

5. Testowanie hipotez

UWAGA! Większość danych pojawiających się w poniższych zadaniach dostępna jest w pliku *dane_hip.csv*

1. Aby mogły pracować urządzenia prądotwórcze elektrowni wiatrowej średnia prędkość wiatru powinna przekraczać 4 m/s. W celu stwierdzenia czy sensowna jest budowa elektrowni wiatrowej mierzono przez okres roku każdego miesiąca przeciętną prędkość wiatru w okolicach Darłowa uzyskując wyniki (w m/s):

5,9 4,4 5,4 3,8 4,0 4,2 3,4 3,6 4,6 6,5 5,6 4,8.

Zakładając, że prędkość wiatru jest zmienną losową o rozkładzie normalnym oraz przyjmując poziom istotności $\alpha=0,05$ sprawdź, czy okolice Darłowa nadają się do budowy elektrowni wiatrowych. W tym celu skonstruuj odpowiednią procedurę testującą. Porównaj otrzymane wyniki z uzyskanymi po zastosowaniu odpowiedniej funkcji z pakietu R.

$t=2,4186$, $p\text{-val}=0,017$, odrzucamy H_0

2. Przyjęto, że współczynnik efektywności pompy ciepłej COP jest zadawalający, gdy jego średnia wartość wynosi co najmniej 3,5 (co oznacza, że ponad 70% dostarczonego przez pompę ciepła pochodzi z naturalnego źródła ciepła, a reszta pochodzi z pracy sprężarki). Potencjalny nabywca pompy ma wątpliwości i wysunął przeproszenie, że współczynnik efektywności pompy ciepłej w jego gospodarstwie domowym jest znacznie mniejszy niż 3,5. Przez pewien okres mierzono współczynnik COP w tym gospodarstwie otrzymując następujące wyniki:

3,5 3,2 3,6 3,0 3,3 3,8 2,5 3,0 3,7 3,9.

Zakładając, że zmienna opisująca wartości współczynnika COP jest zmienną losową o rozkładzie normalnym i na podstawie powyższych wyników (przyjmując poziom istotności $\alpha=0,01$) sprawdź, czy wątpliwości nabywcy są słuszne.

$t=-1,09$, $p\text{-val}=0,15$, brak podstaw do odrzucenia H_0

3. Wiadomo, że rozkład wyników pomiarów głębokości morza w pewnym rejonie jest normalny z odchyleniem standardowym 5 m. Dokonano 5 niezależnych pomiarów głębokości morza w pewnym rejonie i otrzymano następujące wyniki (w m):

862 870 876 866 871.

Przyjmując poziom istotności $\alpha=0,05$ zweryfikuj hipotezę, że średnia głębokość morza w tym rejonie jest różna od 870 m.

$z=-0,447$, $p\text{-val}=0,65$, brak podstaw do odrzucenia H_0

4. Automat produkuje blaszki określonych wymiarów o nominalnej grubości 0,04 mm. Wylosowana próba 40 blaszek dała następujące wyniki:

0,048 0,028 0,037 0,033 0,054 0,046 0,041 0,043 0,044 0,05 0,047 0,052 0,053 0,048
 0,027 0,056 0,058 0,039 0,026 0,034 0,043 0,042 0,047 0,022 0,046 0,04 0,036
 0,043 0,041 0,044 0,043 0,044 0,038 0,046 0,041 0,038 0,047 0,03 0,041 0,049.

Przyjmując poziom istotności $\alpha=0,02$ sprawdź poprawność twierdzenia, że produkowane przez ten automat blaszki są grubsze niż nominalna grubość.

$z=1,64$, $p\text{-val}=0,0504$, brak podstaw do odrzucenia H_0

5. W próbie laboratoryjnej mleka spożywczego wykonano 10 oznaczeń (w %) zawartości tłuszczu i uzyskano:

1,5 1,8 1,5 1,7 1,6 1,6 1,8 1,6 1,7 1,6.

Przyjmując, że rozkład zawartości tłuszczu w mleku spożywczym jest normalny, odpowiedź na poniższe pytania (przyjmij poziom istotności $\alpha = 0,05$):

(a) Czy obserwacje przeczą hipotezie, że średnia zawartość tłuszczu w mleku wynosi 1,7 %?

$t=-1,765$, $p\text{-val}=0,11$, brak podstaw do odrzucenia H_0

(b) Czy można twierdzić, że wariancja zawartości tłuszczu w mleku jest mniejsza niż 0,02 (%)².

