**VARIABILE CANTITATIVE: STATISTICA DESCRIPTIVĂ**

**& REPREZENTARE GRAFICĂ**

***Andreea Gheban-R., Cosmina Bondor, Andrada Urda-C., Sorana D. Bolboacă***

**Obiective educaționale**

La sfârșitul acestei lucrări practice, studentul va fi capabil să:

* Explice semnificația indicatorilor de centralitate (media aritmetică, mediana, modulul, valoarea centrală), dispersie (amplitudine, deviația standard, coeficientul de variație: CV = stdev/medie\*100; și localizare (cuartila 1 și 3), simetrie (asimetria și boltirea).
* Cunoască și să aplice funcții predefinite Excel de calcul al indicatorilor de centralitate (AVERAGE, MEDIAN, MODE, MIN, MAX), dispersie (STDEV.S), asimetrie (SKEW, KURT) și localizare (QUARTILE.INC).
* Cunoască și să aplice funcții definite de utilizator pentru amplitudine, valoare centrală, coeficient de variație.
* Să interpreteze indicatorii de centralitate, dispersie, asimetrie și localizare.
* Să realizeze reprezentări grafice adecvate pentru evaluarea distribuției unei variabile (histograma sau cutia cu mustăți) și a relației dintre două variabile cantitative (graficul de tip nor de puncte).

**SCENARIUL**

S-a evaluat profilul lipidic al pacienților spitalizați în luna martie 2023 în secția de medicină internă a unui spital din județul Cluj (Dej). S-au colectat din fișele medicale ale pacienților următoarele variabile: Vârsta (ani), Sex (F/M), Greutate (kg), Înălțime (cm), Ritmul Cardiac (RC/min), Presiunea Arterială Sistolică (PAS (mmHg)), Presiunea Arterială Diastolică (PAD (mmHg)), Glicemia (mg/dL), Colesterol (mg/dL), Trigliceride (TG (mg/dL)), HDL (mg/dL), Durata Spitalizării (DS (zile)).

Datele colectate sunt disponibile în fișierul ***BD-DateSpitalizare.xlsx***.

**CERINȚE**

**Scop**: Identificarea variabilelor cantitative discrete. Exercițiul 1.

1. În **Tabel nr. 1** *Calculul indicatorilor statistici* din foaia de calcul ***“Data”****(*vezi **BD-DateSpitalizare.xlsx***)* colorați cu fond albastru **variabilele cantitative discrete.**

**Scop**: Calcularea indicatorilor statistici de centralitate, dispersie, simetrie, boltire și localizare. pentru variabile cantitative continue și discrete. Exercițiul 2 și 3.

1. -3. Calculați în **Tabel nr. 1** *Calculul indicatorilor statistici* din foaia de calcul ***“Data”*** *(*vezi **BD-DateSpitalizare.xlsx***)* următorii indicatori statistici pentru toate **variabilele cantitative continue** (vezi fișierul ***Instrucțiuni***– pag. 1-3):

|  |  |  |
| --- | --- | --- |
| **Centralitate**:   * Media aritmetică * Mediana * Modulul * Valoarea centrală |  | **Dispersie**:   * Amplitudinea * Variația * Deviația standard * Eroarea standard * Coeficientul de variație |
|  |  |  |
| **Simetrie**:   * Asimetrie * Boltire |  | **Localizare**:   * Cuartile (Q0 = min; Q1, Q2, Q3, Q4 = max) * Percentile (80%, 90%, 95%) |

1. Conform marcării efectuate la Exercițiul 1, în cazul variabilelor **cantitative discrete** calculăm în Tabelul 1 doar: mediana şi modulul (indicatori de centralitate), coeficientul de variație (*utilizați formula CQV*) și indicatorii de localizare.

**Scop**: Interpretarea indicatorilor statistici de centralitate, dispersie, simetrie, boltire și localizare pentru variabile cantitative. Exercițiul 4a. și 4b.

1. În **Tabel nr. 2** scrieți interpretările indicatorilor statistici calculați pentru
   1. o variabilă cantitativă continuă
   2. o variabilă cantitativă discretă

(vezi foaia „***Data”*** din**BD-DateSpitalizare.xlsx**)

**Scop**: Reprezentarea unei variabile cantitative cu ajutorul graficului: Histogramă. Exercițiile 5-7.

1. În foaia de calcul ***“Histograma”***, creați histograma pentru variabila **Colesterol** pe întregul grup de pacienţi. Adăugaţi pe grafic valorile, titlul axelor (OX *– Clase de colesterol (mg/dL),* OY *– Numărul de subiecți*)şi titlul graficului (*Histograma colesterolului*) (vezi ***Instrucțiuni*** – pag. 4-5).
2. În foaia de calcul ***“Histograma”***, creați histograma pentru variabila **Colesterol** pentru pacienţii de sex feminin. Adăugaţi pe grafic valorile, titlul axelor (OX *– Clase de colesterol (mg/dL),* OY *– Numărul de subiecți*)şi titlul graficului (*Histograma colesterolului la femei*) (vezi ***Instrucțiuni*** – pag. 5-6).
3. În foaia de calcul ***“Histograma”***, creați histograma pentru variabila **Colesterol** pentru pacienţii de sex masculin. Adăugaţi pe grafic valorile, titlul axelor (OX *– Clase de colesterol (mg/dL),* OY *– Numărul de subiecți*)şi titlul graficului (*Histograma colesterolului la bărbați*) (vezi ***Instrucțiuni*** – pag. 6).

