

MICROSOFT EXCEL: CORRELAȚII & REGRESII

Obiective:

- Corelația:
 - Coeficientul de Corelație Pearson (Funcția CORREL, Matricea de corelație, Interpretarea rezultatelor)
 - Grafic de tip nor de puncte
- Analiza de regresie. Interpretarea rezultatelor:
 - Regresia liniară simplă
 - Regresia liniară multiplă

Problemă

Menținerea controlului glicemiei este esențială pentru a evita consecințele pe termen lung ale diabetului la copii (de exemplu probleme neurologice, de vedere și ale funcțiilor excretorii). Ceea ce este mai puțin clar este dacă menținerea controlului glicemiei afectează creșterea și dezvoltarea pe durata copilăriei. În acest scop un grup de băieți adolescenți, cu vârste cuprinse între 9 și 15 ani au fost investigați periodic (de 3 ori). Gradul de control al diabetului a fost estimat prin măsurarea Hemoglobinei A1c (abreviată ca și HbA1c = o formă de hemoglobină folosită în primul rând pentru a identifica concentrația medie a glucozei în plasmă pe perioade mai mari de timp (valori normale: 4%–7%). Valoarea HbA1c crește la pacienți cu diabet și de asemenea la pacienți cu blocaj renal. Cu cât valoarea HbA1c este mai mare cu atât controlul diabetic este mai slab. Vârsta, înălțimea și greutatea fiecărui copil au fost notate la fiecare examinare. Datele au fost colectate și se găsesc în fișierul AP4.xls.

Întrebarea principală este dacă există o relație a controlului glicemic asupra creșterii?

Cerințe

1. Copiați fișierul **PA4.xls** în directorul cu numele dvs.
2. Calculați în **coloana eAG** nivelul mediu de glucoză estimat, exprimat în mg/dl, folosind următoarea formulă [Nathan DM, Kuenen J, Borg R, Zheng H, Schoenfeld D, Heine RJ. Translating the A1C assay into estimated average glucose values. Diabetes Care 2008;31(8):1473-8]:

$$eAG(\text{mg/dl}) = 28.7 \times \text{HbA1c} - 46.7$$

3. Calculați în **coloana BMI** indicele de masă corporală, exprimat în kg/m², folosind următoarea formulă:

$$\text{BMI} = \text{greutatea (kg)} / \text{înălțimea (m}^2\text{)}$$

4. Calculați în coloana **m-HbA1c** media pentru HbA1c pentru fiecare pacient.

$$m\text{-HbA1c} = ((\text{HbA1c la prima examinare}) + (\text{HbA1c la a 2-a examinare}) + (\text{HbA1c la a 3-a examinare}))/3$$

5. Calculați mediana pentru datele din coloana **m-HbA1c** (folosiți funcția predefinită MEDIAN și afișați rezultatul în celula W80).
6. Afișați pentru fiecare pacient dacă statusul controlului a fost bun sau nu (folosiți funcția IF), după următorul criteriu:

Statusul controlului este **da** dacă valoarea m-HbA1c pentru acel pacient este mai mică decât mediana pentru toți cei investigați;

altfel este **nu**.

7. Calculați măsura creșterii pentru fiecare pacient folosind formula:

$$\text{Măsura creșterii} = (\text{greutatea la a 3-a examinare} - \text{greutatea la prima examinare}) / (\text{vârsta la a 3-a examinare} - \text{vârsta la prima examinare})$$

8. Creați o foaie nouă numită Corelație. Copiați următoarele variabile în această foaie (folosiți opțiunea Paste Special – Values and number formats): m-HbA1c, Control bun, Măsura creșterii.
9. Presupunând o distribuție normală a datelor experimentale, calculați coeficientul de corelație (funcția predefinită CORREL) între m-HbA1c și Măsura creșterii:
 - a. Pentru toți pacienții
 - b. Pentru pacienții cu Control bun
 - c. Pentru pacienții fără Control bunInterpretați rezultatele.
10. Creați o foaie nouă numită Regresie Simplă. Copiați în această foaie datele corespunzătoare celei de a treia examinări. Presupunem că toate datele experimentale au o distribuție normală. Pe baza acestor date răspundeți la următoarele întrebări:
 - a. Este înălțimea (variabilă dependentă) liniar dependentă de vârstă (variabilă independentă)? Interpretați următorii parametri (în aceeași foaie lângă valorile parametrilor): R; R²; F.
 - b. Reprezentați grafic relația între înălțime (OY) și vârstă (OX) (folosiți graficul de tip SCATTER). Includeți în reprezentarea grafică R pătrat și de asemenea ecuația de regresie.
11. Inserați o foaie nouă numită Regresie Multiplă. Copiați în această foaie datele corespunzătoare celei de a treia examinări. Presupunem că toate datele experimentale au o distribuție normală. Pe baza acestor date răspundeți la următoarele întrebări:
 - a. Există vreo relație liniară între HbA1c ca și variabilă dependentă și vârstă, înălțime și greutate ca și variabile independente?
 - b. Interpretați următorii parametri în aceeași foaie lângă valorile parametrilor): R; R²; F.
 - c. Este vreuna din aceste variabile legată liniar de HbA1c?