ale planului

Mărimile relative ale dinamicii se utilizează pentru caracterizarea evoluției în timp a fenomenelor social-economice. Cele mai importante componente ale acestei categorii de mărimi relative sunt indicii și ritmurile.

Indicii dinamicii se determină ca raport între nivelurile înregistrate de o variabilă într-o anumită perioadă sau la un moment dat (x_t) și într-o perioadă sau la un moment anterior, considerat bază de comparație (x_0):

$$i_{t/0} = \frac{x_t}{x_0} \times 100$$

Rezultatul raportului arată de câte ori a crescut nivelul variabilei în perioada t față de perioada de bază (dacă sunt supraunitari) sau la ce fracțiune din nivelul de bază a scăzut nivelul variabilei (dacă sunt subunitari). Forma de prezentare a indicilor este cea procentuală.

Ritmuri reprezintă raportul dintre sporul absolut al variabilei în perioada t față de perioada de bază și nivelul înregistrat în perioada de bază:

$$r_{t/0} = \frac{x_t - x_0}{x_0} \times 100$$

și arată cu cât a crescut sau a scăzut nivelul variabilei în raport cu nivelul de bază.

Mărimile relative ale planului sunt utilizate pentru analiza relației dintre nivelul înregistrat de un indicator în perioada de bază (x_0), cel programat (x_{pl}) și cel realizat în perioada curentă (x_1). Se utilizează doi indici ai planului, ambii calculați ca raport între nivelurile indicatorului analizat:

- indicele sarcinii de plan, determinat prin raportarea nivelului planificat la cel realizat în perioada de bază:

$$i_{plan/0} = \frac{x_{plan}}{x_0} \times 100$$

- indicele realizării planului, determinat prin raportarea nivelului realizat la cel planificat pentru perioada curentă:

$$i_{1/plan} = \frac{x_1}{x_{plan}} \times 100$$

Cei doi indici arată proporția modificării programate a fenomenului analizat, respectiv proporția în care a fost realizat nivelul programat. Ambii indici ai planului se calculează sub formă procentuală.

Să ne reamintim

Mărimile relative constituie una dintre principalele categorii de indicatori statistici utilizați în analiza economică. Indicatorii statistici reprezintă expresii numerice ale unor fenomene, procese sau categorii economico-sociale, definite în timp, în spațiu și instituțional-organizatoric.



13.3. Rezumat

Mărimile relative rezultă din compararea a doi indicatori, prin raportarea unuia, numit termen de comparat, la altul numit bază de comparație. Rezultatul obținut este un număr întreg sau fracționar, care arată câte unități din indicatorul de comparat revin la o unitate din indicatorul bază de raportare. În cazul celor mai multe mărimi relative, rezultatul raportului capătă relevanță în analiză prin multiplicarea cu 10^k (respectiv cu 10, 100, 1000, 10000, 100000,, corespunzător lui $k = 1, 2, 3, 4, 5, \dots$).

Mărimile relative calculate ca raport a doi indicatori absoluți similari, exprimați în aceeași unitate de măsură, iau forma unor coeficienți rezultați direct din împărțire sau se exprimă sub formă de procente (%), dacă rezultatul raportului este înmulțit cu 100), promile (‰, dacă rezultatul este multiplicat cu 1000), prodecimile (în cazul înmulțirii cu 10000) sau procentimile (în cazul înmulțirii cu 100000).

Mărimile relative calculate ca raport a doi indicatori diferiți, exprimați în unități de măsură diferite, sunt exprimate în unități de măsură compuse din unitățile de măsură ale celor doi indicatori. Dacă asigurarea relevanței rezultatului impune multiplicarea raportului cu 100, 1000 etc., aceasta este reflectată în unitatea de măsură a indicatorului rezultat.

Principalele tipuri de mărimi relative sunt: mărimile relative de structură, mărimile relative de coordonare, mărimile relative de intensitate, mărimile relative ale dinamicii, mărimile relative ale planului.



