

Statystyka stosowana

Raport 2

Temat: **Testowanie hipotez statystycznych**

Imię i Nazwisko prowadzącego kurs: **Mgr Katarzyna Maraj-Zygmunt**

Imię i Nazwisko, nr indeksu	Szymon Malec, 262276 Filip Oszczepaliński, 262292
Wydział:	Wydział matematyki, W13
Dzień i godzina zajęć:	Wtorek, 7 ³⁰
Kod grupy ćwiczeniowej:	T00-64c
Data oddania raportu:	21.06.2022
Ocena końcowa	

Adnotacje i uwagi:

1. Wstęp

Celem raportu jest rozwiązanie trzech zadań dostępnych na stronie [1]. Zadania te dotyczą dwóch zbiorów danych pochodzących z rozkładów normalnych. Naszym zadaniem jest przeprowadzenie testów statystycznych w celu zweryfikowania prawdziwości podanych w zadaniach hipotez oraz wyznaczenie metodą Monte Carlo prawdopodobieństwa popełnienia błędów I i II rodzaju.

2. Potrzebne definicje

2.1. Obszar krytyczny

Obszarem krytycznym nazywamy taki przedział, do którego należenie statystyki testowej prowadzi do odrzucenia hipotezy zerowej i przyjęcia hipotezy alternatywnej.

2.2. P-wartość

P-wartością prowadzonego testu nazywamy najmniejszy poziom istotności α , przy którym zaobserwowana wartość statystyki testowej prowadzi do odrzucenia hipotezy zerowej.

2.3. Błąd I rodzaju

Odrzucenie hipotezy zerowej, gdy jest ona prawdziwa, nazywamy błędem I rodzaju.

2.4. Błąd II rodzaju

Przy zadanej alternatywnej wartości parametru θ będącego przedmiotem testowania, błędem II rodzaju nazywamy przyjęcie hipotezy zerowej, gdy jest ona fałszywa.

2.5. Moc testu

Dla danej alternatywnej wartości θ , prawdopodobieństwa odrzucenia fałszywej hipotezy zerowej i przyjęcia prawdziwej hipotezy alternatywnej nazywamy mocą testu dla tej wartości.

3. Testowanie hipotez dotyczących wartości średniej

Zadanie 1 polega na zbadaniu próby X_1, \dots, X_n z populacji generalnej o rozkładzie normalnym $\mathcal{N}(\mu, 0.2)$. Dane można odnaleźć na stronie [2]. Długość próby wynosi $n = 1000$. Oznaczmy $\sigma^2 = 0,2$ oraz $\mu_0 = 1,5$. Naszym zadaniem jest zweryfikowanie hipotezy zerowej

$$H_0 : \mu = \mu_0 = 1,5$$

na poziomie istotności $\alpha = 0,05$. W tym celu konstruujemy statystykę

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}},$$

gdzie

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

jest nieobciążonym estymatorem parametru μ . Jeśli hipoteza zerowa jest prawdziwa, to

$$Z \sim \mathcal{N}(0,1).$$

Jeśli H_0 jest fałszywa, statystyka Z będzie miała tendencję do przyjmowania dużych lub małych wartości. Dla naszych danych wartość statystyki wynosi **-3.149**. Hipotezę zerową będziemy testować naprzeciw trzem hipotezom alternatywnym.

