

## Statystyczna analiza danych: egzamin - 11 czerwca 2013

1. Naukowiec dokarmiający zimą ptaki chciał rozstrzygnąć, czy mają one preferencje dotyczące wyglądu ziaren słonecznika, którymi były żywione. W karmniku znajdowały się zarówno zupełnie czarne, jak i pasiaste ziarna, początkowo w równych ilościach. Uważna obserwacja przylatujących ptaków wykazała, że spośród 80 osobników, które skonsumowały ziarna słonecznika, 60 ptaków wybierało czarne ziarna. Za każdym razem dodawano ziarenka, tak aby zrównoważyć skład. Czy naukowiec może przyjąć za prawdziwą hipotezę, że ptaki preferują pewien rodzaj nasion?
2. Przeprowadzono badania sprawdzające ciśnienie tętnicze (skurczowe) w dwóch różnych grupach stosujących diety wegetariańskie oraz w grupie osób odżywiających się standardowo. Badane osoby zostały wylosowane z jednorodnej wiekowo populacji stosującej daną dietę. Pomiar ciśnienia został przeprowadzony w zbliżonych warunkach. W tabeli poniżej znajdują się wyniki ciśnienia dla kobiet i mężczyzn z różnych grup (w milimetrach słupa rtęci mmHg). Przetestuj odpowiednie hipotezy o wpływie diety na ciśnienie oraz jej powiązań z płcią.

dieta	płeć	pomiary ciśnienia
Ścisłe wegetariańska	Mężczyźni	110, 113
Ścisłe wegetariańska	Kobiety	103, 100
Laktowegetariańska	Mężczyźni	115, 119
Laktowegetariańska	Kobiety	105, 109
Standardowa	Mężczyźni	128, 132
Standardowa	Kobiety	119, 127

3. Zespół naukowców chciał zbadać względną przydatność penicyliny i spektynomycyny w leczeniu rzeżączki. Zastosowano trzy terapie: (1) zazwyczaj stosowaną dawkę penicyliny, (2) wysoką dawkę spektynomycyny; oraz (3) niską dawkę spektynomycyny. Po zakończeniu leczenia, zbadano stan pacjentów stosując dwa standardowe testy na rzeżączkę (nazwijmy je T1 i T2). Leczona grupa obejmowała 200 pacjentów, którym aplikowano penicylinę oraz po 100 pacjentów leczonych spek-

tynomycyną (odpowiednio w niskiej i wysokiej dawce). Poniżej wyniki testów dla każdej grupy pacjentów:

- **penicylina**

T1 i T2 wynik pozytywny: 40 pacjentów;

T1 - wynik negatywny, T2 - pozytywny: 30 pacjentów

T1 i T2 - wynik negatywny: 130 pacjentów;

- **niska spektomycyna**

T1 i T2 wynik pozytywny: 10 pacjentów;

T1 - wynik negatywny, T2 - pozytywny: 20 pacjentów

T1 i T2 - wynik negatywny: 70 pacjentów;

- **wysoka spektomycyna**

T1 i T2 wynik pozytywny: 15 pacjentów;

T1 - wynik negatywny, T2 - pozytywny: 40 pacjentów

T1 i T2 - wynik negatywny: 45 pacjentów;

Zbadaj czy istnieje związek między zastosowanym lekiem, a efektem leczenia. Następnie powtórz analizy uznając, że są tylko dwie kategorie pacjentów: wyleczony, jeśli obydwa testy dały wynik negatywny, oraz nie wyleczony w przeciwnym przypadku.

4. Rozważmy dwie niezależne zmienne losowe  $X$  i  $E$  o rozkładzie normalnym o średniej zero i wariancji 1. Zdefiniujmy zmienną  $Y$  jako:  $Y = X + \frac{1}{2}E$ . Policz korelację pomiędzy  $X$  i  $Y$ .
5. Przeprowadzono niewielki eksperyment, żeby zbadać wpływ poziomu ekspresji dwóch genów (nazwijmy zmienne losowe  $X_1$  i  $X_2$ ) na poziom ekspresji trzeciego genu (zmienna  $Y$ ). W tabeli znajdują się wyniki pomiarów ekspresji dla czterech obserwacji, przy czym ekspresja  $X_1$  i  $X_2$  nie jest mierzona dokładnie, tylko w każdym z eksperymentów gen jest albo nokautowany (wartość  $-1$ ) albo jego ekspresja jest sztucznie podwyższona (wartość  $1$ ). Ekspresja  $Y$  jest mierzona dokładnie z użyciem techniki PCR.

lp.	$X_1$	$X_2$	$Y$
1	$-1$	$-1$	2.4
2	$-1$	1	0.4
3	1	$-1$	3.2
4	1	1	1.6

Zapisz model liniowy, który opisuje związek pomiędzy badanymi genami, wraz z niezbędnymi założeniami.

6. Naukowiec zamierza wyjaśnić różnicę w poziomie ekspresji genu A w trzech typach komórek za pomocą modelu regresji linowej. Dla każ-

dego typu komórek dysponuje on dwoma pomiarami. Zaprojektuj odpowiedni model liniowy.

7. Rozważny łańcuch Markowa o następującej macierzy przejść:

$$\begin{matrix} 1 - a & a \\ b & 1 - b \end{matrix}$$

Znajdź rozkład stacjonarny tego łańcucha. Dla jakich parametrów  $a$  i  $b$  ten łańcuch jest odwracalny, a dla jakich jest cykliczny ?

8. W czasie imprezy na kampusie studenckim, podano dwie pizze (serową i pepperoni) do stolika, przy którym siedziało 5-ciu studentów. Każda pizza pocięta została na 6 kawałków. Każdy student sięgnął po jeden kawałek i zjedzono 4 kawałki pepperoni i 1 kawałek pizzy serowej. Dostawca pizzy uznał, że studenci wolą pizzę pepperoni. Użyj testu dokładnego Fishera, żeby sprawdzić, czy hipoteza o preferencjach studentów jest statystycznie uzasadniona.