**Statystyka dla Inżynierów**

**Laboratorium 11**

**Testy Niezależności i Zgodności**

**Dane do zadań 3-5 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

1. Wyznaczyć tablicę oczekiwanych frekwencji przy hipotezie że kostka jest symetryczna.
2. Wyznaczyć realizację statystki testowej dla odpowiedniego testu zgodności.
3. Wyznaczyć wartość *p* dla tego testu.
4. Jaki jest wniosek tego testu?
5. Wykonać ten test za pomocą polecenia chisq.test
6. Tablica poniżej opisuje poziom wykształcenia według płci.

|  |  |  |
| --- | --- | --- |
|  | Wykształcenie Wyższe | Wykształcenie średnie |
| Kobiety | 200 | 300 |
| Mężczyźni | 150 | 350 |

1. Wyznaczyć tablicę oczekiwanych frekwencji przy hipotezie o niezależności tych cech.
2. Wyznaczyć realizację statystki testowej dla testu niezależności Pearsona
3. Wyznaczyć wartość *p* dla tego testu.
4. Jaki jest wniosek tego testu?
5. Wykonać test ten za pomocą polecenia chisq.test
6. Wykonać test ten za pomocą dokładnego testu Fishera (fisher.test)

4. Sporządzić tabelę rozdzielczą 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.

5. 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.

6. Podano poniżej liczbę chłopców w 700 polskich rodzin z czwórką dzieci.

i) Przetestować hipotezę że liczba cjhlopców w rodzinie ma rozkład dwumianowy Bin(4; 0*,* 5) (hipoteza prosta)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Wynik | 0 | 1 | 2 | 3 | 4 |
| Frekwencja | 20 | 30 | 150 | 250 | 250 |

7. Tabela ilustruje poziom wysiłku fizycznego w tygodniu według płci:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Mało aktywni (<3h tygodniowo) | Średnio aktywni (3-5h tygodniowo) | Bardzo aktywni (>5h tygodniowo) |
| Kobiety | 120 | 180 | 150 |
| Mężczyźni | 150 | 140 | 160 |

1. Wyznaczyć tablicę oczekiwanych frekwencji przy hipotezie o niezależności tych zachowań.
2. Wyznaczyć realizację statystki testowej dla testu niezależności Pearsona
3. Wyznaczyć wartość *p* dla tego testu.
4. Jaki jest wniosek tego testu?
5. Wykonać test ten za pomocą polecenia chisq.test
6. Wykonać test ten za pomocą dokładnego testu Fishera (fisher.test)

6. 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.