

În vederea analizei productivității obținute în cadrul unei colectivități de salariați formată din 250 de persoane, s-a extras un eșantion format din 30 de salariați. Datele referitoare la producția zilei precedente sunt prezentate în tabelul următor (datele sunt exprimate în bucăți)

Productia (buc) $X_i$	Nr. Salariați $n_i$
20-30	2
30-40	8
40-50	9
50-60	7
60-70	4
Total	30

### 1. Determinați mărimile medii de calcul.

- Media este o măsură a tendinței centrale, iar valoarea sa calculată sintetizează într-un singur nivel reprezentativ tot ceea ce este tipic, esențial, comun și obiectiv în apariția și manifestarea fenomenelor de masă.

- Din cadrul mărimilor medii de calcul vom determina: media aritmetică, media armonică și media pătratică.

#### CONVENȚIE

- În cazul în care se dispune de o distribuție pe grupe, pentru a calcula o valoare apropiată de valoarea reală a mediei, în locul valorilor  $X_i$  se iau în considerare centrele intervalelor de grupare, notate cu  $x_i'$ .

- Această convenție este indusă de următoarea ipoteză: în interiorul tuturor intervalelor de grupare, frecvențele sunt uniform distribuite. În practică, ipoteza efectuată se verifică foarte rar, motiv pentru care între media calculată pe baza centrelor de interval și cea calculată pe baza datelor inițiale există diferențe.

✓ Pentru determinarea mediei **aritmetice** se utilizează următoarea formulă de calcul.

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

$x_i$  - reprezintă nivelurile individuale ale variabilei;  
 $\sum_{i=1}^n x_i$  - reprezintă volumul centralizat al variabilei;  
 $n$  - reprezintă numărul unităților observate.

În vederea identificării valorii de la numărătorul relației de calcul ne vom ajuta de următorul tabel:

Productia (buc) $X_i$	Nr. Salariați $n_i$	$x'$	$X \cdot n$
20-30	2	25	50
30-40	8	35	280
40-50	9	45	405
50-60	7	55	385
60-70	4	65	260
Total	30		1380

Astfel media aritmetica este egală cu  $\frac{1380}{30} = 46$  buc/salariat.

**Interpretare:** Ca atare, în medie în ziua precedentă cei 30 de salariați au produs 46 de bucăți/salariat.

✓ Pentru determinarea mediei **armonice** se utilizează următoarea formulă de calcul.

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

În vederea identificării valorii de la numitorul relației de calcul ne vom ajuta de următorul tabel:

Productia (buc) $X_i$	Nr. Salariați $n_i$	$x'$	$n/x$
20-30	2	25	0,080
30-40	8	35	0,229
40-50	9	45	0,200
50-60	7	55	0,127
60-70	4	65	0,062
Total	30		0,697

Media armonică este egală cu  $\frac{30}{0,697} = 43,04$ . Deoarece producția este o caracteristică

statistică cu variație discretă (discontinuuă), se va rotunji rezultatul obținut la valoarea întreagă 43 buc/salariat.

- ✓ Pentru determinarea mediei **pătratică** se utilizează următoarea formulă de calcul.

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

În vederea identificării valorii de la numărătorul relației de calcul ne vom ajuta de următorul tabel:

Productia (buc) $X_i$	Nr. Salariați $n_i$	$x'$	$x^2 \cdot n$
20-30	2	25	1250
30-40	8	35	9800
40-50	9	45	18225
50-60	7	55	21175
60-70	4	65	16900
Total	30		67350

Media pătratică este egală cu  $\sqrt{\frac{67350}{30}} = 47,38$ . Deoarece producția este o caracteristică statistică cu variație discretă (discontinuuă), se va rotunji rezultatul obținut la valoarea întreagă 47 buc/salariat.

Așa cum se observă, valorile calculate pentru aceste mărimi medii de calcul nu sunt egale, mai mult se știe că acestea respectă relația de inegalitate  $\bar{x}_h \leq \bar{x}_g \leq \bar{x} \leq \bar{x}_p$ .

Interpretarea mediilor se va realiza după același tipar prezentat în cazul mediei aritmetice.

## 2. Determinați mărimile medii de poziție.

Se vor calcula în continuare valoare dominantă (modul) și mediana.

✓ **Valoarea modală a caracteristicii** (numită și valoare dominantă, valoarea cea mai probabilă sau modul) reprezintă acea valoare a caracteristicii care corespunde celui mai mare număr de unități sau aceea care are cea mai mare frecvență de apariție.