13.4. Test de autoevaluare a cunoștințelor

1. Definiți indicatorii primari și indicatorii derivați.
2. Care sunt principalele mărimi relative?
3. Definiți greutatea specifică.
4. Prezentați formula de calcul a ponderii.
5. Definiți mărimile relative de intensitate. Exemple.
6. Prezentați formulele de calcul a productivității muncii.
7. Definiți mărimile relative de coordonare.
8. Care este unitatea de măsură a diferenței dintre două mărimi exprimate procentual?
9. Definiți mărimile relative ale dinamicii.
10. Prezentați formula de calcul a consumului specific.



13.5. Bibliografie recomandată

- a. Caragea Nicoleta, *Statistica- Concepte si metode de analiza a datelor*, Editura MUSTANG, București, 2015
- b. Voineagu, V. , *Statistică –baze teoretice și aplicații*, Editura Economică, București, 2007
- c. Georgescu-Roegen N., *Metoda statistică*, Editura Expert, București, 1998.
- d. Andrei T., Stancu S., Pele D.T., *Statistică. Teorie și aplicații* (ediția a II-a), Editura Economică, București, 2003.
- e. Biji M., Biji E.M., Lilea E., Anghelache C., *Tratat de statistică*, Editura Economică, București, 2002.

Unitatea de învățare 14.

ELEMENTE DE TEORIA SELECȚIEI



14.1. Introducere

În cadrul acestui capitol, studenților le sunt prezentate elemente de teoria selecției:

- definirea conceptelor de bază și a indicatorilor principali ce caracterizează cercetarea statistică selectivă
- evidențierea principalelor metode de selecție și a tipurilor de sondaj
- definirea conceptului de eroare a cercetării statistice selective
- prezentarea extinderii rezultatelor sondajului prin intervale de încredere și prin estimatii punctuale

Durata medie de parcurgere a acestui capitol este de aproximativ o oră.



14.2. Conținut

Pentru caracterizarea statistică a fenomenelor și proceselor social-economice de masă cu scopul evidențierii legităților specifice acestora, se folosesc date individuale obținute prin metode de înregistrare totală sau parțială. În practică, din diferite motive pentru anumite colectivități nu se poate face înregistrarea totală a unităților ce le compun, fie că ar fi necesitat cheltuieli foarte mari, fie că operativitatea obținerii rezultatelor cercetării nu este asigurată. Toate aceste considerente și altele nespicate, necesită tot mai frecvent organizarea unor înregistrări parțiale de tipul sondajelor statistice.

În cercetarea statistică selectivă se operează cu categorii distincte referitoare la colectivitatea totală și cea de selecție. Colectivitatea totală poartă și denumirea de colectivitate generală, colectivitate generală sau populație, iar colectivitatea de selecție este denumită și eșantion, mostră sau probă.

În general, populațiile statistice sunt **finite**. Statistica operează însă și cu populații ipotetic **infinite**. Acestea sunt populații cărora nu li se cunoaște dimensiunea (populația formată din copacii unei păduri sau producția de televizoare realizată de o firmă într-o perioadă nedefinită de timp) sau

populații atât de mari încât sub aspect statistic pot fi asimilate unora infinite (populații de dimensiuni atât de mari încât eliminarea unui anumit număr de unități nu are un efect sesizabil asupra structurii populației). Gospodăriile sau întreprinderile din țară sau dintr-o regiune, populația unui oraș sau pacienții tratați pentru boli cardiovasculare formează populații care pot fi tratate ca infinite.

Indicatorii care caracterizează populația totală, a căror estimare constituie obiectivul cercetării selective, sunt denumiți **parametri**, indicatorii calculați pe baza informațiilor colectate asupra unităților din eșantionul extras – **estimatori** sau **statistici**, iar valorile concrete ale acestora – **estimații**. Informațiile colectate prin cercetarea selectivă sunt utilizate pentru estimarea a diferiți indicatori, în funcție de scopul analizei.

