
ANALIZA DANYCH ANKIETOWYCH, SEMESTR LETNI 2024/2025

Zadania do sprawozdania 3

Część I oraz II

zadanie 1. Napisz funkcję, która zwraca p-wartość w omówionym na wykładzie warunkowym teście symetrii w przypadku tabeli 2×2 .

zadanie 2. W tabeli 1 umieszczono dane dotyczące reakcji na lek po godzinie od jego przyjęcia dla dwóch różnych leków przeciwbólowych stosowanych w migrenie. Leki zostały zaaplikowane grupie pacjentów w dwóch różnych atakach bólowych. Na podstawie danych zweryfikuj hipotezę, że leki te są jednakowo skuteczne korzystając z testu

- McNemara z poprawką na ciągłość,
- warunkowego (korzystając z funkcji zadeklarowanej w zadaniu 1.).

Reakcja na lek A	Reakcja na lek B	
	Negatywna	Pozytywna
Negatywna	1	5
Pozytywna	2	4

Table 1: Dane do zadania 2.

zadanie 3. Przeprowadź symulacje w celu porównania mocy testu Z i testu Z_0 przedstawionych na wykładzie. Rozważ różne długości prób.

zadanie 4. Dla danych dołączonych do pierwszej listy zadań, na podstawie zmiennych **CZY_ZADW** oraz **CZY_ZADW_2**, zweryfikuj hipotezę, że zadowolenie ze szkoleń w pierwszym badanym okresie i w drugim badanym okresie pierwszego badania odpowiada modelowi symetrii. Czy na podstawie uzyskanych wniosków możemy wnioskować, że poziom zadowolenia ze szkoleń nie uległ zmianie? Przyjmij poziom istotności 0.05.

zadanie 5. W firmie, o której mowa w zadaniu 1 z listy 1, wdrożono pewne działania w celu poprawy komfortu pracy. Następnie badaną grupę respondentów ponownie poproszono o odpowiedź na pytanie dotyczące oceny podejścia firmy do umożliwiania wdrażania wiedzy zdobytej na szkoleniach. W Tabeli 2 przedstawiono tablicę dwudzielczą uwzględniającą odpowiedzi na pytanie w obu tych okresach. Na podstawie danych zweryfikuj hipotezę, że odpowiedzi w pierwszym badanym okresie i w drugim okresie odpowiadają modelowi symetrii. Na podstawie wyników uzyskanych przy weryfikacji hipotezy dotyczącej symetrii, sformułuj wniosek dotyczący hipotezy, że ocena podejścia firmy nie uległa zmianie.

Pytanie 1	Pytanie 2				
	-2	-1	0	1	2
-2	10	2	1	1	0
-1	0	15	1	1	0
0	1	1	32	6	0
1	0	0	1	96	3
2	1	1	0	1	26

Table 2: Dane do zadania 5.

Część III

zadanie 6. W pewnym badaniu porównywano skuteczność dwóch metod leczenia: Leczenie A to nowa procedura, a Leczenie B to stara procedura. Przeanalizuj dane przedstawione w Tabeli 3 (wyniki dla całej grupy pacjentów) oraz w Tabelach 4 i 5 (wyniki w podgrupach ze względu na dodatkową zmienną) i odpowiedz na pytanie, czy dla danych występuje paradoks Simpsona.

Metoda	Wynik leczenia	
	Poprawa	Brak
Leczenie A	117	104
Leczenie B	177	44

Table 3: Dane dla całej grupy

Metoda	Reakcja	
	Poprawa	Brak
Leczenie A	17	101
Leczenie B	2	36

Table 4: Dane dla pacjentów z chorobami współistniejącymi

Metoda	Reakcja	
	Poprawa	Brak
Leczenie A	100	3
Leczenie B	175	8

Table 5: Dane dla pacjentów bez chorób współistniejących

zadanie 7. Dla danych z listy 1, przyjmując za zmienną 1 zmienną **CZY_KIER**, za zmienną 2 – zmienną **PYT_2** i za zmienną 3 – zmienną **STAŻ**, podaj interpretacje następujących modeli log-liniowych: [1 3], [13], [1 2 3], [12 3], [12 13] oraz [1 23].

Część IV oraz V

zadanie 8. Przyjmując model log-liniowy [123] dla zmiennych opisanych w zadaniu 7 oszacuj prawdopodobieństwa:

- że osoba pracująca na stanowisku kierowniczym jest zdecydowanie zadowolona ze szkoleń;
- że osoba o stażu pracy krótszym niż rok pracuje na stanowisku kierowniczym;
- że osoba o stażu pracy powyżej trzech lat nie pracuje na stanowisku kierowniczym.

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

zadanie 9. Dla danych wskazanych w zadaniu 7 zweryfikuj następujące hipotezy:

- zmienne losowe **CZY_KIER**, **PYT_2** i **STAŻ** są wzajemnie niezależne;
- zmienna losowa **PYT_2** jest niezależna od pary zmiennych **CZY_KIER** i **STAŻ**;
- zmienna losowa **PYT_2** jest niezależna od zmiennej **CZY_KIER**, przy ustalonej wartości zmiennej **STAŻ**.

Zadania dodatkowe

zadanie 1* W przypadku zadania 5 występuje problem z zastosowaniem testu Bowkera ze względu na występowanie zer na określonych miejscach w tabeli z danymi. Zastosuj w tym przypadku dokładny test symetrii i opisz w jaki sposób wyznaczana jest wartość poziomu krytycznego w tym teście.

zadanie 2* Na podstawie danych z listy 1 dokonaj wyboru modelu rozważając uwzględnienie zmiennych **PYT_1**, **PYT_2** i **PŁEĆ** w oparciu o:

- testy,
- kryterium AIC,
- kryterium BIC.