

2. INDICATORI DE POZITIE SI DE VARIATIE

INDICATORI DE POZITIE

- **Media.** În Excel se calculeaza folosind functia AVERAGE.
- **Modulul.** În Excel se calculeaza folosind functia MODE.
- **Valoarea mediana.** În Excel se calculeaza folosind functia MEDIAN.

INDICATORI DE VARIATIE

- **Dispersia.** Se calculeaza în Excel folosind functia VAR.
- **Abaterea medie.** Se foloseste formula de calcul $\frac{\sum |x - \bar{X}|}{n-1}$, unde n este dimensiunea esantionului
- **Abaterea standard.** Se utilizeaza functia Excel STDEV
- **Coefficientul de variatie (CV)** exprimat procentual prin formula

$$CV = \frac{S}{\bar{x}} \cdot 100$$

Coefficientul de variatie se foloseste pentru compararea a doua sau mai multe repartitii de frecvente din punctul de vedere al variatiei lor.

Interpretarea CV a unei repartitii de frecvente.

1. $CV < 10\%$ inseamna ca repartitia prezinta o variatie mica.
2. $10\% < CV < 20\%$, repartitia are o variatie medie.
3. $CV > 20\%$, repartitia are o variatie mare.

CUANTILE

$$Q_k = \frac{k(n+1)}{q}$$

unde

- n reprezinta numarul total de date.
- k este indexul cuantilei dorite.
- q reprezinta numarul grupurilor în care se împarte setul de date de catre cuantilele luate în considerare.

CALCULUL CUARTILELOR ÎN EXCEL

În Excel, calculul cuantilelor de ordinul q se poate realiza doar prin implementarea formulei generale de calcul. Pentru calculul cuartilelor ($q = 4$) în Excel avem la dispoziție funcția statistică QUARTILE. Aceasta are sintaxa

=QUARTILE(array,quart)

unde

- array reprezintă domeniul de celule conținând valorile caracteristicii studiate
- quart este un argument ce poate avea valorile 0, 1, 2, 3, sau 4, fiecare valoare având o anumită semnificație.

Pentru valoarea 0 se returnează valoarea minimă din setul de date.

Pentru valoarea 1 se returnează valoarea primei cuartile, Q_1 .

Pentru valoarea 2 se returnează valoarea celei de a doua cuartile, Q_2 , ce coincide cu valoarea mediană.

Pentru valoarea 3 se returnează valoarea celei de a treia cuartile, Q_3 .

Pentru valoarea 4 se returnează valoarea maximă din setul de date.

Intervalul intercuartilic = $Q_3 - Q_1$



Exercițiul 1

Unui lot de 30 de pacienți cu afecțiuni cardiace i se administrează o doză zilnică de 0.25 mg. digoxin. După o perioadă de timp, pentru fiecare pacient se măsoară concentrația plasmatică de digoxin (în ng/ml). Datele obținute sunt prezentate în tabelul de mai jos. Introduceți aceste date în foaia de calcul Excel.

| | | | | |
|------|------|------|------|------|
| 1.56 | 0.5 | 0.77 | 0.85 | 1.06 |
| 0.82 | 0.54 | 0.65 | 1.31 | 0.63 |
| 0.7 | 0.83 | 0.92 | 0.87 | 0.68 |
| 0.89 | 1.12 | 0.83 | 0.62 | 0.92 |
| 0.93 | 1.09 | 1.13 | 0.59 | 0.79 |
| 0.98 | 1.01 | 0.74 | 0.76 | 1.25 |
| 1.22 | 0.79 | 1.06 | 1.01 | 0.53 |
| 1.32 | 0.68 | 0.98 | 1.03 | 0.67 |

Tabel 1. Concentrația plasmatică digoxin ng/ml

1. Calculati media, valoarea mediana si valoarea modala folosind functiile Excel AVERAGE, MEDIAN si respective MODE.
2. Calculati amplitudinea variatiei utilizând valorile extreme ale setului de date într-o formula de calcul Excel.
3. Calculati dispersia si abaterea standard apelând functiile statistice Excel VAR respective STDEV.
4. Calculati deviatia medie implementând formula de calcul corespunzatoare.

