

Kolokwium ze Statystyki matematycznej

Inżynieria i analiza danych GL03

12 czerwca 2023, godz. 9:45

ZADANIE 1.

W zbiorze `gleba.csv` załączonym do zadania znajdują się dane dotyczące stężeń metali ciężkich w wierzchniej warstwie gleby, zebrane na równinie zalewowej rzeki Meuse (Holandia). Zbiór zawiera następujące zmienne:

`X` – numer próbki, `cadium` – stężenie kadmu w wierzchniej warstwie gleby, `copper` – stężenie miedzi, `lead` – stężenie ołowiu, `zinc` – stężenie cynku, `elev` – względne położenie miejsca pobrania próbki nad korytem rzeki, `dist` – odległość miejsca pobrania próbki od brzegu rzeki znormalizowana do przedziału $[0, 1]$.

UWAGA. Do wczytania danych użyj funkcji `read.table()`. Nie zapomnij o argumentach `sep` i `header`.

Na podstawie obserwacji zawartych w zbiorze `gleba` zbadaj, czy stężenie miedzi w badanej glebie jest większe niż stężenie cynku. W tym celu wykonaj poniższe polecenia:

- (A) Narysuj wykres pudełkowy stężeń miedzi i cynku w badanej glebie. (2 pkt)
- (B) Na poziomie istotności 0.04 zweryfikuj hipotezę, że stężenie miedzi w badanej glebie jest większe niż stężenie cynku. (8 pkt)
- (C) Wyznacz estymator punktowy współczynnika korelacji liniowej stężeń miedzi i cynku w badanej glebie, a następnie na poziomie istotności 0.06 zweryfikuj hipotezę, że stężenia miedzi i cynku w badanej glebie są skorelowane. (5 pkt)

ZADANIE 2.

W zbiorze `MASS::anorexia` znajdują się pomiary wagi 72 dziewcząt chorych na anoreksję przed i po terapii behawioralnej (CBT) lub rodzinnej (FT) albo bez terapii (`Cont` – grupa kontrolna). Na podstawie tych danych zbadaj skuteczność terapii behawioralnej. W tym celu wykonaj poniższe polecenia:

- (A) Zbuduj przedział ufności dla różnicy średniej wagi przed i po terapii behawioralnej. (3 pkt)
- (B) Na podstawie punktu (A) postaw i zweryfikuj hipotezę dotyczącą średniej wagi przed i po terapii behawioralnej. (5 pkt)

ZADANIE 3.

Krzyżując czerwone i białe róże otrzymuje się róże czerwone, białe i różowe. Większość genetyków twierdzi, że stosunek róż czerwonych do białych i różowych, powstałych w wyniku krzyżowania, wynosi 3:2:2. Aby sprawdzić to twierdzenie pobrano 80-elementową próbkę róż powstałych z krzyżowania, w której było 35 róż czerwonych, 31 białych i 14 różowych. Na podstawie tej próby, za pomocą odpowiedniego testu, zbadaj, czy genetycy mają rację. (8 pkt)

W rozwiązaniu powyższych zadań skorzystaj z gotowych funkcji programu R.

Na rozwiązanie zadań masz 80 minut. Rozwiązania (poprawnie napisany skrypt programu R nie jest rozwiązaniem zadania) wszystkich zadań zamieść w jednym pliku Rmd. **Plikowi nadaj nazwę, którą jest twoje nazwisko bez znaków diakrytycznych.** Tak nazwany plik prześlij jako odpowiedź do zadania na Teamsach i pamiętaj: *Kto pyta nie błędzi.*