(a) **$H_1 : \mu \neq \mu_0$**

Definiujemy obszar krytyczny jako przedział

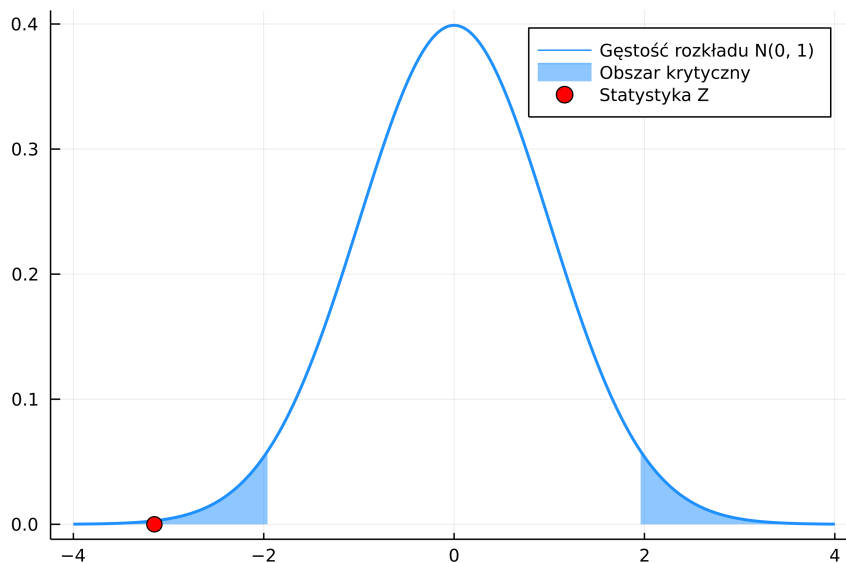
$$C = \left(-\infty ; -z_{1-\frac{\alpha}{2}}\right] \cup \left[z_{1-\frac{\alpha}{2}} ; \infty\right),$$

gdzie $z_{1-\frac{\alpha}{2}}$ jest kwantylem rzędu $1 - \frac{\alpha}{2}$ rozkładu $\mathcal{N}(0,1)$. Jeśli H_0 jest prawdziwa, to prawdopodobieństwo, że wartość statystyki będzie mieściła się w przedziale C , wynosi $\alpha = 0,05$. Dla naszych danych

$$C = (-\infty ; -1.96] \cup [1.96 ; \infty),$$

a więc wartość statystyki Z mieści się w tym przedziale, zatem odrzucamy hipotezę zerową i przyjmujemy hipotezę H_1 . P-wartość tego testu wynosi

$$2 \cdot P(Z > -3.149) \approx 0.0016.$$



Rysunek 1: Wykres gęstości rozkładu $\mathcal{N}(0, 1)$ z zaznaczonym obszarem krytycznym dla $H_1 : \mu \neq \mu_0$ oraz wartością statystyki Z .

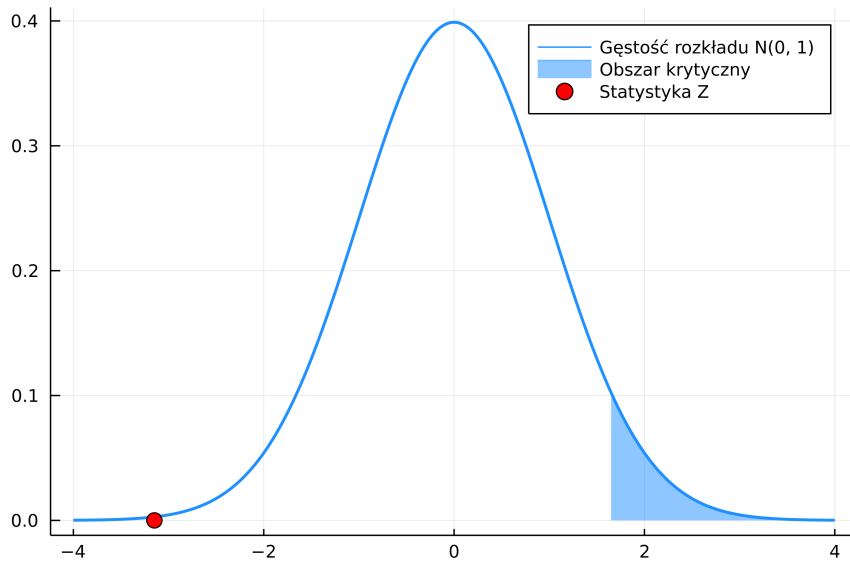
(b) $H_1 : \mu > \mu_0$

W tym przypadku obszar krytyczny ma postać

$$C = [z_{1-\alpha} ; \infty) = [1.645 ; \infty),$$

Wartość statystyki nie należy do tego przedziału, a więc przyjmujemy hipotezę zerową. P-wartość testu wynosi

$$P(Z > -3.149) \approx 0.9991.$$



Rysunek 2: Wykres gęstości rozkładu $\mathcal{N}(0, 1)$ z zaznaczonym obszarem krytycznym dla $H_1 : \mu > \mu_0$ oraz wartością statystyki Z .