$\chi^2=5,2$, $p\text{-val}=0,18$, brak podstaw do odrzucenia H_0

6. Kukułki podrzucają swoje jaja różnym ptakom, między innymi małym strzyżynom. Obserwacje przyrodników wskazują, że kukułka potrafi znieść jajo wielkości podobnej do jaj „adopcyjnych rodziców”. Zmierzono długość [w mm] 21 jaj podrzuconych strzyżynom otrzymując wyniki:

17,93	18,52	19,66	14,30	17,52	20,76	20,26
19,82	21,40	16,54	18,64	17,62	20,79	19,14
16,74	14,93	18,56	15,43	15,19	21,05	20,79.

Wiadomo, że średnia długość jaj strzyżynki wynosi 17 mm, a odchylenie standardowe 2,5 mm. Zakładamy, że badana cecha ma rozkład normalny.

- (a) Na poziomie istotności 0,05 zweryfikuj stawianą przez przyrodników hipotezę dotyczącą:

(i) średniej długości podrzuconych jaj;

$t=2,84$, $p\text{-val}=0,01$ odrzucamy H_0

(ii) wariancji długości podrzuconych jaj.

$\chi^2=15,438$, $p\text{-val}=0,498$, brak podstaw do odrzucenia H_0

- (b) Zbuduj 95% przedział ufności dla średniej długości jaj podrzucanych strzyżynom. Jaki jest związek między skonstruowanym przedziałem ufności a decyzją podjętą przy testowaniu hipotez?

(17,36;19,37)

7. Agencja Ochrony Środowiska ustaliła dopuszczalne średnie zanieczyszczenie na terenach przemysłowych jako 55 miligramów na m^3 , (w promieniu 2 km od fabryki), przy odchyleniu standardowym 18 (miligramów na m^3). Ekolog zmierzył poziom zanieczyszczeń na terenie przemysłowym 100 razy, w różnych dniach i nocach, a następnie obliczył średnią i odchylenie standardowe pomiarów: odpowiednio 60 i 20 miligramów na m^3 . Przyjmując poziom istotności 0,01, zweryfikuj, czy dane potwierdzają, że fabryka działa niezgodnie z prawem (w celu weryfikacji hipotezy dotyczącej wariancji należy przyjąć normalność poziomu zanieczyszczeń).

średnia: $z=2,5$, $p\text{-val}=0,006$, odrzucamy H_0 ; wariancja: $\chi^2=122,22$, $0,05 < p\text{-val} < 0,06$, brak podstaw do odrzucenia H_0

8. Sondaż opinii publicznej na temat frekwencji oczekiwanej na wyborach wykazał, że w losowo wybranej grupie 2500 osób 1600 zamierza uczestniczyć w głosowaniu. Czy na poziomie istotności równym 0,05 próba przeczy twierdzeniu, że 60% ogółu osób zamierza wziąć udział w wyborach? Skonstruuj odpowiednią procedurę testującą. Porównaj otrzymane wyniki z uzyskanymi po zastosowaniu testu dokładnego dostępnego w pakiecie R.

$z=4,08$, odrzucamy H_0 / $p\text{-val}=0$, odrzucamy H_0

9. Przeprowadzono badanie jakości jaj kurzych pochodzących z pewnej fermi. Zakłada się z góry, że 2% jaj jest złej jakości. Wylosowano 1200 jaj do zbadania i wśród nich 16 okazało się złej jakości. Czy obserwacje przeczą przyjętemu założeniu i potwierdzają, że frakcja ta w badanej fermie jest mniejsza? Wnioskuj na poziomie istotności 0,05. W tym celu skonstruuj odpowiednią procedurę testującą. Porównaj otrzymane wyniki z uzyskanymi po zastosowaniu testu dokładnego dostępnego w pakiecie R.

$z=-1,65$, odrzucamy H_0 / $p\text{-val}=0,055$, brak podstaw do odrzucenia H_0

10. W sondażu przeprowadzonym przez pracownię badania opinii społecznych wśród 1100 dorosłych Polaków, 1000 ankietowanych odpowiedziało, że w ubiegłym miesiącu nie przeczytało żadnej książki. Pozostali twierdzili, że przeczytali przynajmniej jedną książkę. Na podstawie tych danych, na poziomie 0,05, stwierdź, czy opinia, że procent Polaków, którzy nie przeczytali żadnej książki jest większy niż 90 jest uzasadniona?

$z=1,005$, brak podstaw do odrzucenia H_0 / $p\text{-val}=0,17$, brak podstaw do odrzucenia H_0