CURSUL IV

MĂSURAREA VARIAŢIEI ÎN SERII UNIDIMENSIONALE ŞI BIDIMENSIONALE

4.0 Introducere
4.1 Indicatorii variaţiei
4.2 Proprietăţile dispersiei
4.3 Dispersia în analiza distribuţiilor bidimensionale
4.4 Forme de repartizare a frecvenţelor
4.4.1 Momentele
4.4.2 Asimetria
4.4.3 Curtozisul

4.0 Introducere

În cadrul analizei statistice, studiul variaţiei fenomenelor social-economice ocupă un loc foarte important; este absolut necesară cunoaşterea domeniului de variaţie a unei caracteristici oarecare pentru **a** aprecia şi estima corect principalele tendinţe ale acesteia, raporturile dintre variabilitatea ei şi cea a altor caracteristici, etc.

Este cunoscut faptul că, gradul de semnificaţie a indicatorilor medii este strâns legat de câmpul de variaţie a valorilor unui fenomen, de modul de dispersare a acestor valori în interiorul câmpului de variaţie, de gradul de aglomerare a cazurilor în jurul unor valori, etc. Aspectele respective sunt cuantificate prin intermediul unor indicatori specifici a căror metodologie de calcul, semnificaţie şi utilitate vor fi prezentate în acest capitol.

4.1 Indicatorii variaţiei

Se deosebesc două grupe:
a)Indicatori simpli
b)Indicatori sintetici

Prima grupă cuprinde:

a1) Amplitudinea variaţiei. Oferă posibilitatea delimitării câmpului de variaţie a unui fenomen. Se întâlnesc două variante:

a$_{1.1}$) Amplitudinea absolută. Se calculează ca diferenţă între varianta maximă şi varianta minimă ale caracteristicii, folosind relaţia:

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

A$_{1.2}$) **Amplitudinea relativă.** Se calculează ca raport între amplitudinea absolută şi media caracteristicii respective, exprimându-se în procente. Deci:

$$A_r = \frac{A_a}{\bar{x}} \cdot 100$$

a2) Abaterea variantelor caracteristicii de la media lor. Oferă posibilitatea cunoaşterii structurii variaţiei la nivelul fiecărei unităţi statistice. Se întâlnesc două variante:

a$_{2.1}$) Abaterea absolută. Se calculează ca diferenţă între variantele caracteristicii şi media lor, folosind relaţia :

$$d_a = x_i - \bar{x} \quad \Rightarrow \{x_1 - \bar{x}\,;\, x_2 - \bar{x}\,;\, x_3 - \bar{x}\,;\, \ldots\ldots$$

a$_{2.2}$) Abaterea relativă. Se calculează ca raport între abaterea absolută şi medie, exprimându-se în procente. Relaţia:

$$d_r = \frac{d_a}{\bar{x}} \cdot 100 = \frac{x_i - \bar{x}}{\bar{x}} \cdot 100$$

Cea de a doua grupă cuprinde:

b1) Abaterea medie lineară. Se calculează ca o medie **aritmetică** – simplă sau ponderată - **a** abaterilor absolute ale variantelor caracteristicii de la media lor, luate sub formă de modul, folosind una dintre relaţiile:

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

$$\bar{d} = \frac{\sum_{j=1}^{m} |x_j - \bar{x}| f_j}{\sum_{j=1}^{m} f_j} \Rightarrow \text{ponderată}$$

în care:

i = numărul nivelurilor caracteristicii X, nerepetabile în cadrul şirului de valori în care se regăsesc ;

j = numărul variantelor caracteristicii X, adică al valorilor care se repetă la mai multe unităţi statistice din colectivitatea studiată.

Semnificaţie: Sintetizează nivelul mediu al abaterilor absolute ale variantelor caracteristicii de la media lor. Datorită faptului că abaterile absolute se iau sub formă de modul, se acordă aceeaşi importanţă atât abaterilor **pozitive** cât şi **abaterilor negative.** Acest lucru se bazează pe proprietatea mediei aritmetice potrivit căreia suma abaterilor variantelor de la media lor este egală cu zero.

Utilitate: Se foloseşte pentru determinarea intervalului mediu de variaţie. Astfel:

$$\bar{x} \pm \bar{d} \Rightarrow \begin{cases} \bar{x} + \bar{d} \\ \Updownarrow \\ \bar{x} - \bar{d} \end{cases}$$

În practică acest indicator este folosit mai rar.

Având în vedere faptul că, în foarte multe cazuri, în analiza economică se determină şi se urmăresc diferenţele dintre nivelurile efective ale diverşilor indicatori, realizate în perioade diferite, sau diferenţele dintre aceste niveluri şi cele standardizate sau planificate, pentru analiza abaterilor respective se poate folosi, cu rezultate foarte bune, un gen de abatere medie lineară, în calculul căreia se ţine seama de semnul " + sau -" al acestor diferenţe.

Sa presupunem că analizăm dinamica producţiei lunare, realizată de o întreprindere oarecare, folosindu-se producţia realizată în ultima lună a anului precedent.

Notăm cu:

n_1= numărul de luni în care s-au realizat depăşiri, abateri pozitive;

n_2= numărul de luni ăn care s-au realizat abateri negative;

n_0=n_1+n_2= numărul total de luni din cadrul intervalului supus analizei.

Abaterea medie lineară pozitivă se determină cu relaţia:

$$(+)\bar{d}_1 = \frac{\sum_{i=1}^{n_1} (x_i - x_r)}{n_1}$$

în care:

i = 1,2, 3,..............,n_1;

x_i - producţiile lunare mai mari decât nivelul referenţial;

x_r = nivel de referinţă (baza de comparaţie).

Abaterea medie lineară **negativă** se determină cu ajutorul relaţiei:

$$(-)\overline{d}_2 = \frac{\sum_{j=1}^{n_2}(x_j - x_r)}{n_2}$$

în care:

i = l,2, 3,.............,n2;

x_j - producţiile lunare mai mici decât nivelul referenţial;

Abaterea medie lineară **generală**, aferentă intervalului n_0, este dată de relaţia:

$$(\pm)\overline{d}_0 = \frac{n_1\overline{d}_1}{n_0} + \left(-\frac{n_2\overline{d}_2}{n_0}\right)$$

ceea ce este echivalent cu :

$$(\pm)\overline{d}_0 = \frac{\sum_{i=1}^{n_1}(x_i - x_r) + \left[-\sum_{j=1}^{n_2}(x_j - x_r)\right]}{n_0}$$

Semnificaţie: Abaterea medie lineară generală, sintetizează nivelul mediu lunar al **soldului** rezultat din suma algebrică a celor două tipuri de abateri - pozitivă şi negativă. Din relaţia sa, se desprinde faptul că, pentru oricare n_0 poate fi determinat nivelul global al abaterilor, fie pozitive, fie negative, prin relaţia:

$$(\pm)\Delta = n_0\overline{d}_0$$

Acest indicator are o valoare informaţională deosebită. Semnul şi nivelul său îl avertizează pe agentul economic, din vreme, asupra tendinţei evolutive a indicatorului supus analizei. El va cunoaşte cu mult timp înainte tendinţa evolutivă a fenomenului care-1 interesează, luând măsurile corespunzătoare.

