

TESTY PARAMETRYCZNE DLA DWÓCH POPULACJI

ZADANIE 4.1 W losowej próbie 233 dorosłych mieszkańców Warszawy znalazło się 40 takich, które regularnie robią zakupy w sklepach sieci Żuczek. W Krakowie na 220 zapytane osoby, 31 okazało się klientami Żuczka.

(a) Czy na podstawie powyższych danych można stwierdzić, że odsetek regularnych klientów Żuczka w Warszawie jest większy niż w Krakowie? Przyjąć poziom istotności $\alpha = 0,05$.

(b) Przypuszczamy, że odsetek regularnych klientów Żuczka w Warszawie wynosi 17%, a w Krakowie – 14%. Jakie jest prawdopodobieństwo, że test z pkt. (a) potwierdzi, że odsetek regularnych klientów Żuczka jest większy w Warszawie niż w Krakowie?

(c) Przypuszczamy, że odsetek regularnych klientów Żuczka w Warszawie wynosi 17%, a w Krakowie – 14%. Ilu mieszkańców Warszawy i ilu mieszkańców Krakowa trzeba by wylosować do próby by, z prawdopodobieństwem nie mniejszym niż 0,8, jednostronny test o poziomie istotności 0,05 porównujący odsetek regularnych klientów Żuczka potwierdził, że odsetek ten jest większy w Warszawie niż w Krakowie?

ZADANIE 4.2 W celu zbadania czy dwa gatunki pszenicy średnio dają takie same plony, 16 poletek podzielono losowo na dwie równoliczne grupy i na poletkach z pierwszej grupy zasiano gatunek I, a na pozostałych - gatunek II. Wielkości uzyskanych plonów zapisano w pliku *yields*. Jakie wyciągniemy wnioski na podstawie tych danych? Przyjąć poziom istotności 0,05. Uwzględnić wyniki wcześniejszych badań, z których wiadomo, że plony badanych dwóch gatunków pszenicy mają rozkłady normalne.

ZADANIE 4.3 W każdym miesiącu roku w wybrany losowo dzień zmierzono stężenie dwutlenku węgla w domu energooszczędnym oraz w domu zbudowanym tradycyjną techniką (domy te położone są w tej samej okolicy). Otrzymane wyniki (w ppm) zapisane są w pliku *domy*: w kolumnie *domE* znajdują się wyniki uzyskanie w domu energooszczędnym, w kolumnie *domS* - wyniki uzyskanie w domu zbudowanym tradycyjną techniką.

(a) Czy dane te potwierdzają, że średnie stężenie dwutlenku węgla w domu energooszczędnym jest wyższe niż w domu zbudowanym tradycyjną techniką? Przyjąć poziom istotności 0,05.

(b) Zakładając, że stężenie dwutlenku węgla w domu energooszczędnym jest średnio o 50 ppm wyższe niż w domu zbudowanym tradycyjną techniką, wyznaczyć moc testu z pkt (a) i podać interpretację otrzymanego wyniku.

Można założyć, że łączny rozkład stężenia dwutlenku węgla w obu typach rozważanych domów jest normalny.

ZADANIE 4.4 (a) Na podstawie poniższych danych zweryfikować hipotezę, że istnieje istotna różnica pomiędzy pomiarami ciśnienia uzyskanymi za pomocą dwóch różnych przyrządów. Przyjąć poziom istotności 0,01.

	Przyrząd A	Przyrząd B
Pacjent 1	144	147
Pacjent 2	165	167
Pacjent 3	125	124
Pacjent 4	149	152
Pacjent 5	128	127
Pacjent 6	159	160

(b) Jaką średnią różnicę pomiędzy pomiarami ciśnienia na tych dwóch przyrządach jest w stanie wykryć test z pkt (a) z prawdopodobieństwem 0,8?

(c) Co najmniej ile pomiarów ciśnienia trzeba by wykonać by test z pkt (a), z prawdopodobieństwem nie mniejszym niż 0,8, wykrywał różnicę pomiędzy pomiarami ciśnienia na tych dwóch przyrządach wynoszącą 1,20 ?

Założyć, że pomiary ciśnienia na tych dwóch przyrządach mają łączny rozkład normalny.

