

Zadania i problemy do wykładu *Testowanie hipotez statystycznych*

(ZESTAW NR 7)

Zadanie 0. W badaniu opinii publicznej często pojawiają się więcej niż dwie opinie. Dla przykład, założmy, że przeprowadzono ankietę zarejestrowanych wyborców, aby zobaczyć, który kandydat prawdopodobnie zostanie wybrany w nadchodzących wyborach. Dla uproszczenia zakładamy że są dwaj kandydaci, A i B. Potencjalny wyborca może wybrać jednogo z nich lub może być niezdecydowany.

1. Jeśli ankietowanych jest 100 osób, a wyniki wynoszą 35 dla A, 40 dla B i 25 niezdecydowanych, czy jest różnica poparciem dla kandydata A oraz B jest znacząca?
2. Założmy na przykład, że chcemy wiedzieć, czy rozkład poparcia dla kandydatów w całej populacji głosujących jest zgodny z rozkładem, w którym 35% wyborców popiera kandydata A, 35% kandydata B oraz 30% wyborców jest niezdecydowanych?

Zadanie 1. W pliku **rekombinacja.xls** znajdują się dane dotyczące liczby miejsc, w których zaszło crossin-over. Przetestuj hipotezę, że rozkład liczby miejsc, w których zachodzi rekombinacja jest rozkładem Bernoulliego z prawdopodobieństwem sukcesu 0,25 oraz liczbą prób równą 5.

Zadanie 2. W pliku **rak.xls** znajdują się dane dotyczące sposobu leczenia kobiet ze wczesnym stadium raka piersi. W przypadku tych kobiet stosowano dwa sposoby leczenia jednym z nich była radykalna mastektomia, drugim częściowa mastektomia z radioterapią.

1. Przetestuj hipotezę o równej proporcji kobiet powyżej 55 roku życia w grupie z radykalną mastektomią i grupie z częściową mastektomią.
2. Sprawdź czy jest zależność pomiędzy grupą wiekową a rodzajem leczenia.

Zadanie 3. W pliku **pasy.xls** znajdują się dane dotyczące skutków wypadków drogowych. Przetestuj zależność pomiędzy zapiętymi pasami bezpieczeństwa podczas wypadku a śmiertelnością na skutek wypadku.

Zadanie 4. W pliku **szpitale.xls** znajdują się dane dotyczące poprawności stwierdzenia przyczyn zgonu w dwóch szpitalach. Szpital A to szpital miejski natomiast B to szpital kliniczny. Stwierdzone przyczyny zgonów zostały sprawdzone i zakwalifikowane do jednej z trzech grup. Poprawnych, niepoprawnych nie wymagających przekwalifikowania oraz nieprawidłowych wymagających przekodowania. Przetestuj hipotezę o różnicach w procedurach stwierdzania przyczyn zgonu w obu szpitalach.

Zadanie 5. Dla danych, z każdej z czterech kolumn w pliku **rozklady.xls**, zaproponuj typ rozkładu, który najlepiej opisuje ich rozkład, a następnie przetestuj tak postawioną hipotezę.

Zadanie 6. W 100 elementowej próbie losowej z diploidalnej populacji otrzymano następujące liczebności genotypów:

	BB	Bb	bb
AA	0	25	0
Aa	25	0	25
aa	0	25	0

Korzystając z testu chi-kwadrat przetestuj hipotezę o niezależności genotypów w obu loci. Oszacuj współczynnik nierównowagi gametycznej D_{AB} .