Această metodă poate fi folosită pentru analiza variabilităţii următorilor indicatori: producţia fabricată, volumul vânzărilor, volumul stocurilor de producţie la producător, volumul desfacerilor de mărfuri ale unei unităţi comerciale, evoluţia stocurilor la unităţile comerciale, etc.

b2) Dispersia. Se calculează ca o medie **aritmetică** - simplă sau ponderată - a **pătratului** abaterilor absolute ale variantelor caracteristicii de la media lor, folosind una dintre relaţiile:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n} \Rightarrow \text{simplă};$$

$$\sigma^2 = \frac{\sum_{j=1}^{m}(x_j - \overline{x})^2 f_j}{\sum_{j=1}^{m} f_j} \Rightarrow \text{ponderată}.$$

unde i şi j au aceeaşi semnificaţie de la tipurile de medii (i - pentru mediile simple, j - pentru mediile ponderate).

Dispersia fiind un **moment centrat** de **ordinul 2,** nu poate fi folosită în caracterizarea variaţiei la seriile simple sau la distribuţiile unidimensionale; în cazul acestora, este folosită ca bază de calcul pentru abaterea medie pătratică. Pentru distribuţiile bidimensionale însă, dată fiind comparabilitatea indicatorilor de acelaşi ordin, dispersia este folosită în caracterizarea variaţiei uneia dintre variabile, de obicei a varibilei secundare, dependente. în cadrul acestei utilizări, dispersia generală se descompune în dispersii factoriale (vezi paragraful 4.3).

În literatura de specialitate, dispersia este cunoscută şi sub numele de **varianţă**.

b3) Abaterea standard sau **Abaterea medie pătratică.** Se calculează ca o medie pătratică - simplă sau ponderată - a abaterilor absolute ale variantelor caracteristicii de la media lor, folosind una dintre relaţiile:

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

$$\sigma = \sqrt{\frac{\sum_{j=1}^{m}(x_j - \bar{x})^2 f_j}{\sum_{j=1}^{m} f_j}} \Rightarrow \text{ponderată .}$$

Analizând cele două relaţii, rezultă că abaterea standard este rădăcina pătrată din dispersie.

Fiind un indicator de gradul l, ca şi media, poate fi folosit în caracterizarea variaţiei, determinându-se cu ajutorul său intervalul mediu de variaţie, după metodologia prezentată la abaterea medie lineară. Deci:

$$\bar{x} \pm \sigma \Rightarrow \begin{cases} \bar{x} + \sigma \\ \Updownarrow \\ \bar{x} - \sigma \end{cases}$$

Dacă pentru aceeaşi serie, se calculează abaterea medie lineară şi abaterea standard, întotdeauna :

$$\sigma > \bar{d}$$

Rezultă că şi intervalul mediu de variaţie, determinat cu ajutorul abaterii standard, este mai larg decât cel stabilit cu abaterea medie lineară; din acest motiv, abaterea standard este preferată în analiza variaţiei fenomenelor social-economice.

b4) Coeficientul de variaţie. Se determină ca raport între abaterea standard şi nivelul mediu al unei variabile oarecare, exprimându-se, de obicei, în procente. El este de fapt expresia relativă a abaterii standard, adică ponderea acesteia în valoarea mediei.
Se calculează după următoarea relaţie:

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

Semnificaţie. Cu cât nivelul său este mai apropiat de zero, cu atât variaţia unui fenomen social-economic este mai redusă, mai slabă. In acest caz colectivitatea studiată este mai omogenă; indicatorii sintetici folosiţi în analiza diferitelor caracteristici ale acestor colectivităţi, au un grad de semnificaţie din ce în ce mai ridicat. Cu cât nivelul coeficientului de variaţie este mai îndepărtat de zero, cu atât variaţia fenomenului studiat este mai pronunţată; eterogenitatea colectivităţii statistice este din ce în ce mai accentuată; indicatorii sintetici au în acest caz un grad de semnificaţie din ce în ce mai redus. Spre exemplu, dacă v = 0, atunci valorile x_i sunt egale; între unităţile statistice, privite din acest punct de vedere, nu există nici-o deosebire; media are gradul de semnificaţie (de reprezentaţivitate) de 100 %.

În practică se consideră semnificative valorile lui v de până la 30 %. Nivelurile mai mari, presupun faptul că indicatorii sintetici, folosiţi mai ales în calculele estimative, introduc erori din ce în ce mai mari. Pentru evitarea acestor situaţii, atunci când v > 30 %, se recomandă împărţirea colectivităţii în grupe omogene şi determinarea indicatorilor sintetici pentru fiecare grupă.

Coeficientul de variaţie este folosit pe scară largă în analiza statistică, alături de abaterea standard, apreciindu-se cu ajutorul său nu numai intensitatea variaţiei unui fenomen ci şi gradul de semnificaţie a unor valori tipice, cum este - spre exemplu - media.

4.2 Proprietăţile dispersiei

Dispersia fiind calculată ca o medie aritmetică, înseamnă că proprietăţile mediei aritmetice sunt aplicabile şi în caz dispersiei. Reţin atenţia în mod deosebit proprietăţile care facilitează simplificarea calculului dispersiei.

Prima proprietate. *Dispersia calculată din abaterile variantelor x_j de la o constantă „a", este mai mare decat dispersia reală cu pătratul diferenţei dintre medie şi constanta "a"* ceea ce simbolic se poate scrie astfel:

$$\frac{\sum_{j=1}^{m}(x_j - a)^2 f_j}{\sum_{j=1}^{m} f_j} > \sigma^2 \; cu \; (\bar{x} - a)^2$$

Pentru egalare, se alege una dintre cele două posibilităţi:

$$\sigma^2 = \frac{\sum_{j=1}^{m}(x_j \quad a)^2 f_j}{\sum_{j=1}^{m} f_j} - (\bar{x} - a)^2$$

Deci, pentru calculul dispersiei există şi această posibilitate, care este mai complicată însă decât calculul obişnuit, în cadrul căruia dispersia este determinată ca o medie aritmetică ponderă a patratului abaterilor variantelor xj de la media lor.