Cei mai frecvent utilizați sunt:

Media

$$m = \frac{\sum_{i=1}^N x_i}{N} \quad \text{pentru colectivitatea totală și}$$

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad \text{pentru eșantion}$$

Dispersia

$$\sigma^2 = \frac{\sum_{i=1}^N (x_i - m)^2}{N} \quad \text{pentru colectivitatea totală și}$$

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} \quad \text{pentru eșantion}$$

Unde **N** reprezintă volumul colectivității totale (numărul de unități), iar **n** volumul eșantionului.

Estimarea parametrilor populației totale pe baza rezultatelor sondajului poartă denumirea de **inferență statistică**.

Cercetarea selectivă are trei componente: selecția unităților eșantionului, culegerea și prelucrarea informațiilor referitoare la eșantion (calculul indicatorilor specifici eșantionului) și inferența statistică (extinderea rezultatelor asupra colectivității totale, prin estimarea parametrilor acesteia).

14.2.1. Selecția

Rezultatele cercetării selective pot fi extinse la întreaga colectivitate numai dacă eșantionul este reprezentativ, ceea ce înseamnă că structura sa este identică cu structura colectivității generale sau diferă nesemnificativ de aceasta, iar indicatorii calculați pe baza eșantionului coincid cu parametrii populației totale. Un eșantion poate fi considerat satisfăcător sub aspectul reprezentativității dacă indicatorii de structură a populației de sondaj se abat cu maximum 5% față de parametrii corespunzători ai populației totale.

Asigurarea reprezentativității impune utilizarea unui eșantion suficient de mare, astfel încât acesta să poată reproduce structura populației, precum și selecția aleatoare a unităților eșantionului.

14.2.1.1. Metode de selecție

Principala metodă de selecție utilizată în cercetarea statistică este **selecția aleatoare**. Aceasta presupune alegerea unui eșantion de n unități din cele N unități ale populației după un procedeu care să asigure tuturor unităților colectivității totale aceeași șansă (probabilitate) de a fi selecționată în eșantion. Aceasta înseamnă, de asemenea, ca fiecare eșantion de volum n din mulțimea de eșantioane de același volum care pot fi extrase din populația de volum N să aibă aceeași probabilitate de selecție¹⁰.

Tehnic, selecția aleatoare poate fi realizată prin trei procedee: tragerea la sorți, utilizarea tabelelor cu numere aleatoare sau mecanic. Selecția prin tragere la sorți constă în extragerea treptată dintr-o urnă a unor bile, discuri, bilete cu numele sau numerele de înregistrare a unităților populației. Al doilea procedeu presupune numerotarea unităților populației și alegerea celor cu numerele de ordine cuprinse într-o porțiune sau alta din tabelul cu numere aleatoare (întâmplătoare)¹¹. În fine, selecția mecanică sau sistematică este alegerea unităților eșantionului din lista unităților populației, la un interval predeterminat (fiecare a zecea, a douăzecea, a o suta etc. unitate), numit pas de numărare și egal cu $N:n$.

În principiu, există două tipuri de selecție: cu sau fără revenire (și selecția nerepetată). Prima, numită și **selecție repetată** sau bernoulliană, presupune ca o unitate selectată să fie reintrodusă în populație înaintea unei noi extrageri, iar cea de-a doua, numită **selecție nerepetată** sau nebernoulliană, se realizează fără reintroducerea în populație a unei unități odată selectate.

¹⁰ Numărul eșantioanelor de volum n care pot fi extrase dintr-o populație de volum N este $k = C_n^N$.
De exemplu, dintr-o populație formată din 40 unități se pot extrage 658008 eșantioane diferite formate din 5 unități.

¹¹ Tabelele cu numere aleatoare cuprind coloane de numere formate din 4-5 cifre, generate la întâmplare, fără o regulă prestabilită.

Eșantioanele sunt constituite, în general, prin selecții nerepetate. Atunci când populația este de volum mare, putând fi asimilată uneia infinite, selecția poate fi tratată ca selecție repetată.

