

TESTY ZGODNOŚCI – Chi kwadrat Pearsona c.d., Kołmogorowa-Smirnova, Shapiro-Wilka;
FUNKCJA *APPLY*

Zad. 1

Zweryfikuj hipotezę H_0 że pewna cecha ma rozkład dwumianowy ($\text{binom}(k, n, p)$) w pakiecie R) z parametrami $n=10$ i $p=0.4$ na podstawie próby: 7, 5, 4, 3, 3, 4, 2, 3, 3, 3, 6, 5, 2, 4, 1, 4, 5. Użyj testu Chi kwadrat Pearsona.

Zad. 2 (Test Kołmogorowa, test Shapiro-Wilka - dla jakich rozkładów je stosujemy? Jakich hipotez dotyczą?)

Zbadano pojemność partii kondensatorów (w jednostkach pikoFaradach) i otrzymano wyniki: 62, 57, 70, 58, 59, 67, 65, 69, 55, 57, 60, 54, 72, 66, 74.

Zweryfikuj hipotezę H_0 , że rozkład pojemności jest normalny. Parametry rozkładu oszacuj na podstawie próbek. Użyj testu Kołmogorowa oraz testu Shapiro-Wilka.

Zad. 3 (Test Kołmogorowa-Smirnova)

Przeprowadzono badania dotyczące pojemności płuc (w cm^3) młodzieży szkolnej w pewnym mieście i sąsiedniej wsi. Wyniki które otrzymano znajdują się w pliku 'pluca.txt' – zaimportuj dane do zmiennej *pluca*.

Założmy, że pojemność płuc jest zmienną losową o ciągłej dystrybucji. Zweryfikuj hipotezę H_0 , że pojemności płuc młodzieży ze wsi i miast mają taki sam rozkład (przy standardowym poziomie istotności $\alpha=0.05$, czyli poziomie ufności 0.95).

Jaką statystykę testową wykorzystuje użyty do weryfikacji hipotezy test?

Zad. 4

Niech wektor *a* przechowuje próbkę 60 liczb wygenerowanych z rozkładu Cauchy'ego z parametrami 0 i 1 (parametry: przesunięcia i skali, odpowiednio) natomiast wektor *b* – próbkę 55 liczb wygenerowanych z rozkładu normalnego z parametrami 0 i 1 (średnia i odchylenie standardowe). Którego testu możemy użyć by zweryfikować hipotezę H_0 , że próby te (pobrane niezależnie) pochodzą z populacji o jednakowych rozkładach?

Zad. 5 (Funkcja *apply(...)*)

Wygeneruj ramkę *frame_* z 5 kolumnami: A, B, C, D, E zawierającymi odpowiednio wartości:

w kolumnie A: 1, 4, 8,

w kolumnie B: 2, 1, -6,

w kolumnie C: -7, 0, -3,

w kolumnie D: 3, -2, 4

oraz w ostatniej: 5, -9, 1.

Utworzymy nowe kolumny o nazwach: 'Suma' (zwracającą sumę elementów z pięciu kolumn dla danego wiersza), 'Suma(A, C, E)' (zwracającą sumę elementów z kolumny pierwszej, trzeciej i piątej), 'Max' (zwracającą element największy spośród pięciu pierwszych kolumn w danym wierszu) oraz 'MaxAbsValue' (zwracającą maximum z wartości bezwzględnych elementów).

Zrobimy to szybciej niż zwykle – używając funkcji *apply(...)*. Komenda ta, aby wygenerować zadaną kolumnę, jako pierwszy argument przyjmuje ramkę (bądź jej część) na której ma zostać wykonana operacja (pewna funkcja) *f*, jako drugi argument przyjmuje liczbę 1 lub liczbę 2

(jedynkę - gdy chcemy aby operacja f wykonywana była na kolejnych kolumnach przy ustalonym wierszu; dwójkę – gdy chcemy aby operacja f wykonana była na wierszach przy ustalonej kolumnie); w końcu jako trzeci argument przyjmuje nazwę samej operacji f .