Dispersia calculată pe baza abaterilor variantelor x_j de la constanta „a„ este momentul obişnuit de ordinul 2.

A doua proprietate. *Dispersia calculată din abaterile variantelor x_j de la media lor, micşorate în prealabil prin împărţire la o constantă „k" este mai* ***mica*** *decât dispersia reală de "k^2 "ori* , ceea ce simbolic se poate scrie astfel:

$$\frac{\sum_{j=1}^{m}\left(\frac{x_j - \bar{x}}{k}\right)^2 f_j}{\sum_{j=1}^{m} f_j} < \sigma^2 \; de \; k^2 \; ori$$

Pentru egalare, se alege una dintre cele două posibilităţi:

$$\sigma^2 = \frac{\sum_{j=1}^{m}\left(\frac{x_j - \bar{x}}{k}\right)^2 f_j}{\sum_{j=1}^{m} f_j} \cdot k^2$$

Iată încă o posibilitate de calcul pentru dispersie, care este însă mai complicată decât cea obişnuită.

Aplicând concomitent cele două proprietăţi, se ajunge la relaţia de calcul simplificat al dispersiei:

$$\sigma^2 = \frac{\sum_{j=1}^{m}\left(\frac{x_j - a}{k}\right)^2 f_j}{\sum_{j=1}^{m} f_j} \cdot k^2 - (\bar{x} - a)^2$$

La prima vedere, această relaţie pare mult mai complicată decat relaţia obişnuită; practic însă, calculul se simplifică într-o măsură considerabilă datorită faptului că: după stabilirea poziţiei lui "a" în cadrul seriei, pentru rapoartele (xj - a) / k se vor trece automat valorile: **zero** - în dreptul poziţiei lui a ; -1,-2,-3 etc. deasupra şi 1,2,3 etc**.** sub zero. Este acelaşi avantaj, pe care l-am evidenţiat şi la calculul simplificat al mediei aritmetice ponderate.

În paragraful 4.3 este prezentat un exemplu în cadrul căruia se foleseşte procedeul de calcul simplificat al dispersiei.

4.3 Dispersia în analiza distribuţiilor bidimensionale

După cum am arătat în & 4.1, pentru analiza variaţiei caracteristicii secundare dintr-o distribuţie bidimensională, se poate folosi dispersia, fără a mai fi nevoie de calculul altor indicatori sintetici ai variaţiei.

Dispersia oferă posibilitatea separării şi determinării influenţei factorilor întâmplători şi esenţiali, prin descompunerea dispersiei generale în dispersii factoriale.

Este cunoscut faptul că, în general, distribuţiile bidimensionale cuprind caracteristici de grupare între care există raporturi de cauzalitate. În aceste cazuri, analizând modul de dispunere a frecvenţelor f_{xy} în interiorul tabelului, se va observa faptul că, pe măsură ce valorile caracteristicii factoriale(principale) înregistrează o tendinţă de creştere, asistăm la o deplasare a frecvenţelor respective de-a lungul uneia dintre diagonalele tabelului. Este semnalată astfel, existenţa legăturii cauzale dintre cele două variabile. Influenţa variabilităţii caracteristicii principale X asupra variaţiei caracteristicii secundare Y, poate fi separată şi cuantificată cu ajutorul dispersiei. Variabila principală este considerată factor esenţial.

Dar valorile caracteristicii secundare sunt şi rezultatul acţiunii altor factori care - în raport cu factorul esenţial - sunt consideraţi factori întâmplători; şi influenţa acestor factori poate fi cuantificată cu ajutorul dispersiei în acest scop se folosesc următoarele tipuri de dispersie:

1.Dispersia de grupă. Se calculează ca o medie aritmetică ponderată a pătratului abaterilor variantelor y_j de la media de grupă, folosind relaţia:

$$\sigma_i^2 = \frac{\sum (y_{ij} - \bar{y}_i)^2 f_{ij}}{f_i} = \frac{\sum (y_{ij} - \bar{y}_i)^2 f_{ij}}{\sum f_{ij}}$$

Mediile de grupă sunt medii aritmetice ponderate, cu următoarea relaţie de bază:

$$\bar{y} = \frac{\sum y_{ij} f_{ij}}{f_i} = \frac{\sum y_{ij} f_{ij}}{\sum f_{ij}}$$

în care:

i = 1, 2, 3 ..... n; reprezintă numărul de grupe constituite după variabila principală X;

j = 1, 2, 3, ... m ; reprezintă numărul de grupe constituite după variabila secundară Y;

y_{ij} = variantele caracteristicii Y pentru care există cazuri în cadrul grupelor „i„ adică pentru situaţiile în care $f_{ij} > 0$;

$\bar{y}_i$ = media valorilor yi j, corespunzătoare fiecărei grupe " i";

f_{ij} = frecvenţele corespunzătoare valorilor yj j;

f_i = frecvenţele fiecărei grupe „i" (sunt frecvenţele f_x din tabelul bidimensional).

Cunoscând semnificaţia simbolurilor din relaţia de calcul al dispersiei de grupă prezentarea ei sub formă desfăşurată facilitează înţelegerea conţinutului acestui indicator.

Astfel, pentru i = 1 , adică pentru prima grupă după X:

$$\sigma_1^2 = \frac{\sum (y_{1j} - \bar{y}_1)^2 f_{1j}}{f_1} =$$

$$= \frac{(y_{11} - \bar{y}_1)^2 f_{11} + (y_{12} - \bar{y}_1)^2 f_{12} + \ldots\ldots + (y_{1m} - \bar{y}_1)^2 f_{1m}}{f_1}$$

$$unde\ f_1 = f_{11} + f_{12} + \cdots + f_{1m}$$

pentru i=2:

$$\sigma_2^2 = \frac{\sum (y_{2j} - \bar{y}_2)^2 f_{2j}}{f_2} =$$

$$= \frac{(y_{21} - \bar{y}_2)^2 f_{21} + (y_{22} - \bar{y}_2)^2 f_{22} + \cdots + (y_{2m} - \bar{y}_2)^2 f_{2m}}{f_2}$$

$$unde\ f_2 = f_{21} + f_{22} + \cdots + f_{2m}$$

Pentru i = 3 ;4 ; 5 ; etc calculul dispersiilor de grupă este similar.

Semnificaţie. Dacă admitem faptul că grupele "i", constituite după variabila principală, sunt omogene, ar trebui ca unităţile statistice din fiecare grupă să se concentreze într-o singură grupă (clasă) a valorilor caracteristicii secundare Y. În acest caz, dispersiile de grupă ar fi egale cu zero, ceea ce presupune faptul că variabila Y depinde în exclusivitate de variabila X.