În practica statistică se utilizează și selecția nealeatoare, **selecția dirijată**. Aceasta înseamnă alegerea unităților care compun eșantionul după un anumit principiu, prestabilit, astfel încât să se includă unități reprezentative pentru toate categoriile de unități ale populației. O astfel de selecție este mai economicoasă, întrucât nu necesită un eșantion la fel de mare ca în cazul selecției aleatoare, dar riscul apariției unor erori de sistematice de reprezentativitate este mai mare în cazul selecției dirijate. Totuși este posibil ca un sondaj dirijat, bine fundamentat și organizat, să dea erori mai mici decât un sondaj aleator. În cazul eșantioanelor mari selecția aleatoare este creditată însă cu un grad mai mare de încredere.

14.2.1.2. Tipuri de sondaj

Prin aplicarea unui singur procedeu de selecție se realizează un **sondaj simplu (aleator sau dirijat)**. Necesitatea îmbinării avantajelor celor două tipuri de selecții, în sensul asigurării unui grad mare de încredere a rezultatelor cu cheltuieli cât mai mici, a condus la dezvoltarea unor **tehnici mixte de sondaj**. Un astfel de sondaj este cel **stratificat (tipic)**, realizat în condițiile în care populația este împărțită în grupe (straturi), iar selecția se face prin extragerea aleatoare a unor subeșantioane din fiecare grupă. Dacă numărul unităților extrase din grupe este proporțional cu numărul unităților cuprinse în grupele populației totale, sondajul este denumit **tipic proporțional**, iar dacă mărimea subeșantioanele de grupă este stabilită astfel încât eficiența sondajului să fie maximă (sub aspectul preciziei rezultatelor), sondajul este denumit **tipic optim**.

Un tip special de sondaj, aplicat în cazul populațiilor de mare volum, în care unitățile sunt grupate instituțional sau administrativ-teritorial, este **sondajul în faze**. Acesta se realizează prin selectarea, în prima fază, a unui număr de grupe (unități statistice complexe) din numărul total de grupe din care este compusă populația, iar în faza următoare a unui număr de unități statistice simple din numărul total de unități care compun fiecare din grupele selectate în prima fază. Dacă sondajul se realizează în mai mult de două faze, în fazele intermediare (de la a doua la penultima) sunt selectate succesiv subgrupe din grupele delectate în fazele precedente.

În România, sondajul în două faze este aplicat în anchetele în gospodării. În prima fază a fost construit, pe baza rezultatelor recensământului populației și locuințelor, un eșantion de zone teritoriale (denumit Eșantion Multifuncțional de Zone Teritoriale – EMZOT), selectate aleator din sectoarele de recensământ. Eșantionul este compus din 780 centre de cercetare (427 centre în mediul urban și 353 centre în mediul rural), repartizate în toate județele țării și în sectoarele din

Municipiul București. Locuințele aflate în aceste centre de cercetare constituie baza de eșantionare pentru toate anchetele în gospodării.

În a doua fază, din totalul locuințelor aflate în centrele de cercetare se extrage aleator un număr de locuințe stabilit în mod diferențiat de la o anchetă la alta (18036 locuințe, câte 36 în fiecare centru, în cazul Anchetei asupra forței de muncă în gospodării; 36072 locuințe, câte 72 în fiecare centru, în cazul Anchetei bugetelor de familie).

Din punctul de vedere al stabilității în timp, eșantioanele pot fi de tip **transversal** sau **longitudinal (panel)**. Primele sunt eșantioane selectate de fiecare dată când o anchetă este repetată, ultimele sunt selectate la început și sunt menținute în cercetare la toate repetările unei anchete. Există de asemenea anchete în care o parte a eșantionului este de tip panel, iar cealaltă se reînnoiește la fiecare repetare a anchetei, precum și anchete cu eșantioane de tip rotațional, în care o parte a unităților sunt menținute în cercetare la prima repetare a anchetei, după care sunt scoase din eșantion pentru următoarele repetări și reintroduse ulterior pentru una sau două repetări (Ancheta trimestrială asupra forței de muncă în gospodării).

Algoritmul de calcul al indicatorilor eșantionului și cel de estimare a parametrilor populației prezintă unele diferențe de la un tip de sondaj la altul. În acest capitol sunt prezentate numai cele care se aplică în cazul sondajului aleator simplu.