Relația de calcul în cazul unei distribuții grupate pe intervale, este:

$$Mo = x_0 + h \frac{\Delta_1}{\Delta_1 + \Delta_2}$$

în care:

- $x_0$  - reprezintă limita inferioară a intervalului modal;
- $h$  - mărimea intervalului modal;
- $\Delta_1$  - diferența dintre frecvența intervalului modal și a celui precedent;
- $\Delta_2$  - diferența dintre frecvența intervalului modal și a celui următor.

Într-o primă fază se identifică intervalul modal analizând coloana frecvențelor  $n_i$  și identificând valoarea cea mai mare. În cazul nostru frecvența cea mai mare este 9, aceasta fiind corespunzătoare intervalului (40-50), ca atare, acesta va fi intervalul modal.

Astfel, valoarea dominantă este egală cu  $40 + 10 \frac{(9-8)}{(9-8) + (9-7)} = 43,3$  Deoarece producția

este o caracteristică statistică cu variație discretă (discontinuuă), se va rotunji rezultatul obținut la valoarea întreagă 43 buc.

Interpretare: Astfel, producția obținută de către cei mai mulți salariați, din rândul celor 30 analizați, a fost de 43 buc.

✓ Mediana împarte numărul unităților investigate în două părți egale.

În vederea determinării acestui indicator se parcurg următoarele etape:

a. Calculul unității mediane  $U^{me} = \frac{n+1}{2} = \frac{\sum_{i=1}^k n_i + 1}{2}$ ,  $U^{me} = \frac{30+1}{2} = 15,5$

b. se determină frecvențele cumulate  $N_i$

Productia (buc) $X_i$	Nr. Salariați $n_i$	Frecv. cumulate descrescator $N_i$
20-30	2	2
30-40	8	10
40-50	9	19
50-60	7	26
60-70	4	30
Total	30	

c. se identifică intervalul median prin comparare  $N_i \geq U^{me}$

Astfel, frecvența cumulată cea mai apropiată de unitatea mediană și care respectă condiția de mai sus este 19, aferentă intervalului 40-50.

d. Utilizarea următoarei relații pentru determinarea valorii medianei

$$Me = x_0 + h \cdot \frac{\frac{\sum_{i=1}^k n_i + 1}{2} - \sum_{i=1}^{m-1} n_i}{n_m}$$

unde:

$h$  - mărimea intervalului median;

$m$  - indexul intervalului median;

$\sum_{i=1}^{m-1} n_i$  - suma frecvențelor precedente intervalului median (frecvența cumulată a intervalului precedent celui median);

$n_m$  - frecvența absolută a intervalului median.

$$Me = 40 + 10 \frac{15,5 - 10}{9} = 46,11$$

Interpretare: Astfel, valoarea producției ce împarte distribuția aferentă celor 30 de salariați în 2 părți de volum egal este 46 buc.

În cazul unei distribuții unimodale ușor asimetrice, frecvențele sunt ușor deplasate într-o parte sau alta, între cei trei indicatori ai tendinței centrale există următoarea relație  $\bar{x} - M_0 = 3(\bar{x} - M_e)$ , care se respectă și în cadrul aplicației noastre.

### 3. Caracterizați variabilitatea față de tendința centrală cu ajutorul indicatorilor simpli și sintetici ai variației.

✓ Un indicator simplu al variației, care oferă o imagine generală asupra dispersiei, este distanța dintre nivelul maxim și nivelul minim al variabilei sau **amplitudinea variației**.

$$A_x = X_{\max} - X_{\min}$$

$$A_x = 70 - 20 = 50 \text{ buc}$$

Interpretare: Distanța dintre nivelul maxim și nivelul minim al producției realizate este de 50 de bucăți-producție.

✓ abaterile *nivelurilor individuale* ale unei variabile față de nivelul lor mediu, notate cu  $d_i$ , și care se pot calcula după formula:

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

Media aritmetică a distribuției noastre a fost determinată și are valoarea  $\bar{x} = 46$  buc/salariat

Determinarea abaterilor individuale față de media aritmetica a distribuției se prezintă în următorul tabel:

Productia (buc) $X_i$	Nr. Salariați $n_i$	$x'$	$d_i$	$d_i\%$
20-30	2	25	-21	-46%
30-40	8	35	-11	-24%
40-50	9	45	-1	-2%
50-60	7	55	9	20%
60-70	4	65	19	41%
Total	30			

✓ Abaterile  $d_i$  se pot exprima și în mărimi relative prin raportarea lor la nivelul mediu al variabilei, astfel:

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

În practic aceste abateri se utilizează pentru calculul ulterior al indicatorilor sintetici ai variației

**Interpretare:** Se observă că prima grupă de salariați au reușit să producă cu 21 de bucăți mai puțin decât media colectivității sau putem prezenta în formă relativă că au produs cu 46% mai puțin decât media colectivității. În mod asemănător se interpretează și valorile determinate pentru intervalele de grupare următoare cu precizarea că valorile pozitive exprimă depășiri ale nivelului mediu al colectivității.

