

Statystyka dla Inżynierów
Laboratorium 11
Testy Niezależności i Zgodności

Dane do zadań 3-6 znajdują się w pliku mieszkania.xlsx

1. Tablica poniżej opisuje wyniki z rzutów kostką

Wynik	1	2	3	4	5	6
Frekwencja	171	200	168	213	226	222

Za pomocą R

- Wyznaczyć tablicę oczekiwanych frekwencji przy hipotezie że kostka jest symetryczna.
- Wyznaczyć realizację statystyki testowej dla odpowiedniego testu zgodności.
- Wyznaczyć wartość p dla tego testu.
- Jaki jest wniosek tego testu?
- Wykonać ten test za pomocą polecenia `chisq.test`

2. Tablica poniżej opisuje poziom wykształcenia według płci.

	Wykształcenie Wyższe	Wykształcenie średnie
Kobiety	200	300
Mężczyźni	150	350

- Wyznaczyć tablicę oczekiwanych frekwencji przy hipotezie o niezależności tych cech.
 - Wyznaczyć realizację statystyki testowej dla testu niezależności Pearsona
 - Wyznaczyć wartość p dla tego testu.
 - Jaki jest wniosek tego testu?
 - Wykonać test ten za pomocą polecenia `chisq.test`
 - Wykonać test ten za pomocą dokładnego testu Fishera (`fisher.test`)
3. a) Sporządzić tabelę rozdzielną opisującą rozkład liczby pokoi w zależności od dzielnicy (rejonu).
b) Skoro mieszkania rzadko mają więcej niż 4 pokoje, stworzyć nową zmienną, która przyjmuje wartość 4 gdy liczba pokoi jest co najmniej 4, a inaczej równa się liczbie pokoi.
c) Przetestować hipotezę, iż liczba pokoi jest niezależna od dzielnicy.
4. a) Tworzyć zmienną, która wskazuje czy cena za m² jest większa niż 6000zł czy nie.
b) Przetestować hipotezę iż prawdopodobieństwo tego, że cena za m² jest większa niż 6000zł zależy od dzielnicy.
5. a) Przetestować hipotezę, iż cena za m² ma rozkład normalny.
b) Przetestować hipotezę, iż metraż mieszkań na Śródmieściu ma rozkład normalny.
W obu przypadkach, narysować estymator gęstości odpowiedniej zmiennej.
6. a) Wygenerować próbę 1000 realizacji z rozkładu wykładniczego o wartości oczekiwanej 1.
b) Przetestować hipotezę iż próba ta pochodzi i) z rozkładu normalnego o średniej $\mu = 1$ i odchylenie $\sigma = 1$, ii) z rozkładu wykładniczego z parametrem $\lambda = 1$.
c) Wygenerować próbę 1000 realizacji z rozkładu Gamma o parametrze kształtu 100 a parametrze skali 1.
d) Przetestować hipotezę iż próba ta pochodzi i) z rozkładu normalnego o średniej $\mu = 100$ i odchylenie $\sigma = 10$, ii) z rozkładu Gamma z parametrami 100 i 1.

Uwaga: Suma n zmiennych o rozkładzie $\text{Exp}(\lambda)$ ma rozkład $\text{Gamma}(n, \lambda)$. Zadanie to ilustruje centralne twierdzenie graniczne. Parametry rozkładów z hipotez zerowych odpowiadają średniej i odchyleniu standardowemu rozkładów, z którego wygenerowano realizacje.