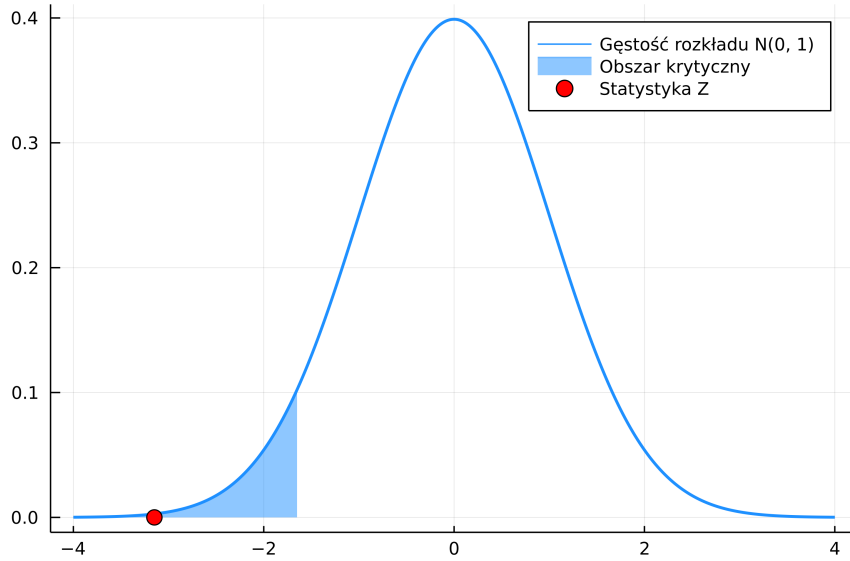
(c) $H_1 : \mu < \mu_0$

Obszar krytyczny w tym przypadku to

$$C = (-\infty ; -z_{1-\alpha}] = (-\infty ; -1.645].$$

Wartość statystyki Z mieści się w tym przedziale, zatem przyjmujemy H_1 . P-wartość dla tego testu wynosi

$$P(Z < -3.149) \approx 0.0008.$$



Rysunek 3: Wykres gęstości rozkładu $\mathcal{N}(0, 1)$ z zaznaczonym obszarem krytycznym dla $H_1 : \mu < \mu_0$ oraz wartością statystyki Z .

Gdy α będzie równa 0,01, to $z_{1-\frac{\alpha}{2}} = 2,576$, a $z_{1-\alpha} = 2,326$, więc wraz ze zmniejszaniem α dojdziemy do momentu, że w każdym z podpunktów hipoteza zerowa będzie przyjmowana. Gdy α będzie równa 0,01, to $z_{1-\frac{\alpha}{2}} = 1,645$, a $z_{1-\alpha} = 1,282$, więc wraz ze zwiększaniem α nic nie zmieni się w podpunkcie (a), natomiast dla α bliskiej 1, w podpunkcie (b), będziemy odrzucać hipotezę zerową.

Podsumowując, odrzuciliśmy hipotezę $H_1 : \mu > \mu_0$ przyjmując H_0 w podpunkcie (b) oraz przyjęliśmy hipotezy alternatywne z podpunktów (a) i (c). Bardzo małe p-wartości dla tych hipotez utwierdzają nas w przekonaniu, że są one prawdziwe. Zatem z dużą dozą pewności możemy stwierdzić, że $\mu < \mu_0 = 1,5$.

