

Kolokwium ze Statystyki matematycznej

Inżynieria i analiza danych GL03

18 czerwca 2023, godz. 11:30

UWAGA. Ten plik zawiera dwie strony. Proszę zapoznać się z całością tekstu.

ZADANIE 1. Zbiór danych *car::Prestige* zawiera 6 zmiennych między innymi *education* - czas edukacji (w latach) i *income* - średni dochód w dolarach w 1971 r. Zakładając, że czas edukacji i średni dochód w 1971 r. mają rozkład normalny wykonaj poniższe polecenia i odpowiedz na zadane pytania.

(A) Na podstawie danych zawartych w zbiorze *car::Prestige* wyznacz wartość estymatora punktowego współczynnika korelacji liniowej czasu edukacji i średniego dochodu w 1971 r. i dokonaj jego interpretacji.

(4 pkt)

(B) Na poziomie ufności 0.97 zbuduj przedział ufności dla współczynnika korelacji liniowej czasu edukacji i średniego dochodu w 1971 r. Jakie wnioski można sformułować na podstawie otrzymanego przedziału?

(4 pkt)

(C) Na poziomie istotności 0.03, zweryfikuj hipotezę sformułowaną w punkcie (B).

(4 pkt)

W rozwiązaniu zadania skorzystaj z gotowych funkcji programu R.

ZADANIE 2. W przedszkolu „Wesołe skrzaty” przeprowadzono ankietę wśród 120 dzieci, pytając o ulubione zajęcia spośród czterech propozycji: malowanie, śpiew, taniec, czytanie bajek. Dyrekcja przedszkola zakłada, że dzieci wybierają te zajęcia z następującymi proporcjami: malowanie wybiera 30% przedszkolaków, śpiew - 25%, taniec - 25%, a czytanie bajek tylko 20%.

W ankiecie uzyskano następujące liczby wyborów: malowanie wybrało 45 dzieci, śpiew - 25, taniec - 30, a czytanie bajek - 20.

Na poziomie istotności 0.05 zweryfikuj hipotezę, że preferencje dzieci są zgodne z założeniami dyrekcji przedszkola. Do przeprowadzenia testu wykorzystaj odpowiednią funkcję programu R.

(8 pkt)

ZADANIE 3. W celu sprawdzenia, poglądu, że pracownicy zmian nocnych śpią mniej niż pracownicy zmian dziennych zmierzono czas snu wśród 40 losowo wybranych pracowników zmiany nocnej i 45 - zmiany dziennej, a wyniki zestawiono w tabeli

zmiana	\bar{x} (h)	s (h)
dzienna	7.2	1.1
nocna	6.5	1.3

gdzie \bar{x} oznacza średnią z pobranej próby, a s odchylenie standardowe obliczone na podstawie obciążonej wariancji. Zakładając, że czas snu ma rozkład normalny wykonaj poniższe polecenia.

(A) Na poziomie ufności 0.94 zbuduj przedział ufności dla różnicy średniego czasu snu pracowników zmiany dziennej i nocnej. Jakie wnioski możesz sformułować na podstawie otrzymanego przedziału?

(6 pkt)

(B) Na poziomie istotności 0.06 zweryfikuj hipotezę sformułowaną w punkcie (A). Weryfikację hipotezy przeprowadź na podstawie odpowiedniego obszaru krytycznego. (4 pkt)

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. Rozwiązania zadań wpisz do pliku zachowując ich kolejność.