ZADANIE 4.5 Niezależne pomiary dokonane na losowych próbach dwóch gatunków papierosów dały następujące wyniki zawartości nikotyny (w miligramach):

Gatunek A: 26.4, 22.5, 24.9, 23.7, 21.5

Gatunek B: 25.1, 29.0, 23.4, 27.6, 22.3

(a) Na poziomie istotności 0,05 zweryfikować hipotezę, że gatunek B ma wyższą zawartość nikotyny niż gatunek A.

(b) Zakładając, że gatunek B ma zawartość nikotyny średnio o 2 miligramy większą niż gatunek A, obliczyć prawdopodobieństwo, że test z pkt (a) da błędną odpowiedź.

(c) Załóżmy, że gatunek B ma zawartość nikotyny średnio o 2 miligramy większą niż gatunek A. Jak liczne próby losowe tych gatunków papierosów trzeba by pobrać, by na ich podstawie, test z pkt. (a), z prawdopodobieństwem nie mniejszym niż 0,75, dawał poprawną odpowiedź?

Przyjąć, że, w przypadku obu badanych gatunków papierosów, zawartość nikotyny ma rozkład normalny.

ZADANIE 4.6 Zbiór *nlschools*, znajdujący się w bibliotece MASS, zawiera dane dotyczące wybranych uczniów szkół holenderskich kończących ósmą klasę:

IQ - wynik testu na IQ werbalne (w pkt.),

SES - społeczno-ekonomiczny status rodziny ucznia.

Czy na podstawie powyższych danych możemy stwierdzić na poziomie istotności 0,05, że wśród uczniów kończących ósmą klasę, ci pochodzący z domów o społeczno-ekonomicznym statusie powyżej mediany, mają wyższy poziom inteligencji werbalnej niż pozostali?

ZADANIE 4.7 Do badania wybrano w sposób losowy 15 dzieci chorych na cukrzycę. Poddano je kuracji podając nowo opracowany lek. W pliku *hemoglobina* zapisano poziom hemoglobiny glikowanej (w %) u tych dzieci przed (zmienna *przed*) oraz po kuracji (zmienna *po*). Wiadomo, że poziomy te mają łączny rozkład normalny.

(a) Czy dane te potwierdzają, że nowy lek obniża poziom hemoglobiny glikowanej? Przyjąć poziom istotności 0,05.

(b) Zakładając, że nowy lek obniża poziom hemoglobiny glikowanej o średnio 1,5 %, wyznaczyć moc testu z pkt (a) i podać interpretację otrzymanego wyniku.

ZADANIE 4.8 Zbadano losowo wybranych 1000 osób w mieście i 2000 osób na wsi. W mieście odsetek zwolenników partii „Beta” wyniósł 20%, zaś na wsi znalazło się 450 zwolenników tej partii.

(a) Czy na poziomie istotności 0,05 można stwierdzić, że partia „Beta” cieszy się mniejszym poparciem w mieście niż na wsi?

(b) Przypuszczamy, że poparcie dla partii „Beta” wynosi w mieście 20% a na wsi 23%. Ilu mieszkańców miasta i ilu mieszkańców wsi trzeba by wylosować do próby aby jednostronny test porównujący proporcje poparcia z poziomem istotności 0,05 przyjmował, z prawdopodobieństwem nie mniejszym niż 0,8, hipotezę, że poparcie partii „Beta” jest większe na wsi niż w mieście?

ZADANIE 4.9 Zbiór *Cushings* z pakietu *MASS* zawiera wyniki testów przeprowadzonych na pacjentach z tzw. zespołem Cushinga. Dane zawierają 27 obserwacji dla 3 zmiennych:

Tetrahydrocorisone - tempo wydalania tetrahydrokortyzonu z moczem (w mg/24h),

Pregnanatiol - tempo wydalania pregnanetriolu z moczem (w mg/24h),

Type - zmienna grupująca oznaczająca typ zespołu (u oznacza postać nieznana).

(a) Zweryfikować hipotezę o jednakowej wariancji tempa wydalania tetrahydrokortyzonu z moczem u pacjentów z postacią zespołu zakodowaną jako *a* i *b*. Przyjąć poziom istotności $\alpha = 0,1$.

(b) Zweryfikować hipotezę o równości średniego tempa wydalania tetrahydrokortyzonu z moczem u pacjentów z typem zespołu *a* i *b*. Przyjąć poziom istotności $\alpha = 0,02$.

Przyjąć, że tempo wydalania tetrahydrokortyzonu z moczem u pacjentów z rozważanymi typami zespołu Cushinga ma wielowymiarowy rozkład normalny.