4. Testowanie hipotez dotyczących wariancji

W zadaniu 2 mamy zbadać próbę X_1, \dots, X_n z populacji generalnej o rozkładzie normalnym $\mathcal{N}(0,2, \sigma^2)$. Dane dostępne są na stronie [3]. Długość próby wynosi $n = 1000$. Oznaczmy $\mu = 0,2$ oraz $\sigma_0^2 = 1,5$. Naszym zadaniem jest zweryfikowanie hipotezy zerowej

$$H_0 : \sigma^2 = \sigma_0^2 = 1,5$$

na poziomie istotności $\alpha = 0,05$. W tym celu konstruujemy statystykę

$$\chi^2 = \frac{nS^2}{\sigma_0^2},$$

gdzie

$$S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$$

jest nieobciążonym estymatorem wariancji. Zauważmy, że

$$\chi^2 = \frac{nS^2}{\sigma_0^2} = \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma_0} \right)^2 \sim \chi^2(n),$$

pod warunkiem zachodzenia H_0 . Jeśli H_0 jest fałszywa, statystyka χ będzie miała tendencję do przyjmowania dużych lub małych wartości. Dla naszych danych wartość statystyki wynosi **1112,58**. Hipotezę zerową będziemy testować naprzeciw trzem hipotezom alternatywnym.

(a) **$H_1 : \sigma^2 \neq \sigma_0^2$**

Definiujemy obszar krytyczny jako przedział

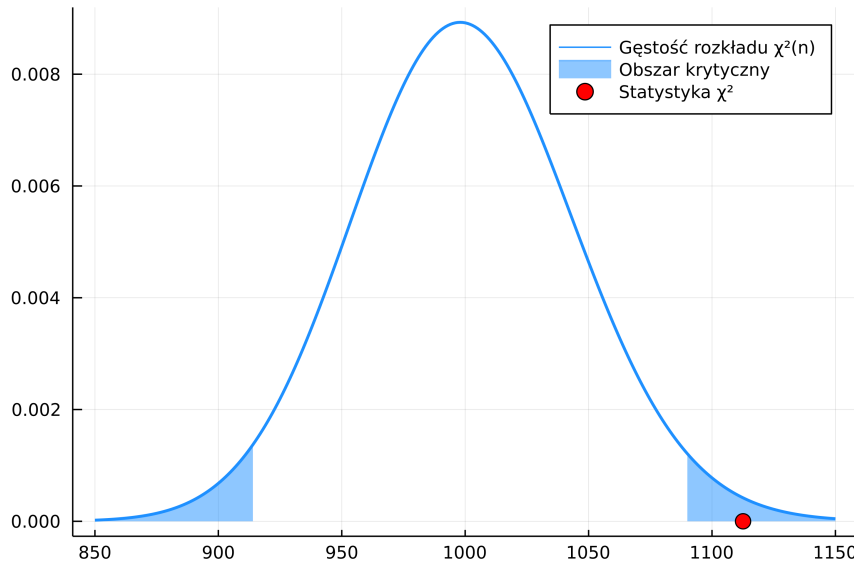
$$C = \left(-\infty ; x_{\frac{\alpha}{2},n}^2 \right] \cup \left[x_{1-\frac{\alpha}{2},n}^2 ; \infty \right),$$

gdzie $x_{\frac{\alpha}{2},n}^2$ i $x_{1-\frac{\alpha}{2},n}^2$ są kwantylami odpowiednio rzędu $\frac{\alpha}{2}$ i $1 - \frac{\alpha}{2}$ rozkładu χ^2 z n stopniami swobody. Jeśli H_0 jest prawdziwa, to prawdopodobieństwo, że wartość statystyki będzie mieściła się w przedziale C , wynosi $\alpha = 0,05$. Dla naszych danych

$$C = (-\infty ; 914,25] \cup [1089,53 ; \infty),$$

a więc wartość statystyki χ mieści się w tym przedziale, zatem odrzucamy hipotezę zerową i przyjmujemy hipotezę H_1 . P-wartość tego testu wynosi

$$2 \cdot P(\chi^2 > 1112,58) \approx 0,0145.$$



Rysunek 4: Wykres gęstości rozkładu $\chi^2(n)$ z zaznaczonym obszarem krytycznym dla $H_1 : \sigma^2 \neq \sigma_0^2$ oraz wartością statystyki χ^2 .