Exercitiul 2

În tabelul de mai jos sunt prezentate valorile concentratiilor de colesterol în sânge, masurate în mg/dl, pentru un esantion de 50 de pacienti.

| | | | | |
|-----|-----|-----|-----|-----|
| 250 | 200 | 240 | 210 | 180 |
| 160 | 210 | 170 | 240 | 140 |
| 160 | 220 | 150 | 260 | 150 |
| 180 | 170 | 140 | 180 | 190 |
| 145 | 220 | 150 | 170 | 210 |
| 220 | 210 | 230 | 140 | 220 |
| 230 | 180 | 250 | 230 | 230 |
| 240 | 170 | 260 | 240 | 20 |
| 190 | 160 | 180 | 250 | 180 |
| 160 | 190 | 220 | 260 | 200 |

Tabel 2. Concentratiile colesterolului din sânge pentru un lot de 50 de pacienti

1. Calculati media, mediana si valoarea modala.
2. Calculati amplitudinea variatiei, dispersia si abaterea standard.
3. Calculati frecventele absolute si relative ale valorilor variabilei aleatoare “concentratia colesterolului” si realizati histogramele frecventelor.
4. Determinati cuartilele Q_1 , Q_2 si Q_3 care impart setul de date în patru sectiuni cu numar egal de valori: fiecare parte contine 25% din numarul total de date. Apelati functia statistica Excel QUARTILE. Determinati intervalul intercuartilic.

5. Determinati cvintilele Q1, Q2, Q3 si Q4 care împart setul de date în cinci sectiuni cu numar egal de date: fiecare parte contine 20% din totalul datelor. Utilizati formula generala de calcul a cuantilelor pentru $q = 5$ implementând-o ca formula de calcul Excel.



Exercitiul 3

În procesul de productie a remediilor homeopate una dintre etape presupune evaluarea dimensiunilor granulelor inerte. Aceasta operatiune se realizeaza prin cântariri repetate a câte 80 de granule. O greutate de $1 \text{ gr.} \pm 15\%$ a acestora indica încadrarea dimensiunilor în limitele normale.

Se doreste compararea dimensiunilor fizice a granulelor inerte homeopate preparate în doua loturi diferite. În acest scop se realizeaza 30 de cântariri a câte 80 de granule din fiecare lot. În tabelul de mai jos sunt prezentate datele obtinute.

| Esantion din lotul 1 | | | Esantion din lotul 2 | | |
|-----------------------|------|------|-----------------------|------|------|
| gr./80 granule inerte | | | gr./80 granule inerte | | |
| 1.5 | 1.45 | 1.11 | 1.12 | 1.11 | 0.93 |
| 1.5 | 1.23 | 1.1 | 1.1 | 1.1 | 0.91 |
| 1.02 | 1.22 | 1.08 | 1.07 | 1.08 | 0.89 |
| 1.45 | 1.21 | 1.06 | 1.05 | 1.06 | 0.87 |
| 1.23 | 1.2 | 1.04 | 1.03 | 1.04 | 0.85 |
| 1.09 | 1.19 | 1.02 | 1 | 1.02 | 0.83 |
| 1.11 | 1.18 | 1 | 0.98 | 1 | 0.81 |
| 1 | 1.17 | 0.98 | 0.95 | 0.98 | 0.79 |
| 1.5 | 1.16 | 0.96 | 0.93 | 0.96 | 0.77 |
| 1.19 | 1.2 | 0.94 | 0.9 | 0.94 | 0.76 |

1. Pentru fiecare esantion în parte realizati o împartire în clase a datelor si calculati frecventele absolute si relative ale claselor. Realizati histogramele corespunzatoare.
2. Calculati mediile, dispersiile si abaterile standard corespunzatoare fiecarui esantion.
3. Calculati coeficientii de variatie pentru cele doua seturi de date în ce tipuri de variatie se înscriu cele doua esantioane.