14.2.2. Erorile cercetării selective

Eroarea din cercetarea selectivă este diferența dintre indicatorii calculați pe baza datelor din eșantion și parametrii corespunzători al populației totale (ale căror valori ar fi calculate dacă ar exista date referitoare la toate unitățile populației):

➡ abatere absolută a indicatorului de selecție față de parametrul populației,

d = valoarea indicatorului de sondaj – parametrul populației,

➡ și abatere relativă

$d\% = (\text{indicatorul de selecție} - \text{parametru}) / \text{parametru} \times 100.$

În cazul mediei, de exemplu, eroarea absolută de selecție este

$$d_{\bar{x}} = \bar{x} - m$$

iar cea relativă

$$d_{\bar{x}}\% = \frac{\bar{x} - m}{m} 100$$

Unele erori se pot produce în cursul observării și prelucrării datelor. O categorie specială de erori, proprii cercetării selective, este formată din erorile de reprezentativitate. Acestea sunt sistematice și întâmplătoare.

Erorile sistematice sunt determinate de încălcarea principiului fundamental al selecției aleatoare, cerinței ca selecția să se facă în așa fel încât toate unitățile colectivității totale să aibă șanse egale de includere în eșantion. Dacă selecția este afectată de erori sistematice, extinderea rezultatelor cercetării la întreaga populație poate să conducă la rezultate eronate.

Erorile întâmplătoare țin de specificul sondajului și pot să apară chiar dacă se respectă regulile selecției aleatoare, întrucât o colectivitate cu un volum considerabil mai mic nu poate reproduce fidel structura populației. Probabilitatea apariției erorilor întâmplătoare se are în vedere atunci când se face extinderea rezultatelor asupra colectivității generale, prin determinarea intervalului de încredere.

14.2.3. Mărimea eșantionului

Din formula de calcul a erorii limită se poate remarca faptul că precizia estimării depinde de siguranța acesteia, de dispersia variabilei și de mărimea eșantionului. Precizia estimării este mai mică (eroarea este mai mare), dacă siguranța (nivelul de încredere) și dispersia sunt mari, crescând odată cu mărimea eșantionului. Cu alte cuvinte, dispersia fiind dată, pentru a obține estimatii cu un grad mai mare de precizie, cu o anumită siguranță, este necesar un volum mare al eșantionului. Pe de altă parte, cercetările pe eșantioane mari se realizează cu cheltuieli mari, iar timpul necesar pregătirii anchetei, observării și prelucrării datelor este mai mare. În consecință, se impune utilizarea unui eșantion de volum optim, care să permită obținerea preciziei și siguranței necesare, cu cheltuieli minime.

În cazul sondajului aleator simplu, mărimea eșantionului minim necesar se determină pe baza unor relații derivate din formulele de calcul ale erorii limită. Astfel, în ipoteza că abaterea standard a variabilei observate (sau a principalelor variabile, în cazul unei cercetări complexe) este cunoscută, dacă populația este infinită, volumul eșantionului se poate stabili pe baza relației:

$$n = \frac{z_{\alpha/2}^2 \sigma^2}{\Delta_{\bar{x}}^2}$$

în care $z_{\alpha/2}$ se determină din tabelul funcției Gauss-Laplace, pentru nivelul de încredere dorit, iar este $\Delta_{\bar{x}}$ eroarea maximă admisă.

Dacă populația este finită, de volum N , relația de calcul a volumului eșantionului este:

$$n = \frac{Nz_{\alpha/2}^2 \sigma^2}{(N-1)\Delta_{\bar{x}}^2 + z_{\alpha/2}^2 \sigma^2} \quad \text{sau}$$

$$n = \frac{z_{\alpha/2}^2 \sigma^2}{\Delta_{\bar{x}}^2 + \frac{z_{\alpha/2}^2 \sigma^2}{N}}$$

dacă volumul populației este mare.