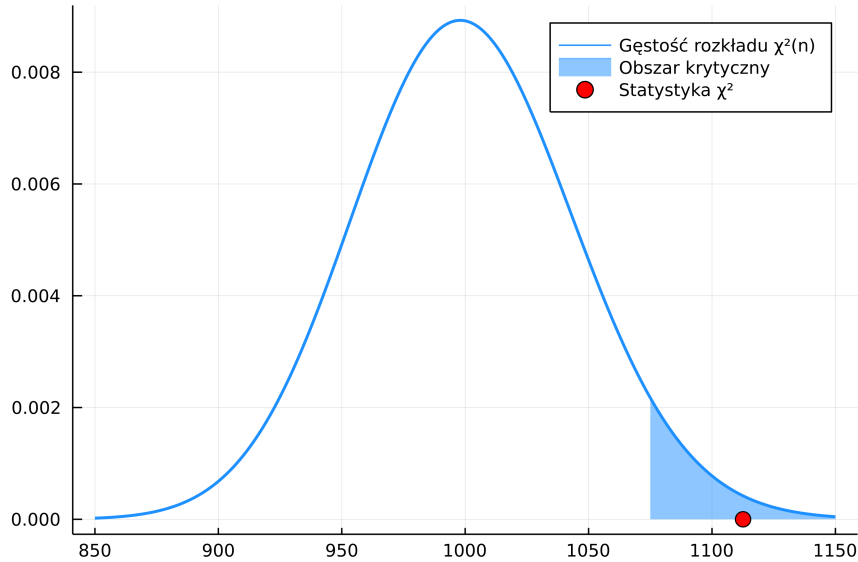
(b) **$H_1 : \sigma^2 > \sigma_0^2$**

W tym przypadku obszar krytyczny ma postać

$$C = \left[x_{1-\alpha,n}^2 ; \infty \right) = [1074,68 ; \infty),$$

czyli również wartość statystyki należy do tego przedziału, a więc przyjmujemy hipotezę H_1 . P-wartość testu wynosi

$$P(\chi^2 > 1112,58) \approx 0,0073.$$



Rysunek 5: Wykres gęstości rozkładu $\chi^2(n)$ z zaznaczonym obszarem krytycznym dla $H_1 : \sigma^2 > \sigma_0^2$ oraz wartością statystyki χ^2 .

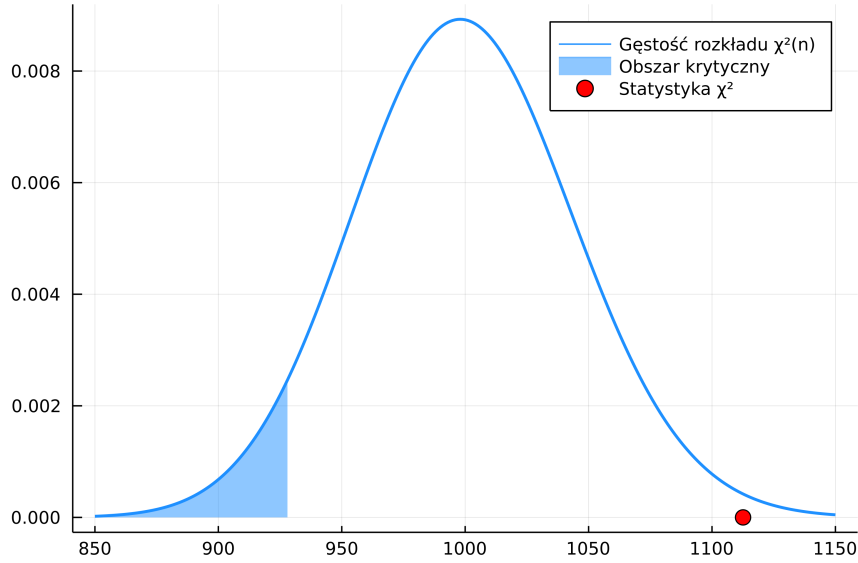
(c) $H_1 : \sigma^2 < \sigma_0^2$

Obszar krytyczny w tym przypadku to

$$C = (-\infty ; x_{\alpha,n}^2] = (-\infty ; 927,6].$$

Wartość statystyki χ^2 nie mieści się w tym przedziale, zatem przyjmujemy hipotezę zerową. P-wartość dla tego testu wynosi

$$P(\chi^2 < 1112,58) \approx 0,9927.$$



Rysunek 6: Wykres gęstości rozkładu $\chi^2(n)$ z zaznaczonym obszarem krytycznym dla $H_1 : \sigma^2 < \sigma_0^2$ oraz wartością statystyki χ^2 .

