

RECAPITULARE STATISTICĂ DESCRIPTIVĂ

Mădălina Văleanu

OBIECTIVE EDUCAȚIONALE

La sfârșitul acestei lucrări practice, studentul va recapitula următoarele noțiuni parcurse până la acest moment:

- Crearea de variabile derivate utilizând formule predefinite Excel sau funcții definite de utilizator
- Realizarea graficelor corespunzătoare tipului și numărului de variabile reprezentate
- Interpretarea reprezentărilor grafice
- Calcularea frecvențelor
- Calcularea și interpretarea indicatorilor de statistică descriptivă
- Calcularea și interpretarea intervalului de încredere

SCENARIUL

S-a evaluat profilul lipidic în relație cu nivelul sangvin de calciu, fumat și hepatita cronică la un eșantion de 122 subiecți. Au fost colectate datele corespunzătoare următoarelor variabile: HepatitaCronica (da/nu), Fumator (da/nu), Sex (F/M), Vârstă (ani), Greutate (kg), Înălțime (cm), Colesterol (mg/dL), TG (mg/dL), HDL (mg/dL), și CalciuTotal (mg/dL), unde TG = trigliceride (valori normale 35-149 mg/dL la femei și 40-149 mg/dL bărbați), HDL = *High-Density Lipoprotein* (< 40 mg/dL se asociază cu risc crescut de boală cardiovasculară și se consideră ca valori scăzute; valorile între 41 și 59 sunt considerate valori de graniță iar valorile ≥ 60 sunt considerate normale). Datele colectate se regăsesc în fișierul BD_ProfilLipidic.xlsx.

- Creați un folder LP09.

- Salvați baza de date în noul folder creat cu denumirea BD_ProfilLipidic_PrenumeNUME.xlsx, unde PrenumeNUME este prenumele si NUMELE dvs.

CERINȚE

1. Creați la dreapta coloanei Înălțime o nouă coloană denumită IMC (indice de masă corporală). Pentru fiecare pacient calculați indicele de masă corporală (IMC) utilizând formula:

$$IMC = \text{greutate(kg)} / (\text{înălțime(m)})^2$$

2. Creați la dreapta coloanei HDL o nouă coloană denumită LDL (LDL = *Low-Density Lipoprotein*). Pentru fiecare pacient calculați valoarea LDL colesterol utilizând formula:

$$LDL \text{ colesterol} = \text{Colesterol} - HDL - \frac{TG}{5}$$

3. Pentru fiecare pacient verificați dacă valoarea colesterolului este *Normală* (până la 200 mg/dL), *La Limită* (între 201 – 240 mg/dL) sau *Crescută* (peste 240 mg/dL). Codificați această stare în variabila **NivelColesterol** pe care o creați la dreapta coloanei **Colesterol**.
4. Creați tabelul de frecvență pentru variabila **HepatitaCronica**.
5. Creați tabelul de contingență pentru variabilele **HepatitaCronica** și **Sex**. Pe baza acestui tabel de contingență calculați:
 - a. proporția bărbaților în grupul celor care au avut diagnosticul de hepatită cronică
 - b. raportul bărbați / femei în grupul celor cu hepatita cronică
 - c. calculați riscul de a avea hepatită cronică la bărbați față de femei. Calculați intervalul de încredere de 95% asociat riscului relativ. Interpretați riscul relativ și intervalul lui de încredere de 95% într-o celulă alăturată rezultatelor.
6. Realizați histograma variabilei **CalciuTotal**.

7. Pentru variabila **CalciuTotal** calculați, folosind funcțiile predefinite ale Excel: media aritmetică, mediana, modulul, deviația standard, cvartila 1, cvartila 3, asimetria și boltirea, coeficientul de variație. Apreciați pe baza acestor valori dacă variabila **CalciuTotal** urmează o distribuție normală de probabilitate. Interpretați nivelul de omogenitate pentru variabila **CalciuTotal**. Aprecierea distribuției și a nivelului de omogenitate se va include în într-o celulă alăturată rezultatelor indicatorilor.
8. Calculați intervalul de încredere de 95% asociat mediei **calciului total** și interpretați-l într-o celulă alăturată intervalului.
9. Reprezentați grafic valorile **CalciuTotal** în funcție de **NivelColesterol**.
10. Reprezentați grafic distribuția în funcție de **Sex**.
11. Reprezentați grafic repartiția **Fumătorilor** în funcție de **Sex**.

RAPORTARE REZULTATE

Creați un document Word pe care îl veți denumi *Rezultate-LP09-PrenumeNUME* și care va respecta următoarele cerințe:

1. Toate paginile să aibă format A4.
2. Prima pagină să fie formatată diferit (**different first page**).
3. Pagina 1: introduceți un titlu (dimensiune font: 26), autorul (dimensiune font: 20), afilierea autorului (dimensiune font: 12). Treceți la pagină nouă folosind Page Break.
4. Pagina 2: această pagină va conține titlul de subcapitol: **Cuprins**, iar apoi cuprinsul inserat automat al documentului realizat. Treceți la pagină nouă (Page Break).
5. Pagina 3: Această pagină va conține titlul de subcapitol: **Scenariu**. Copiați în această pagină scenariul. Titlul va fi formatat pentru a intra automat în cuprins. Treceți la pagină nouă.

6. Paginile următoare vor conține graficele/tabelele și interpretările corespunzătoare cerințelor de la punctele 4-11 (fiecare cerință va începe pe o pagină nouă și va avea un titlu format astfel încât să fie inserat automat în cuprins).
7. Includeți următorul antet începând cu pagina 2: *Recapitulare statistica descriptivă*.
8. Inserați începând cu pagina 2 numărul paginii.

CERINȚE SUPLIMENTARE

12. Care este procentul de valori ale **Calciului total** în eșantionul studiat pe care ne așteptăm să-l regăsim în intervalul: $[medie - 1 \cdot DS, medie + 1 \cdot DS]$? Argumentați răspunsul în *Rezultate-LP09-PrenumeNUME*. Calculați acest procent folosind funcția NORM.DIST.
13. Separat pentru bărbați, respectiv femei realizați:
 - a. reprezentarea grafică pentru variabila **Colesterol**;
 - b. calculul indicatorilor de statistică descriptivă pentru **Colesterol**, **HDL** și **TG** (folosind Data Analysis) pentru fiecare sex în parte;
 - c. calcularea și aprecierea nivelului de omogenitate pentru **Colesterol**, **HDL** și **TG** pentru fiecare sex în parte;
 - d. interpretarea asimetriei, boltirii și a distribuției teoretice normale (utilizând indicatorii statistici media aritmetică, mediana și modulul) pentru **Colesterol**, **HDL** și **TG** pentru fiecare sex în parte;
 - e. calculul intervalelor de încredere corespunzătoare mediei pentru **Colesterol**, **HDL** și **TG** pentru fiecare sex în parte;
 - f. În fișierul de raportare interpretarea intervalele de încredere obținute.