Să ne reamintim

Atunci când pentru analiza unei caracteristici de interes utilizăm un eșantion, și nu întreaga populație statistică, absența informației complete face ca rezultatul la care ajungem să fie supus unui grad de incertitudine. Inferența statistică este procesul prin care extrapolăm concluziile pe care le-am obținut - referitor la o anumită caracteristică - pornind de la un eșantion, pentru a ne informa cu privire la caracteristica similară a întregii populații statistice.



14.3. Rezumat

Pentru caracterizarea statistică a fenomenelor și proceselor social-economice de masă cu scopul evidențierii legităților specifice acestora, se folosesc date individuale obținute prin metode de înregistrare totală sau parțială. În practică, din diferite motive pentru anumite colectivități nu se poate face înregistrarea totală a unităților ce le compun, fie că ar fi necesitat cheltuieli foarte mari, fie că operativitatea obținerii rezultatelor cercetării nu este asigurată. Toate aceste considerente și altele nespicate, necesită tot mai frecvent organizarea unor înregistrări parțiale de tipul sondajelor statistice.

În cercetarea statistică selectivă se operează cu categorii distincte referitoare la colectivitatea totală și cea de selecție. Colectivitatea totală poartă și denumirea de colectivitate generală, colectivitate generală sau populație, iar colectivitatea de selecție este denumită și eșantion, mostră sau probă.



14.4. Test de autoevaluare a cunoștințelor

1. Ce înțelegeți prin cercetare statistic selectivă?
2. Cum se stabilește mărimea eșantionului?
3. Care sunt cei mai utilizați indicatori estimați în cadrul cercetării selective?
4. Ce reprezintă selecția?
5. Ce metode de selecție cunoașteți?



14.5. Bibliografie recomandată

- a. Caragea Nicoleta, *Statistica- Concepte si metode de analiza a datelor*, Editura MUSTANG, București, 2015
- b. Voineagu, V. , *Statistică –baze teoretice și aplicații*, Editura Economică, București, 2007
- c. Georgescu-Roegen N., *Metoda statistică*, Editura Expert, București, 1998.
- d. Andrei T., Stancu S., Pele D.T., *Statistică. Teorie și aplicații* (ediția a II-a), Editura Economică, București, 2003.
- e. Biji M., Biji E.M., Lilea E., Anghelache C., *Tratat de statistică*, Editura Economică, București, 2002.

Răspunsuri la testele de autoevaluare a cunoștințelor

Capitolul 1. Domeniul și metoda statisticii

- 1) e
- 2) b
- 3) b

Capitolul 2. Sursele de date statistice

- 1) a
- 2) a
- 3) e

Capitolul 3. Sistematizarea informației statistice

- 1) e
- 2) c
- 3) f

Capitolul 4. Serii de distribuție

- 1) d
- 2) d
- 3) a
- 4) b
- 5) d
- 6) a

Capitolul 5. Indicatorii frecvenței

- 1) b
- 2) d
- 3) e
- 4) a
- 5) e
- 6) a
- 7) b

Capitolul 6. Indicatorii tendinței centrale. Media

- 1) d
- 2) c
- 3) b
- 4) b
- 5) a
- 6) a

Capitolul 7. Alte tipuri de medie. Proprietățile mediei

- 1) e
- 2) e
- 3) e
- 4) b
- 5) b

- 6) e
- 7) a
- 8) b

Capitolul 8. Mediana

- 1) e
- 2) c
- 3) a
- 4) c
- 5) a
- 6) a
- 7) b
- 8) c
- 9) a
- 10) b

Capitolul 9. Modul

- 1) c
- 2) c
- 3) d
- 4) a
- 5) b

Capitolul 10. Indicatorii de nivel. Quartilele

- 1) e
- 2) e
- 3) e

Capitolul 11. Indicatorii simpli ai variației

- 1) b
- 2) b
- 3) e
- 4) a
- 5) e
- 6) b
- 7) b
- 8) a
- 9) a
- 10) d

Capitolul 12. Indicatorii sintetici ai variației

- 1) b
- 2) a
- 3) d
- 4) e
- 5) e
- 6) e
- 7) e
- 8) b
- 9) e
- 10) c