Ilustrăm această situaţie printr-un caz concret: presupunem că analizăm variaţia salariilor pentru muncitorii unei formaţii de lucru, grupaţi bidimensional după nivelurile productivităţii muncii şi salariului efectiv, realizate într-o lună oarecare. Evident, muncitorii care au realizat acelaşi nivel al productivităţii, ar trebui să primească acelaşi salariu, în condiţiile plăţii cu tariful unic pe unitatea de produs. Acest lucru presupune faptul că, variaţia salariului este determinată, în exclusivitate, de variaţia productivităţii muncii.

În practică însă, nivelurile şi dinamica salariului sunt influenţate şi de alţi factori: categoria de calificare cu salariul tarifar de încadrare corespunzător, timpul lucrat de fiecare muncitor, gradul de dotare tehnică, etc. Toţi aceşti factori sunt consideraţi, în raport cu productivitatea muncii, factori **întâmplători.** Influenţa lor se concretizează în gradul de dispersare, de împrăştiere a numărului de muncitori din fiecare grupă " i ", într-un interval mai mic sau mai mare de variaţie a salariilor.

În cazul în care aceşti factori nu ar fi existat sau nu ar fi exercitat nici-o influenţă asupra salariului, dispersiile de grupă ar fi fost egale cu zero; existenţa acţiunii acestor factori este semnalată, în primul rând, de prezenţa frecvenţelor f_{ij} la mai multe valori y_j în cadrul fiecărei grupe "i". Intensitatea acţiunii acestor factori este relevată de faptul că : cu cât nivelurile dispersiilor de grupă sunt mai îndepărtate de zero, cu atât influenţa factorilor respectivi este mai pronunţată.

Deci, generalizând şi sintetizând semnificaţia dispersiei de grupă, se poate spune că: dispersia de grupă sintetizează intensitatea influenţei factorilor **întâmplători,** care au determinat variaţia caracteristicii secundare Y în cadrul fiecărei grupe constituite după caracteristica principală X.

Prin compararea nivelurilor acestor dispersii, se poate face o ierarhizare a intensităţii factorilor întâmplători pentru fiecare grupă " i "; dispersiile cu nivelurile cele mai mari, evidenţiază faptul că factorii respectivi au avut o influenţă mai pronunţată în raport cu grupele la care dispersiile au valori mai mici.

2. Media dispersiilor de grupă. Se calculează ca o medie aritmetică ponderată a dispersiilor de grupă, folosind relaţia:

$$\overline{\sigma}^2 = \frac{\sum \sigma_i^2 f_i}{\sum f_i}$$

Toate elementele din cadrul relaţiei sunt deja cunoscute; precizăm doar faptul că $\sum f_i$ reprezintă numărul total de unităţi statistice din colectivitatea generală. Făcând legătura cu notaţiile prezentate în capitolul 1, paragraful referitor la construcţia seriilor de distribuţie, rezultă că:

$$\sum_i f_i = \sum_x f_x = \sum_j f_j = \sum_y f_y = \sum_i \sum_j f_{ij} = \sum_x \sum_y f_{xy} = N_0$$

Media dispersiilor de grupă este un indicator sintetic determinat la nivelul întregii colectivităţi; relaţia de calcul desfăşurată, are următoarea formă de prezentare:

$$\overline{\sigma}^2 = \frac{\sigma_1^2 f_1 + \sigma_2^2 f_2 + \cdots + \sigma_n^2 f_n}{f_1 + f_2 + \cdots + f_n}$$

Din formula desfăşurată rezultă faptul că, în locul frecvenţelor absolute $f_1, f_2, f_3, \ldots, f_n$, se pot folosi frecvenţele relative fr_1 , fr_2 , fr_3 ,...., fr_n , ceea ce permite cunoaşterea contribuţiei fiecărei dispersii de grupă la variaţia caracteristicii secundare determinată la nivelul colectivităţii generale de acţiunea factorilor **întâmplători.** Deci:

$$\overline{\sigma}^2 = \sigma_1^2 fr_1 + \sigma_2^2 fr_2 + \cdots + \sigma_n^2 fr_n$$

Contribuţia fiecărei dispersii, exprimată în procente:

$$C_1 = \frac{\sigma_1^2 fr_1}{\overline{\sigma}^2} \cdot 100\,;\; C_2 = \frac{\sigma_2^2 fr_2}{\overline{\sigma}^2} \cdot 100\,; \cdots$$

După cum se va observa din analiza lucrărilor practice şi programelor aplicative, folosirea frecvenţelor relative în locul frecvenţelor absolute prezintă o serie de avantaje, motiv pentru care aceste frecvenţe sunt preferate.

Semnificaţie. Dacă dispersiile de grupă evidenţiază variaţia caracteristicii secundare Y determinată de influenţa factorilor întâmplători **la nivelul fiecărei grupe,** atunci media acestor dispersii sintetizează **tot influenţa factorilor întâmplători,** însă la nivelul colectivităţii generale.

3. Dispersia mediilor de grupă. Se calculează ca o medie aritmetică ponderată a pătratului abaterilor mediilor de grupă de la media generală, folosind relaţia:

$$\delta^2 = \frac{\sum (\overline{y}_i - \overline{y}_0)^2 f_j}{\sum f_j}$$

Şi în această relaţie toate elementele sunt cunoscute; media generală este - ca şi mediile de grupă - tot o medie aritmetică ponderată, adică:

$$\bar{y}_0 = \frac{\sum y_j f_j}{\sum f_j}$$

în care, frecvenţele fj, nu sunt altele decât frecvenţele f_y corespunzătoare valorilor yj;

$\sum f_j$ reprezintă volumul colectivităţii generale .

Semnificaţie. Ştiind că media sintetizează ceea ce este esenţial în variabilitatea unui fenomen, rezultă că dispersia mediilor de grupă, comparând ceea ce este esenţial la nivelul grupelor - prin mediile de grupă - cu ceea ce este esenţial la nivelul colectivităţii generale - prin media generală , sintetizează variaţia caracteristicii secundare Y, determinată de acţiunea factorilor **esenţiali,** la nivelul întregii colectivităţi.

Dacă ne referim din nou la distribuţia muncitorilor după productivitatea muncii şi după salariul realizat şi presupunând că mediile de grupă sunt egale, ar înseamnă **că** variaţia productivităţii muncii nu ar fi avut nici-o influenţă asupra salariului; în acest caz media generală ar fi fost egală cu mediile de grupă şi deci, dispersia mediilor de grupă ar fi fost egală cu zero. Dar productivitatea muncii este însă unul dintre factorii determinanţi ai salariului, întrucât o parte însemnată a acestuia este strâns legată de nivelul productivităţii muncii prin sistemul de salarizare; această legătură este sintetizată de dispersia mediilor de grupă, al cărei nivel este cu atât mai mare cu cât influenţa productivităţii asupra variaţiei salariului este mai consistentă.