Podobnie, jak w przypadku hipotez dotyczących μ , gdy będziemy zmniejszać wartość α , będzie zmniejszał się obszar krytyczny, a więc w pewnym momencie dla każdego przypadku hipoteza zerowa będzie akceptowana. Jeśli z kolei zwiększymy wartość α , doprowadzimy do tego, że poza H_1 z podpunktów (a) i (b) zostanie przyjęta także hipoteza alternatywna z podpunktu (c), jednak wartość ta musiałaby być równa aż 0,9927.

Podsumowując, odrzuciliśmy hipotezę $H_1 : \sigma^2 < \sigma_0^2$ przyjmując H_0 w podpunkcie (c) oraz przyjęliśmy hipotezy alternatywne z podpunktów (a) i (b). Małe p-wartości dla tych hipotez wyraźnie sugerują, że są one prawdziwe. Zatem możemy stwierdzić, że z dużym prawdopodobieństwem $\sigma^2 > \sigma_0^2 = 1,5$.

5. Wyznaczenie symulacyjnie prawdopodobieństwa popełnienia błędów I i II rodzaju

Zadanie 3 polega na obliczeniu metodą Monte Carlo prawdopodobieństwa popełnienia błędów I i II rodzaju dla hipotez z dwóch poprzednich zadań. W przypadku hipotez dotyczących μ musimy wygenerować $N = 1000$ prób z rozkładu normalnego $\mathcal{N}(\mu_0, \sigma^2)$ o długości $n = 1000$. Z każdej próby wyliczamy statystykę Z , a więc otrzymujemy Z_1, \dots, Z_n . Następnie definiujemy zbiór

$$K_I = \{Z_i : Z_i \in C, i = 1, \dots, n\},$$

gdzie C jest obszarem krytycznym dla danej hipotezy alternatywnej. Prawdopodobieństwo popełnienia błędu I rodzaju wynosi

$$p_I = \frac{\#K_I}{N}.$$

Jeżeli chcemy obliczyć prawdopodobieństwo popełnienia błędu II rodzaju, wykonujemy te same kroki, tylko tym razem generujemy próby z rozkładu $\mathcal{N}(\mu_0 + \theta, \sigma^2)$, $\theta \neq 0$. Definiujemy zbiór

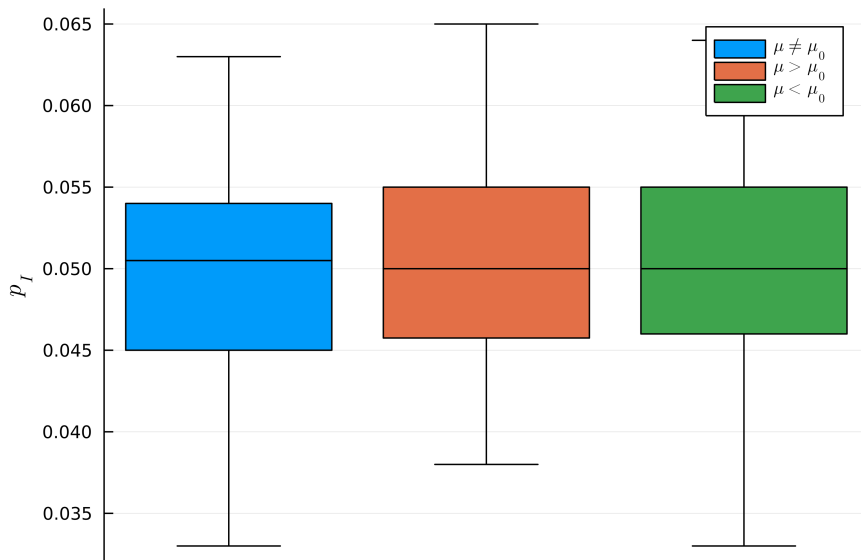
$$K_{II} = \{Z_i : Z_i \notin C, i = 1, \dots, n\}.$$

Szukane prawdopodobieństwo jest równe

$$p_{II} = \frac{\#K_{II}}{N}.$$

W przypadku hipotez dotyczących wariancji postępujemy analogicznie, tylko zamiast statystyki Z obliczamy statystykę χ^2 .