**Indicatorii sintetici ai variației** se caracterizează prin faptul că, în calculul lor, se includ abaterile tuturor nivelurilor individuale ale unei variabile față de nivelul lor mediu.

✓ **Abaterea medie liniară**, notată cu  $\bar{d}$ , se determină ca o medie aritmetică (simplă sau ponderată) a abaterilor nivelurilor individuale ale unei variabile față de media lor aritmetică, considerate

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

Abaterea medie liniară se poate calcula și față de mediană.

Determinarea valorii de la numărătorul relației o vom realiza cu ajutorul următorului tabel:

Productia (buc) $X_i$	Nr. Salariați $n_i$	$x'$	$d_i$	$ x_i - \bar{x}  n_i$
20-30	2	25	-21	42
30-40	8	35	-11	88
40-50	9	45	-1	9
50-60	7	55	9	63
60-70	4	65	19	76
Total	30			278

$\bar{d} = \frac{278}{30} = 9,26$  Deoarece producția este o caracteristică statistică cu variație discretă (discontinuu), se va rotunji rezultatul obținut la valoarea întreagă 9 buc.

**Interpretare:** Putem concluziona că fiecare salariat, din rândul celor 30 analizați, s-a abătut de la producția medie a eșantionului selectat, în medie cu 9 buc/salariat.

✓ **Dispersia**, notată cu  $\sigma^2$ , se calculează ca o medie aritmetică a pătratelor abaterilor nivelurilor individuale ale unei variabile de la nivelul lor mediu (media aritmetică).

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

Productia (buc) $X_i$	Nr. Salariați $n_i$	$x'$	di	$(x_i - \bar{x})^2 n_i$
20-30	2	25	-21	882
30-40	8	35	-11	968
40-50	9	45	-1	9
50-60	7	55	9	567
60-70	4	65	19	1444
Total	30			3870

$$\sigma^2 = \frac{3870}{30} = 129$$

**Interpretare:** Dispersia este un indicator cu un caracter economic abstract, rezultatul obținut s-ar exprima în  $\text{buc}^2$ , ceea ce din pct. de vedere economic nu poate fi interpretat. Astfel, valoarea acestui indicator nu poate fi interpretată din pct. de vedere economic și nici nu se va prezenta vreo unitate de măsură.

✓ **Abaterea medie pătratică**, numită și abatere standard sau abatere tip se notează cu  $\sigma$  și se calculează prin extragerea rădăcinii pătrate din dispersie. Abaterea medie pătratică reprezintă media pătratică a abaterilor nivelurilor individuale ale caracteristicii față de media lor.

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

$\sigma = \sqrt{\frac{3870}{30}} = 11,35$  Deoarece producția este o caracteristică statistică cu variație discretă (discontinuu), se va rotunji rezultatul obținut la valoarea întreagă 11  $\text{buc/salariat}$ .

**Interpretare:** Interpretarea abaterii pătratice se realizează după modelul prezentat în cazul abaterii liniare.

✓ **Coeficientul de variație**, notat cu  $v_\sigma$ , se calculează ca raport între abaterea medie și nivelul mediu al unei variabile, iar pentru a se interpreta mai ușor se exprimă în procente:

$$v_\sigma = \frac{\sigma}{\bar{x}} \cdot 100 \quad v_d = \frac{\bar{d}}{\bar{x}} \cdot 100$$

- Cu cât valorile sale sunt mai apropiate de zero, cu atât seria este mai omogenă (media este mai reprezentativă)
- Cu cât valorile sale sunt mai apropiate de 100 cu atât ansamblul valorilor individuale observate este mai eterogen (împrăștierea este mai mare, iar media calculată este mai puțin reprezentativă).
- Practica utilizării coeficientului de variație a stabilit pragul de trecere de la starea de omogenitate la cea de eterogenitate: dacă  $Cv \leq 35\%$  colectivitatea este omogenă; dacă  $Cv > 35\%$  colectivitatea este eterogenă.

$$v_{\sigma} = \frac{11}{46} \cdot 100 = 24,69\% \qquad v_d = \frac{9}{46} \cdot 100 = 20,14\%$$

Interpretare: Valorile determinate ale coeficientului de variație se situează sub valoarea de 35%, astfel concluzionăm că eșantionul celor 30 de salariați analizați este unul omogen din perspectiva variabilei producție și, totodată, rezultă că și media producției este reprezentativă din punct de vedere statistic.