Generalizând aceste aprecieri, se poate spune că, valorile dispersiei mediilor de grupa apropiate de zero, evidenţiază o influenţă slabă a factorilor esenţiali, iar valorile din ce în ce mai mari - o infuenţă a cărei intensitate este din ce în ce mai mare.

Din cele trei tipuri de dispersie prezentate, se reţine faptul că, **media dispersiilor de grupă şi dispersia mediilor de grupă** (2 şi 3), pot fi comparate între ele, întrucât se calculează la nivelul întregii colectivităţi. Prin comparaţie se poate observa care dintre grupele de factori - întâmplători sau esenţiali - au influenţat într-o măsură mai mare variaţia caracteristicii secundare Y.

Evident, o atenţie deosebită trebuie acordată influenţei factorilor **întâmplători,** pentru a cunoaşte cauzele care au avut ca efect dispersarea - mai mult sau mai puţin pronunţată - a unităţilor statistice din cadrul grupelor ; pot fi depistate astfel cauzele subiective şi obiective care au provocat o deplasare a frecvenţelor f_{ij} din cadrul fiecărei grupe "**i**" spre zona valorilor y_{ij} mai mici decât media grupei; pot fi cunoscute şi condiţiile care au favorizat deplasarea acestor frecvenţe spre zona valorilor y_{ij} mai mari decât media grupei.

4. Dispersia generală. Se calculează ca o medie aritmetică ponderată a pătratului abaterilor variantelor y_j de la media generală, folosind relaţia:

$$\sigma_0^2 = \frac{\sum (y_j - \bar{y}_0)^2 f_j}{\sum f_j}$$

Semnificaţie. Dispersia generală sintetizează variaţia caracteristicii secundare Y determinată de acţiunea concomitentă **atât** a factorilor **întâmplători,** cât şi a factorilor **esenţiali,** la nivelul colectivităţii generale. Cu alte cuvinte,dispersia generală evidenţiază variaţia globală, integrală a caracteristicii secundare.

Ţinând seama de faptul că dispersiile de la punctele 2 şi 3 evidenţiază separat influenţele celor două grupe de factori, rezultă că :

$$\sigma_0^2 = \bar{\sigma}^2 + \delta^2$$

de unde şi posibilitatea determinării ponderii variaţiilor aferente influenţelor respective în variaţia totală. Astfel:

$$P_1 = \frac{\overline{\sigma}^2}{\sigma_0^2} \cdot 100 \; ; \; P_2 = \frac{\delta^2}{\sigma_0^2} \cdot 100 \Rightarrow P_1 + P_2 = 100$$

Această metodologie este folosită şi în studiul corelaţiei dintre fenomene şi mai ales în calculul raportului de corelaţie, după cum vom vedea în capitolele urmatoare.

Pentru a uşura înţelegerea modului de folosire a dispersiei în analiza variaţiei, vom aborda elementele de ordin practic folosind distribuţia bidimensională prezentată în tabelul 4.1.

Tab. 4.1. Distribuţia muncitorilor dintr-o formaţie de lucru după productivitatea muncii (X) şi salariul efectiv (Y) – Octombrie

Y mii lei / X buc	30-50 $y_1 = 40$	50-70 $y_2 = 60$	70-90 $y_3 = 80$	90-110 y4=100	110- 130 y5 =120	130-150 y6 = 140	Total f_i
16-25	5	3	2	-	-	-	10
26-35	2	10	4	2	-	-	18
36-45	-	4	10	4	8	4	40
46-55	-	-	2	5	4	3	14
56-65	-	-	-	-	4	-	4
Total f_j	7	17	18	21	16	7	86

La determinarea mediilor de grupă, mediei generale, dispersiilor de grupă şi dispersiei generale, se va folosi procedeul de calcul simplificat; datele necesare acestui mod de calcul sunt prezentate în tabelul 4.2.

Mediile de grupă sunt calculate cu ajutorul relaţiei:

$$\overline{y}_i = \frac{\sum \left(\frac{y_{ij} - a}{k} \right) f_{ij}}{f_i} \cdot k + a$$

Media generală este calculată cu relaţia:

$$\overline{y}_0 = \frac{\sum \left(\frac{y_j - a}{k} \right) f_j}{\sum f_j} \cdot k + a$$

Dispersiile de grupă s-au calculat cu ajutorul relaţiei:

$$\sigma_i^2 = \frac{\sum \left(\frac{y_{ij} - a}{k} \right)^2 f_{ij}}{f_i} \cdot k^2 - \left(\overline{y}_i - a \right)^2$$

Dispersia generală a fost calculată cu relaţia:

$$\sigma_0^2 = \frac{\sum \left(\frac{y_j - a}{k} \right)^2 f_j}{\sum f_j} \cdot k^2 - \left(\overline{y}_0 - a \right)^2$$

Despre modul în care au fost stabilite datele din tabelul 4.2, sunt necesare câteva explicaţii.

Pentru prima grupă (i = 1), nivelul constantei "**a**" este centrul intervalului y cu frecvenţa 5 - care este cea mai mare din grupa respectivă; deci **a** = 40. Constanta " k " este mărimea intervalului de variaţie ales pentru variabila Y; deci k = 20 - menţinându-se constant pentru toate grupele .

Valorile corespunzătoare rapoartelor (y_j - a) / k sunt determinate astfel (redăm metodologia de bază):

$$a = 40; \quad (y_1 - a)/k = (40 - 40)/20 = 0;$$
$$y_1 = 40; \quad (y_2 - a)/k = (60 - 40)/20 = 1;$$
$$y_2 = 60; \quad (y_3 - a)/k = (80 - 40)/20 = 2.$$
$$y_3 = 80;$$

Acest calcul este o confirmare a principiului potrivit căruia, după stabilirea poziţiei lui " **a** " în cadrul distribuţiei, pentru rapoartele (y-a)/k se trec **automat** valorile: zero - în dreptul poziţiei lui " **a** ", ... (vezi paragraful 4.2). Vom urmări aplicarea acestui principiu, de pildă, la grupa a treia (i = 3): observăm că **a** = y_4 = 100**,** valoare corespunzătoare frecvenţei maxime (14); în cadrul minidistribuţiei grupei respective, această valoare ocupă locul 3; deci, rapoartele (y_j - a) / k vor avea următoarele rezultate: **zero** - în dreptul poziţiei lui " **a** ", -l, -2 - deasupra şi 1, 2 sub zero.

