

Sprawozdanie 3

Lista 10

Zadanie 1

W tabeli 1 zawarte są wyniki (w skali pozytywny, negatywny) z pierwszego i drugiego kolokwium w pewnej grupie studentów. Przyjmując, że poziom trudności zadań na pierwszym i drugim kolokwium był taki sam, na podstawie tych danych, zweryfikować hipotezę, na poziomie istotności 0.05, że studenci byli tak samo przygotowani do obu kolokwiów.

Tablica 1: Dane do zadania 1.

Wynik z kolokwium 2	Wynik z kolokwium 1		Suma
	Negatywny	Pozytywny	
Negatywny	32	44	76
Pozytywny	22	38	60
Suma	54	82	136

Zadanie 2

W tabeli 2 zawarte są dane dotyczące reakcji po godzinie od przyjęcia dwóch różnych leków przeciwbólowych (powiedzmy A i B) stosowanych w migrenie, zaaplikowanych grupie pacjentów w dwóch różnych atakach bólowych. Na podstawie tych danych, zweryfikować hipotezę, że leki te są jednakowo skuteczne korzystając z testu:

1. McNemary'ego z poprawką na ciągłość,
2. dokładnego (opisanego w sekcji 2.1.3 wykładu 9. do wydruku)

W drugim przypadku, najpierw napisać deklarację funkcji, której wartością będzie wartość poziomu krytycznego (p wartość) w tym warunkowym teście dokładnym.

Tablica 2: Dane do zadania 2.

Reakcja na lek B	Reakcja na lek A		Suma
	Negatywna	Pozytywna	
Negatywna	1	5	6
Pozytywna	2	4	6
Suma	3	9	12

Zadanie 3

Przeprowadzić symulacje, w celu porównania mocy testu Z (opisanego w sekcji 2.1.1) i testu Z_0 (opisanego w sekcji 2.1.2). Wyniki przedstawić w tabeli lub/i na wykresach i napisać odpowiednie wnioski.

Lista 11

Zadanie 1

W tabeli 1 zawarte są wyniki (w skali 2, 3, +3, 4, +4, 5) z pierwszego i drugiego kolokwium w pewnej grupie studentów. Korzystając z odpowiedniego testu, na poziomie istotności $\alpha = 0.05$, zweryfikować hipotezę, że dane w tabeli 1 podlegają modelom:

1. symetrii,
2. quasi-symetrii
3. quasi-niezależności.

Zwrócić uwagę na problem z zastosowaniem do analizowanych danych testu Bowkera

0.0.1. Testowanie symetrii za pomocą testu ilorazu wiarygodności

```

symmetry <- glm(count ~ Symm(kol1, kol2), data=dane33,
                 family = poisson)
x = symmetry$deviance #wartość statystyki  $G^2$ 
r=15 #Liczba stopni swobody
1-pchisq(x,r) #p-wartość
## [1] 0.1004656

```

Do przetestowania symetrii dla danych w Tabeli 1. wykorzystaliśmy test ilorazu wiarygodności (IW), aby to zrobić musieliśmy przekształcić dane do postaci ramki danych. Następnie za pomocą funkcji **glm** z biblioteki **gnm** przeprowadziliśmy test symetrii. P-wartość owego testu wyznaczyliśmy za pomocą wzoru : **1-pchisq(x,r)**, gdzie **pchisq** jest dystrybuantą rozkładu χ^2 , **x** to wartość statystyki G^2 z testu, a **r** to ilość stopni swobody. Na poziomie istotności 0.05 nie mamy podstaw do odrzucenia hipotezy o symetrii. Wartość poziomu krytycznego wynosi 0.1004656.

0.0.2. Testowanie quasi-symetrii

Hipotezę zerową, że dane podlegają modelowi quasi-symetrii również zweryfikujemy korzystając z funkcji `glm`.

```
quasi.symm <- glm(count ~ kol1+kol2 + Symm(kol1,kol2),  
                  data=dane33, family =poisson)  
x = quasi.symm$deviance #wartość statystyki  $G^2$   
r=10 #Liczba stopni swobody  
1-pchisq(x,r) #p-wartość  
## [1] 0.9589187
```

Korzystając z testu ilorazu wiarygodności (IW), na poziomie istotności 0.05, nie ma podstaw do odrzucenia hipotezy o quasi-symetrii. Wartość poziomu krytycznego w teście wynosi 0.9589187.

0.0.3. Testowanie quasi-niezależności

Hipotezę zerową, że dane podlegają modelowi quasi niezależności również zweryfikujemy korzystając z funkcji `glm`.

```
quasi.indep <- glm(count ~ kol1 + kol2 + Diag(kol1, kol2),  
                   data=dane33, family = poisson)  
x=quasi.indep$deviance #wartość statystyki  $G^2$   
r = 19 #Liczba stopni swobody  
1-pchisq(x,r) #p-wartość  
## [1] 0.00962481
```

Korzystając, z testu ilorazu wiarygodności (IW), na poziomie istotności 0.05, hipotezę o quasi-niezależności należy odrzucić. Wartość poziomu krytycznego w tym teście wynosi 0.00962481.

