

**TESTY PARAMETRYCZNE DLA JEDNEJ I DWÓCH POPULACJI**

**ZADANIE 3.1** Narysować wykres gęstości standardowego rozkładu normalnego w zakresie od kwantyla rzędu 0,001 do 0,999. Następnie na wykresie tym nanieść różnymi kolorami wykresy gęstości rozkładu t-Studenta o liczbie stopni swobody równej 5, 10, 20, 50, 100. Zwrócić uwagę na charakter zmiany kształtu tych wykresów, gdy wzrasta liczba stopni swobody.

**ZADANIE 3.2** W kolumnie *WeightInitial* w pliku *goats.txt* zapisano wagę (w kg) losowo wybranych młodych kóz hodowanych w Australii. Wiadomo, że rozkład badanej cechy jest normalny.

(a) Na poziomie istotności 0,05 przetestować hipotezę, że średnia waga młodych kóz hodowanych w Australii przekracza 23 kg.

(b) Zakładając, że rzeczywista średnia waga młodych kóz hodowanych w Australii wynosi 24 kg, wyznaczyć prawdopodobieństwo, że przeprowadzając test na poziomie istotności 0,05 i na podstawie 40 obserwacji, błędnie uznamy, że średnia waga takich kóz nie przekracza 23 kg.

(c) Załóżmy, że rzeczywista średnia waga młodych kóz hodowanych w Australii wynosi 24 kg. Ile trzeba by zebrać pomiarów wag takich kóz, by prawdopodobieństwo, że test (przeprowadzony na poziomie istotności 0,05) wykrywał, z prawdopodobieństwem nie mniejszym niż 0,8, że średnia waga takich kóz przekracza 23 kg?

**ZADANIE 3.3** \* Napisać funkcję, która dla danej próby losowej z rozkładu normalnego  $\mathcal{N}(\mu, \sigma^2)$  z nieznanymi  $\mu$  i  $\sigma^2$ , weryfikuje hipotezę dotyczącą wariancji. Niech funkcja ta zwraca wartość statystyki testowej oraz *p-value* testu i dopuszcza rozważanie trzech hipotez alternatywnych. Następnie używając tej funkcji rozwiązać zadanie 3.4.

**ZADANIE 3.4** Korzystając z danych z zadania 3.2, sprawdzić czy

(a) można przyjąć, że wariancja wagi młodych kóz hodowanych w Australii wynosi 20 kg<sup>2</sup>;

(b) odchylenie standardowe wagi młodych kóz hodowanych w Australii przekracza 3 kg.

Przyjąć poziom istotności 0,1.

**ZADANIE 3.5** W pakiecie *faraway* znajduje się zbiór danych *uswages* zawierający kolumnę *wage* z rzeczywistymi tygodniowymi dochodami (w dolarach amerykańskich) 2000 mężczyzn wybranych do próby losowej podczas badania przeprowadzonego w USA w 1988 roku.

(a) Wczytać te dane.

(b) Na poziomie istotności 0,05 przetestować hipotezę, że średni tygodniowy dochód mężczyzn pracujących w USA w 1988 roku przekraczał 600 \$.

(c) Zakładając, że średni tygodniowy dochód mężczyzn pracujących w USA w 1988 roku wynosił 610 \$, obliczyć prawdopodobieństwo, że test z pkt. (b) da nam błędną odpowiedź.

(d) Ile musiałby wynosić średni tygodniowy dochód mężczyzn pracujących w USA w 1988 roku, aby test z pkt. (b) z prawdopodobieństwem 0,8 przyjmował hipotezę, że badany średni dochód jest większy niż 600 \$?

**ZADANIE 3.6** Pełnomocnik rządu Alfalandii d/s równego statusu kobiet i mężczyzn podejrzewa, że udział mężczyzn wśród pracowników przedszkoli jest niższy niż minimum przewidziane w ustawie, a wynoszące 35%.

(a) Czy na poziomie istotności 0,05 można uznać to stwierdzenie za uzasadnione, jeśli wśród losowo zbadanych 400 pracowników przedszkoli było 128 mężczyzn?

(b) Czy odpowiedź uzyskana w pkt. a). zmieniłaby się, gdyby pełnomocnik pobrał reprezentatywną próbkę 10 pracowników przedszkoli i 3 z nich okazałoby się mężczyznami?

**ZADANIE 3.7** Ornitolog, badający określony gatunek, pobrał próbę losową 10 dorosłych ptaków i zmierzył ich wagę, otrzymując następujące wyniki (w kg):

5, 21; 5, 15; 5, 20; 5, 48; 5, 19; 5, 25; 5, 09; 5, 17; 4, 94; 5, 11.

Można uznać, że waga ptaków badanego gatunku ma rozkład normalny.

(a) Czy na poziomie istotności 0,05 można stwierdzić, że średnia waga ptaków badanego gatunku jest mniejsza niż 5, 20 kg?

(b) Z jakim prawdopodobieństwem test, przeprowadzony w pkt. (a), przyjmie na poziomie istotności 0,05 hipotezę, że średnia waga ptaków badanego gatunku jest mniejsza niż 5, 20 kg, w sytuacji, gdy w rzeczywistości ta średnia waga wynosi 5, 15 kg?

(c) Ile by musiała wynosić średnia waga ptaków tego gatunku by test z pkt. (a) z prawdopodobieństwem 0,8, na poziomie istotności 0,05, przyjmował hipotezę, że średnia waga jest mniejsza niż 5, 20 kg.

(d) Załóżmy, że rzeczywista średnia waga ptaków jest równa 5, 15 kg. Wyznaczyć minimalną licznosc próby, która zagwarantuje, że test na poziomie istotności 0,05, z prawdopodobieństwem nie mniejszym niż 0,8, będzie przyjmował hipotezę, że średnia waga jest mniejsza niż 5, 20 kg.

(e) Czy na poziomie istotności 0,05 można stwierdzić, że odchylenie standardowe wagi ptaków badanego gatunku wynosi 0, 20 kg?

**ZADANIE 3.8** Agencja, badająca opinię publiczną, przeprowadziła badanie na reprezentatywnej próbie 1000 dorosłych Polaków. 721 z nich było przeciwnych paleniu w miejscach publicznych. Czy w świetle tych danych, możemy na poziomie istotności 0,1 uznać, że frakcja dorosłych Polaków przeciwnych paleniu w miejscach publicznych jest większa niż 70%?

**ZADANIE 3.9** 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%.

• Jak jest prawdopodobieństwo, że test z pkt. (a) potwierdzi, że odsetek regularnych klientów Żuczka jest większy w Warszawie niż w Krakowie?

•• 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 3.10** 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 3.11** W każdym miesiącu roku w wybrany losowo dzień zmierzono stężenie dwutlenku węgla w domu energooszczędnym oraz w domu zbudowanym tradycyjną techniką. 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 3.12 (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 rozkłady normalne.

**ZADANIE 3.13** Pomiary dokonane na próbach losowych 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 3.14** 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? Jeśli tak, to czy możemy stwierdzić, że ci pierwsi podczas testu na IQ werbalne uzyskują średnio ponad 1 pkt. więcej niż ci drudzy?

**ZADANIE 3.15** Do badania wybrano w sposób losowy 15 dzieci chorych na cukrzycę. Poddano ich 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.