Calculul mediilor de grupă şi mediei generale.

$$\bar{y}_1 = \frac{7}{10}\cdot 20 + 40 = 54\,;\ \bar{y}_2 = \frac{6}{18}\cdot 20 + 60 = 66{,}66\,;\ \bar{y}_3 = 99\,;$$
$$\bar{y}_4 = 111{,}43\,;\ \bar{y}_5 = 120\,;\ \bar{y}_0 = 90$$

1. **Dispersiile de grupă.**

$$\sigma_1^2 = \frac{11}{10}\cdot 20^2 - (54-40)^2 = 244\,;\ \sigma_2^2 = \frac{14}{18}\cdot 20^2 - (66{,}66-60)^2 = 267\,;$$
$$\sigma_3^2 = 499\,;\ \sigma_4^2 = 384\,;\ \sigma_5^2 = 0$$

Interpretare : în grupa a treia factorii întâmplători au avut influenţa cu intensitatea cea mai mare, întrucât dispersia grupei respective are nivelul cel mai ridicat (499); urmează grupele 4, 2 si 1. In grupa a cincea, influenţa factorilor întâmplători a fost nulă, toţi muncitorii s-au plasat în aceeaşi grupă de salarii. In funcţie de aceste rezultate, se poate trece la analiza cauzelor care le-au determinat.

2. **Media dispersiilor de grupă.**

$$\bar{\sigma}^2 = \frac{244\times 10 + 267\times 18 + \cdots + 0\times 4}{86} = 379$$

3. **Dispersia mediilor de grupă.**

$$\delta^2 = \frac{(54-90)^2\times 10 + (66{,}66-90)^2\times 18 + \cdots}{86} = 419$$

Tab. 4.2. Datele necesare calculului simplificat al mediilor şi dispersiilor folosite în analiza distribuţiei din Tab. 4.1.

Grupa după variabila X	$\frac{y_{ij}-a}{k}$	$\left(\frac{y_{ij}-a}{k}\right)f_{ij}$	$\left(\frac{y_{ij}-a}{k}\right)^2 f_{ij}$
Gr. 1 i = 1	0	0	[illegible]
a = 40	1	3	[illegible]
k = 20	2	4	[illegible]
	Tot	**7**	**[illegible]**
Gr. 2 i = 2	-	-	[illegible]
a = 60	0	0	[illegible]
k = 20	1	4	[illegible]
	2	4	[illegible]
	To	**6**	**[illegible]**
Gr. 3 i = 3	-2	-	[illegible]
a = 100	-	-	[illegible]
k = 20	0	0	[illegible]
	1	8	[illegible]
	2	8	[illegible]
	Tot	-	**[illegible]**
Gr. 4 i = 4	-	-	[illegible]
a= 100	0	0	[illegible]
k = 20	1	4	[illegible]
	2	6	[illegible]
	Total:	8	**[illegible]**
Gr. 5 i = 5	0	0	[illegible]
a= 120;k =	**Tot**	0	**[illegible]**
Pentru	**-**	-	[illegible]
întreaga	-2	-	[illegible]
a =100	-	-	
k = 20	0	**0**	
	1	1	
	2	1	[illegible]
	Tot	-	**[illegible]**

Interpretare: Comparând acest rezultat cu cel aferent mediei dispersiilor de grupă, se observă că productivitatea muncii a avut o influenţă mai mare decât cea exercitată de factorii întâmplători (419 > 379).

4. Dispersia generală.

Este calculată prin procedeul simplificat, pe baza datelor din tabelul 4.2.

$$\sigma_0^2 = \frac{193}{86}\cdot 20^2 - (90-100)^2 = 798$$

Se observă că:

$$\sigma_0^2 = \bar{\sigma}^2 + \delta^2 \text{ ; } 798 = 379 + 419$$

Ponderea influenţelor factoriale asupra variaţiei salariului:

$$P_1 = \frac{379}{798}\cdot 100 = 47{,}5\ \% \text{ ; } P_2 = 100 - P_1 = 52{,}5\ \%$$

Considerăm exemplul respectiv ca suficient de relevant pentru evidenţierea facilităţilor practice pe care le oferă analiza dispersională a distribuţiilor bidimensionale.

4.4 Forme de repartizare a frecvenţelor
4.4.1 Momentele.

Momentele sunt indicatori statistici calculaţi ca medii aritmetice ponderate ale abaterilor variantelor x_i de la o anumită valoare, folosită ca nivel de referinţă, în care abaterile respective sunt luate la puteri diferite.

Momentele calculate din abaterile variantelor x_i de la o constantă oarecare, să presupunem constanta " **a** " cu care ne-am mai întâlnit, sunt considerate **momente obişnuite.** Un exemplu îl constituie **momentul de ordinul 1** folosit în calculul simplificat al mediei aritmetice ponderate corespunzător primei propietăţi,adică:

$$\mu_1^* = \frac{\sum(x_i - a)f_i}{\sum f_i} = \frac{\sum x_i f_i}{\sum f_i} - \frac{a\sum f_i}{\sum f_i} = \bar{x} - a$$

Momentele calculate din abaterile variantelor x_i de la media lor, sunt denumite **momente centrate** şi au o importanţă deosebită în analiza statistică. Un exemplu îl constituie **momentul centrat de ordinul 2,** care nu este altceva decât dispersia cu care ne-am întâlnit în paragrafele precedente, adică:

$$\mu_2 = \frac{\sum(x_i - \bar{x})^2 f_i}{\sum f_i} = \sigma^2$$

Între momentele obişnuite şi momentele centrate pot fi stabilite o serie de relaţii, fiind posibil calculul unora cu ajutorul celorlalte. Folosind notaţiile:

$$x_i - a = c\ ;\ x_i - \bar{x} = d\ ;\ \bar{x} - a = m$$

Atunci:

$$(x_i - a) = (x_i - \bar{x}) + (\bar{x} - a)\ ;\ c = d + m$$

Momentul obişnuit de ordinul **n** va fi:

$$\mu_n^* = \frac{\sum c^n f_i}{\sum f_i} = \frac{\sum(d+m)^n f_i}{\sum f_i} = \frac{\sum d^n f_i}{\sum f_i} + \frac{C_n^1 m \sum d^{n-1} f_i}{\sum f_i} +$$

$$\frac{C_n^2 m^2 \sum d^{n-2} f_i}{\sum f_i} + \cdots + m^n = \mu_n + C_n^1 m \mu_{n-1} + C_n^2 m^2 \mu_{n-2} + \cdots + m^n$$