Zadanie 2

W tabeli 1 zawarte są wyniki (w skali 2, 3, +3, 4, +4, 5) z pierwszego i drugiego kolokwium w pewnej grupie studentów. Przyjmując, że poziom trudności zadań na pierwszym i drugim kolokwium był taki sam, na podstawie tych danych, zweryfikować hipotezę, na poziomie istotności 0.05, że studenci byli tak samo przygotowani do obu kolokwiów.

Tablica 3: Dane do zadania 1 i 2.

Wyniki z kolokwium 2	Wyniki z kolokwium 1						Suma
	2	3	+3	4	+4	5	
2	5	2	1	0	0	0	8
3	6	3	2	2	0	0	13
+3	1	4	5	5	2	2	19
4	0	10	15	18	5	2	50
+4	1	2	5	3	2	2	15
5	0	1	3	4	3	2	13
Suma	13	22	31	32	12	8	118

Lista 12,13 i 14

Wszystkie poniższe zadania należy wykonać w oparciu o dane w pliku Ankieta.csv, które zawierają, a wyniki ankietowania 40 losowo wybranych studentów PWr. Ankieta zawierała trzy pytania, które dotyczyły jakości snu (odpowiedź 1 oznaczała, że student sypia dobrze, 0, że źle), czy regularnie biega (1 – tak, 0 – nie) oraz czy posiada psa (1 – tak, 0 – nie).

Zadanie 1

W przypadku powyższych danych, podać interpretację następujących modeli logliniowych:

■ [1 3],

$$l_{ij} = \lambda + \lambda_i^{(1)} + \lambda_j^{(3)}, \forall i \in \{1, \dots, R\} \text{ i } j \in \{1, \dots, C\},$$

Zmienne W_1 i W_3 mają dowolne rozkłady oraz zmienne te są niezależne.

■ [13],

$$l_{ij} = \lambda + \lambda_i^{(1)} + \lambda_j^{(3)} + \lambda_{ij}^{(13)}, \forall i \in \{1, \dots, R\} \text{ i } j \in \{1, \dots, C\},$$

Zmienne W_1 i W_3 mają dowolne rozkłady oraz zmienne te nie są niezależne.

■ [1 2 3],

$$l_{ijk} = \lambda + \lambda_i^{(1)} + \lambda_j^{(2)} + \lambda_k^{(3)}, \forall i \in \{1, \dots, R\} \text{ i } j \in \{1, \dots, C\} \text{ i } k \in \{1, \dots, L\},$$

Zmienne W_1 i W_2 i W_3 są wzajemnie niezależne.

■ [12 3],

$$l_{ijk} = \lambda + \lambda_i^{(1)} + \lambda_j^{(2)} + \lambda_k^{(3)} + \lambda_{ij}^{(12)},$$

$$\forall i \in \{1, \dots, R\} \text{ i } j \in \{1, \dots, C\} \text{ i } k \in \{1, \dots, L\},$$

Zmienna W_3 jest niezależna od zmiennej W_1 i W_2 , ale zmienne W_1 i W_2 nie są niezależne.

■ [12 13],

$$l_{ijk} = \lambda + \lambda_i^{(1)} + \lambda_j^{(2)} + \lambda_k^{(3)} + \lambda_{ij}^{(12)} + \lambda_{ik}^{(13)},$$

$$\forall i \in \{1, \dots, R\} \text{ i } j \in \{1, \dots, C\} \text{ i } k \in \{1, \dots, L\},$$

Przy ustalonej wartości zmiennej W_1 , zmienne W_2 i W_3 są niezależne. Mówimy wówczas, że zmienne W_2 i W_3 są warunkowo niezależne.

■ [1 23].

$$l_{ijk} = \lambda + \lambda_i^{(1)} + \lambda_j^{(2)} + \lambda_k^{(3)} + \lambda_{jk}^{(23)},$$

$$\forall i \in \{1, \dots, R\} \text{ i } j \in \{1, \dots, C\} \text{ i } k \in \{1, \dots, L\},$$

Zmienna W_1 jest niezależna od zmiennej W_2 i W_3 , ale zmienne W_2 i W_3 nie są niezależne.

Zadanie 2

Przyjmując model log-liniowy [12 3], na podstawie danych Ankieta.csv, oszacować prawdopodobieństwo:

1. dobrej jakości snu studenta, który regularnie biega,
2. tego, że student biega regularnie, gdy posiada psa.

Jakie byłyby oszacowania powyższych prawdopodobieństw przy założeniu modelu [12 23]?

Zadanie 3

Na podstawie danych Reakcja3.csv zweryfikować następujące hipotezy:

1. zmienne losowe Sen, Bieganie i Pies są wzajemnie niezależne,
2. zmienna losowa Pies jest niezależna od pary zmiennych Sen i Bieganie,
3. zmienna losowa Sen jest niezależna od zmiennej Pies, przy ustalonej zmiennej Bieganie.

Zadanie 4

Na podstawie danych Ankieta.csv dokonać wyboru modelu w oparciu o:

1. testy,
2. kryterium AIC,
3. kryterium BIC.

W przypadku, gdy wybrane modele w punktach 1–3 są różne, dokonać ich porównania.