**4. Determinați eroarea maximă admisibilă între nivelul mediu al producției descoperit la nivelul eșantionului și nivelul mediu al producției realizat la nivelul întregii colectivități de salariați format, așa cum se precizează în enunț, din 250 persoane. Determinați ulterior nivelul eșantionului necesar a fi extras dacă se dorește o reducere a erorii maxime admisibile descoperite anterior cu 30%.**

Calculul se va realiza cu o probabilitate de lucru de 95,45% pentru care coeficientul  $z=2$ .

**Prezentați calculele atât pentru procedeul extragerii repetate cât și pentru procedeul extragerii nerepetate.**

✓ Procedeul extragerii repetate

Eroarea maximă admisibilă se determină cu ajutorul relației  $\Delta_x = z \cdot s_x^-$

În cadrul acestei relații  $s_x^-$  semnifică eroarea standard (eroarea medie de reprezentativitate), ce poate fi determinată cu ajutorul relației de mai jos:

$$s_x^- = \sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{129}{30}} = 2,07 \text{ buc/salariat}$$

$\Delta_x = 2,07 \cdot 2 = 4,14$ . Deoarece producția este o caracteristică statistică cu variație discretă (discontinuu), se va rotunji rezultatul obținut la valoarea întregă 4.

Interpretare: Se garantează cu o probabilitate de 95,45% o eroare maximă între nivelul mediu al producției descoperit la nivelul eșantionului și nivelul mediu al producției realizat la nivelul întregii colectivități de salariați, în cazul unei extrageri repetate, de 4 buc/ salariat.

Nivelul eșantionului, pentru extragerea repetată, se determină cu ajutorul relației

$$n = \frac{z^2 \sigma^2}{\Delta_x^2}$$

Eroarea maximă admisibilă obținută anterior se reduce cu 30%, astfel  $\Delta_x = 2,8$  buc/salariat



$$n = \frac{2^2 * 129}{2,8^2} = 65,81$$

. Fiind vorba despre un număr de salariați va trebui sa rotunjim rezultatul la 66 persoane necesare a fi extrase pentru a forma un eșantion care să garanteze cu o probabilitate de 95,45% o eroare maximă admisibilă de 2,8 buc/salariat.

✓ **Pocedeul extragerii nerepetate**

Eroarea maximă admisibilă se determină cu ajutorul relației  $\Delta_x = z \cdot s_x^-$

În cadrul acestei relații  $s_x^-$  semnifică eroarea standard (eroarea medie de reprezentativitate), ce poate fi determinată cu ajutorul relației de mai jos:

$$s_x^- = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)} = \sqrt{\frac{129}{30} \left(1 - \frac{30}{250}\right)} = 1,94 \text{ buc/salariat}$$

$\Delta_x = 1,94 * 2 = 3,88$ . Deoarece producția este o caracteristică statistică cu variație discretă (discontinuu), se va rotunji rezultatul obținut la valoarea întreagă 4.

**Interpretare:** Se garantează cu o probabilitate de 95,45% o eroare maximă între nivelul mediu al producției descoperit la nivelul eșantionului și nivelul mediu al producției realizat la nivelul întregii colectivități de salariați, în cazul unei extrageri nerepetate, de 4 buc/ salariat.

Nivelul eșantionului, pentru extragerea nerepetată, se determină cu ajutorul relației

$$n = \frac{z^2 \sigma^2}{\Delta_x^2 + \frac{z^2 \sigma^2}{N}}$$

Eroarea maximă admisibilă obținută anterior se reduce cu 30%, astfel  $\Delta_x = 2,8 \text{ buc/salariat}$

$$n = \frac{2^2 * 129}{2,8^2 + \frac{2^2 * 129}{250}} = 52,10. \text{ Fiind vorba despre un număr de salariați va trebui sa rotunjim}$$

rezultatul la 52 persoane necesare a fi extrase pentru a forma un eșantion care să garanteze cu o probabilitate de 95,45% o eroare maximă admisibilă de 2,8 buc/salariat.