În mod similar, momentul centrat de ordinul **n** va fi dat de relaţia:

$$\mu_n = \frac{\sum d^n f_i}{\sum f_i} = \frac{\sum(c-m)^n f_i}{\sum f_i} = \frac{\sum c^n f_i}{\sum f_i} - \frac{C_n^1 m \sum c^{n-1} f_i}{\sum f_i} +$$

$$\frac{C_n^2 m^2 \sum c^{n-2} f_i}{\sum f_i} - \cdots (-1)^n m^n = \mu_n^* - C_n^1 m \mu_{n-1}^* + C_n^2 m^2 \mu_{n-2}^* - \cdots (-1)^n m^n$$

Particularizând aceste relaţii, pentru valorile lui n = 0, 1, 2, 3, 4, momentele obişnuite în funcţie de momentele centrate şi invers, momentele centrate în funcţie de momentele obişnuite, se determină astfel:

- pentru n = 0

$$\mu_0^* = \frac{\sum c^0 f_i}{\sum f_i} = 1 \quad ; \quad \mu_0 = \frac{\sum d^0 f_i}{\sum f_i} = 1$$

- pentru n = 1

$$\mu_1^* = \frac{\sum c f_i}{\sum f_i} = \frac{\sum (d+m) f_i}{\sum f_i} = \frac{\sum d f_i}{\sum f_i} + \frac{m \sum f_i}{\sum f_i} = \mu_1 + m = m$$

$$\mu_1 = \frac{\sum d f_i}{\sum f_i} = \frac{\sum (c-m) f_i}{\sum f_i} = \mu_1^* - m = 0$$

In mod analog se fac calculele pentru n=3 si 4.

Momentele prezentate sunt folosite foarte frecvent în statistică, atât ca bază de calcul în determinarea unor indicatori specifici seriilor de distribuţie (cum este -spre exemplu - coeficientul de boltire), cât şi ca modalitate practică de simplificare a calculului multora dintre indicatorii sintetici utilizaţi în analiza statistică - cum ar fi coeficientul de corelaţie lineară. La rândul lor, momentele obişnuite pot fi determinate cu ajutorul procedeului de calcul simplificat, care - în principiu - nu diferă de cel prezentat la calculul mediei aritmetice ponderate şi dispersiei. Astfel:

$$\mu_1^* = \frac{\sum \left[\frac{x_i - a}{k}\right] f_i}{\sum f_i} k \quad ; \quad \mu_2^* = \frac{\sum \left[\frac{x_i - a}{k}\right]^2 f_i}{\sum f_i} k^2 \quad ;$$

$$\mu_3^* = \frac{\sum \left[\frac{x_i - a}{k}\right]^3 f_i}{\sum f_i} k^3 \quad ; \quad \mu_4^* = \frac{\sum \left[\frac{x_i - a}{k}\right]^4 f_i}{\sum f_i} k^4 .$$

Cu ajutorul acestor rezultate vor fi calculate foarte uşor momentele centrate.

$$\mu_1 = \mu_1^* - m = 0 \quad ; \quad \mu_2 = \mu_2^* - m^2 = \sigma^2 \quad ;$$

$$\mu_3 = \mu_3^* - 3m\mu_2^* + 2m^3 \quad ; \quad \mu_4 = \mu_4^* - 4m\mu_3^* + 6m^2\mu_2^* - 3m^4 .$$

Ilustrăm modul de calcul al acestor indicatori folosind datele din tabelul 4.4.

Tab. 4.4. Date necesare calculului simplificat al momentelor.

xi	*fi*	$\frac{x_i - a}{k}$	$\left(\frac{x_i - a}{k}\right) f_i$	$\left(\frac{x_i - a}{k}\right)^2 f_i$	$\left(\frac{x_i - a}{k}\right)^3 f_i$	$\left(\frac{x_i - a}{k}\right)^4 f_i$
10	2	-3	-6	18	-54	162
20	8	-2	-16	32	-64	128
30	12	-1	-12	12	- 12	12
40	15	0	0	0	0	0
50	10	1	10	10	10	10
60	**7**	2	14	28	56	112
70	5	3	15	45	135	405
80	1	4	4	16	64	256
Total	60	*	9	161	135	1085

Momentele obişnuite vor fi:

$$\mu_1^* = \frac{9}{60}10 = 1{,}5 = m \; ; \; \mu_2^* = \frac{161}{60}10^2 = 268{,}333 \; ;$$

$$\mu_3^* = \frac{135}{60}10^3 = 2250 \; ; \; \mu_4^* = \frac{1085}{60}10^4 = 180833{,}33.$$

Momentele centrate vor fi:

$$\mu_1 = 1{,}5 - 1{,}5 = 0 \; ; \; \mu_2 = 268{,}33 - 1{,}5^2 = 266{,}1 = \sigma^2 \; ;$$

$$\mu_3 = 1049{,}25 \quad ; \quad \mu_4 = 170940{,}64.$$

3.4.2 Asimetria

Asimetria vizează în special seriile de distribuţie.

Prin asimetrie, se înţelege abaterea de la simetrie a acestor serii, abatere care poate fi **moderată** - ca în cazul distribuţiilor **moderat asimetrice,** sau o abatere foarte **pronunţată** - ca în cazul distribuţiilor **extrem asimetrice.** Asimetria este cunoscută şi sub numele de oblicitate. Dintre indicatorii folosiţi în măsurarea gradului de asimetrie, **coeficientul lui Pearson** are întrebuinţarea cea mai largă. El se determină cu ajutorul relaţiei:

$$k_{as} = \frac{\bar{x} - Do}{\sigma}$$

In locul lui **Do** poate fi folosit simbolul **Mo** - de la **Mode,** denumirea englezească a **Dominantei.**

Semnificaţie. Când $k_{as} < 0$ - atunci distribuţia prezintă o **asimetrie** spre stânga sau **negativă;** când $k_{as} = 0$ - seria este **simetrică** (cu condiţia ca frecvenţele dispuse de o parte şi de alta a frecvenţei maxime, să fie egale două cate două); când $k_{as} > 0$ - seria prezintă o asimetrie **spre dreapta** sau **pozitivă.**

In figura 4.2 - a, b, c sunt prezentate formele distribuţiilor pentru cele trei grupe de valori ale lui k_{as}.

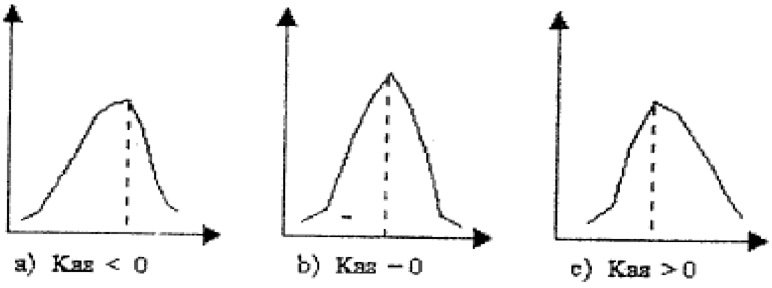


Fig. 4.2 Tipuri de distribuţie după gradul de asimetrie.