**Scop**: Reprezentarea unei variabile cantitative defalcate pe grupuri (conform unei variabile calitative) cu ajutorul graficului: Cutie cu mustăți. Exercițiile 8-9.

1. În foaia de calcul ***“Cutie cu mustăți”***, reprezentați grafic distribuția variabilei **Colesterol** în funcție de **Sex**. Adăugaţi pe grafic valorile, titlul axelor (OX – *Sex,* OY – *Colesterol (mg/dL)* şi titlul graficului (*Distribuția colesterolului în funcție de sex*) (vezi ***Instrucțiuni*** – pag. 6-8).
2. Completați Tabelul 3, folosindu-vă de graficul creat la cerința 8 (vezi foaia de calcul **Cutia cu mustăți** și fișierul ***Instrucțiuni*** – pag. 8):

|  |  |  |
| --- | --- | --- |
| **Indic. localizare** | **Femei** | **Bărbați** |
| Minim |  |  |
| Cuartila 1 |  |  |
| Media aritmetică |  |  |
| Mediana |  |  |
| Cuartila 3 |  |  |
| Maxim |  |  |

**Scop**: Reprezentarea relației dintre două variabile cantitative cu ajutorul graficului: Nor de puncte. Exercițiul 10.

1. În foaia de calcul ***“Nor de puncte”***, reprezentaţi grafic relația dintre **Greutate** și **Colesterol** cu ajutorul unui grafic de tip **Nor de puncte (scatter)**. Adăugaţi pe grafic titlul axelor (OX – *Greutate (kg)*, OY – *Colesterol (mg/dL)*) şi titlul graficului (*Relația între Greutate și Colesterol*) (vezi ***Instrucțiuni*** – pag. 8-10).

**Exerciții recapitulative:**

1. Calculați în foaia de calcul ***“Recapitulare”*** indicatorii statistici de la punctul 1 pentru variabilele cantitative, separat pentru grupul de femei și pentru grupul de bărbați (vezi **BD-DateSpitalizare.xlsx**,)
2. Alegeți o variabilă cantitativă continuă și o variabilă cantitativă discretă de la exercițiul anterior. Interpretați indicatorii calculați pentru acestea la punctul anterior. Faceți interpretarea pentru fiecare sex în parte. (vezi **BD-DateSpitalizare.xlsx**, foaia de calcul ***“Recapitulare”***)
3. Într-o nouă foaie de calcul ***“Histograma2”***, creați histograma pentru variabila **PAS**. Adăugaţi pe grafic titlul axelor, valorile şi titlul graficului.
4. În foaia de calcul ***“Histograma2”***, creați histograma pentru variabila **PAS** pentru pacienţi de sex masculin. Adăugaţi pe grafic titlul axelor, valorile şi titlul graficului.
5. În foaia de calcul ***“Nor de puncte”***, reprezentaţi grafic pentru pacienții de sex feminin relația dintre **Greutate** și **Colesterol** cu ajutorul unui grafic de tip **Nor de puncte (scatter)**. Adăugaţi pe grafic titlul axelor şi titlul graficului.
6. În foaia de calcul ***“Nor de puncte”***, reprezentaţi grafic pentru pacienții de sex masculin relația dintre **Greutate** și **Colesterol** cu ajutorul unui grafic de tip **Nor de puncte (scatter)**. Adăugaţi pe grafic titlul axelor şi titlul graficului.

**DE REȚINUT!**

* Statistica descriptivă este utilizată pentru a descrie distribuţia unei variabile.
* Histograma este adecvată pentru descrierea distribuției unei SINGURE variabile CANTITATIVE.
* Relația dintre DOUĂ variabile CANTITATIVE va fi descrisă printr-un grafic nor de puncte (Scatter). Datele sunt prezentate ca o colecție de puncte, fiecărui punct corespunzându-i valoarea variabile de pe axa OX şi respectiv valoarea celei de-a doua variabile de pe axa OY. Variabila de pe axa OX este variabila independentă; Variabila de pe axa OY este variabila dependentă. Această regulă se aplică atunci când există o explicație rațională de dependență dintre două variabile.