Pentru distribuţiile moderat asimetrice, valorile lui sunt cuprinse în intervalul:

$$-0{,}3 < k_{as} < 0{,}3$$

Valorile situate în afara acestui interval semnalează faptul că distribuţiile tind spre (sau au) forme extrem asimetrice.

Coeficientul de asimetrie are o importanţă deosebită în analiza distribuţiilor, putând fi folosit atât pentru aprecierea semnificaţiei mediei, cât şi pentru verificarea normalităţii distribuţiilor empirice, alături de indicatorii curtozisului.

4.4.3 Curtozisul

Se poate observa foarte uşor faptul că distribuţiile unimodale au bolta mai largă sau mai îngustă, după cum gradul de concentrare a frecvenţelor în jurul valorilor tipice - medie, mediana, dominanta - este mai slab sau mai intens.

Curtozisul îşi are rădăcina etimologică în cuvântul grecesc kurtos = cocoşat. Gradul de concentrare a frecvenţelor în jurul valorilor tipice este cunoscut sub numele de exces. Intre exces şi forma distribuţiilor există o legătură foarte strânsă. Astfel:

- dacă excesul este mai puţin pronunţat, frecvenţele fiind destinse pe un interval mai mare al valorilor x_i, atunci distribuţia este **platicurtică** – de la **platos = larg + kurtos** - adică această distribuţie are o boltă largă, cu coada scurtă - ca în figura 4.3 - a ;
- dacă excesul este foarte pronunţat, atunci distribuţia este **leptocurtică - de** la **leptos = îngust + kurtos** - adică distribuţia are o boltă îngustă, cu coada lungă - ca în figura 4.3 - b.

Coeficientul de boltire se calculează ca raport între momentul centrat de ordinul 4 şi pătratul momentului centrat de ordinul 2, folosind relaţia:

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\mu_4}{\left(\sigma^2\right)^2}$$

Excesul este determinat cu ajutorul relaţiei:

$$\gamma_2 = \beta_2 - 3$$

în care 3 reprezintă valoarea coeficintului de boltire aferent repartiţiei normale. Deci:

► în cazul în care $\beta_2 < 3$ - distribuţiile sunt **platicurtice;**

► în cazul în care $\beta_2 > 3$ - distribuţiile sunt **leptocurtice.**

Pentru exemplificare vom folosi datele din tabelul 4.5, în care sunt reluate valorile x_i din tab. 4.4 însă cu o altă repartizare a frecvenţelor. In final vom compara rezultatele aferente celor două distribuţii (tab 4.4 şi tab. 4.5), observând foarte uşor care dintre ele este platicurtică şi care este leptocurtică.

Tab. 4.5. Date necesare pentru calculul momentelor şi curtozisuiui.

xi	*fi*	$\frac{x_i - a}{k}$	$\left(\frac{x_i - a}{k}\right) f_i$	$\left(\frac{x_i - a}{k}\right)^2 f_i$	$\left(\frac{x_i - a}{k}\right)^3 f_i$	$\left(\frac{x_i - a}{k}\right)^4 f_i$
10	1	**-3**	- 3	9	-27	81
20	3	-2	-6	12	-24	48
30	15	- 1	- 15	15	- 15	15
40	22	0	0	0	0	0
50	12	1	12	12	12	12
60	4	2	8	16	32	64
70	2	**3**	6	18	54	162
80	1	4	4	16	64	256
	•		6	98	96	638

$$\bar{x} = \frac{6}{60}10 + 40 = 41\,;\; m = 41 - 40 = 1\,;$$

$$\mu_2^* = \frac{98}{60}10^2 = 163{,}33\,;\; \mu_3^* = \frac{96}{60}10^3 = 1600\,;\; \mu_4^* = \frac{638}{60}10^4 = 106333{,}33\,;$$

$$\mu_2 = 163{,}33 - 1 = 162{,}33\,;\; \mu_4 = 106333{,}33 - 4*1*1600 + 6*1*163{,}33 + -3*1 = 100910{,}3.$$

Coeficintii de boltire şi curtozisul pentru cele două distribuţii au următoarele niveluri:

- pentru distribuţia din tab. 4.4:

$$\beta_2 = \frac{170940{,}64}{266{,}1^2} = 2{,}41\,;\; \gamma_2 = 2{,}41 - 3 = -0{,}59$$

Deci, distribuţia din tab.4.4 este **platicurtică** ; graficul acestei distribuţii este prezentat în fig. 4.3 - a .

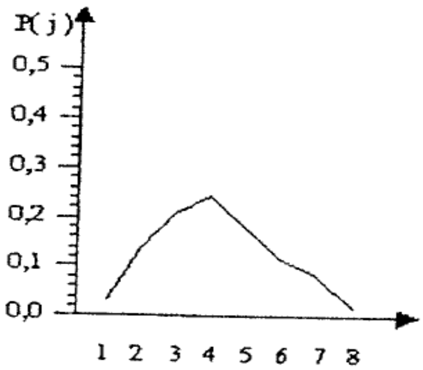
- pentru distribuţia din tab. 4.5:

$$\beta_2 = \frac{100910{,}3}{163{,}33} = 3{,}78\,;\; \gamma_2 = 3{,}78 - 3 = 0{,}78$$

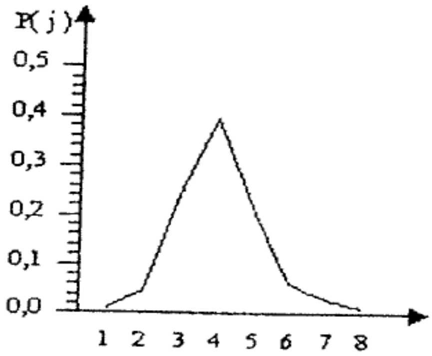
Distribuţia din tabelul 4.5 este **leptocurtică**; graficul corespunzător este prezentat în fig. 4.3 - b .

Cele două diagrame din fig.4.3 sunt comparabile întrucât s-a folosit aceeaşi scară de reprezentare.

Aspectele de ordin practic sunt uşor de sesizat.



a) Distribuţie **platicurtică** ($\beta_2 < 3$)



b) Distribuţie **leptocurtică** ($\beta_2 > 3$)

Fig. 4.3 Forma distribuţiilor după gradul de boltire.