Gdy znamy już prawdopodobieństwo popełnienia błędu II rodzaju, w prosty sposób możemy policzyć moc testu, ponieważ jest ona równa $1 - p_{II}$. Poniżej widoczne są wyniki przeprowadzonych symulacji.



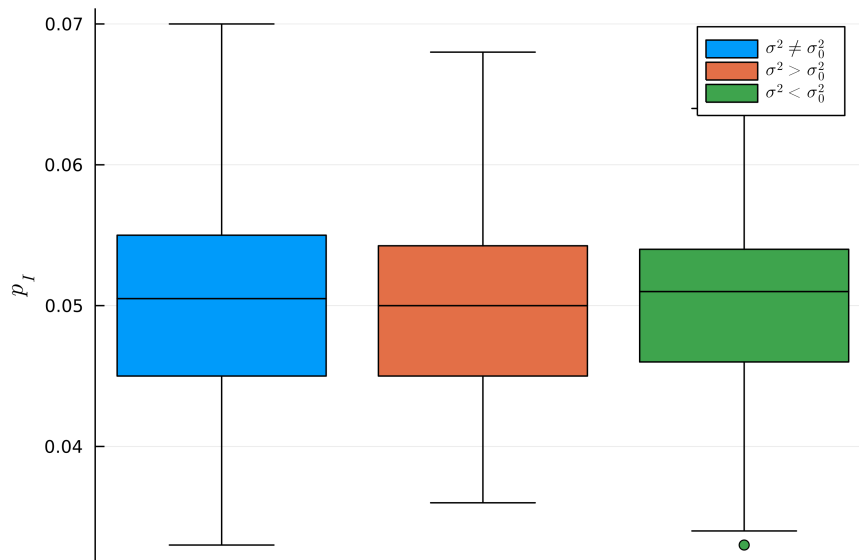
Rysunek 7: Wykresy pudełkowe 100 wysymulowanych wartości prawdopodobieństwa popełnienia błędu I rodzaju dla każdej z hipotez dotyczących parametru μ w pierwszym zadaniu.

μ	$\mu \neq \mu_0$	$\mu > \mu_0$	$\mu < \mu_0$
1,47	0,457	–	0,331
1,48	0,730	–	0,621
1,49	0,891	–	0,815
1,51	0,895	0,835	–
1,52	0,690	0,575	–
1,53	0,418	0,304	–

Tabela 1: Wartości prawdopodobieństwa popełnienia błędu II rodzaju dla podanych w 2. zadaniu hipotez alternatywnych, wyznaczone symulacyjnie dla różnych wartości μ .

μ	$\mu \neq \mu_0$	$\mu > \mu_0$	$\mu < \mu_0$
1,47	0,543	–	0,669
1,48	0,270	–	0,379
1,49	0,109	–	0,185
1,51	0,105	0,165	–
1,52	0,310	0,425	–
1,53	0,582	0,696	–

Tabela 2: Wartości mocy testu dla podanych w 2. zadaniu hipotez alternatywnych, wyznaczone symulacyjnie dla różnych wartości μ .



Rysunek 8: Wykresy pudełkowe 100 wysymulowanych wartości prawdopodobieństwa popełnienia błędu I rodzaju dla każdej z hipotez dotyczących parametru σ^2 w drugim zadaniu.

σ^2	$\sigma^2 \neq \sigma_0^2$	$\sigma^2 > \sigma_0^2$	$\sigma^2 < \sigma_0^2$
1,47	0,931	–	0,891
1,48	0,951	–	0,921
1,49	0,947	–	0,933
1,51	0,922	0,92	–
1,52	0,952	0,925	–
1,53	0,927	0,891	–

Tabela 3: Wartości prawdopodobieństwa popełnienia błędu II rodzaju dla podanych w 2. zadaniu hipotez alternatywnych, wyznaczone symulacyjnie dla różnych wartości σ^2 .

σ^2	$\sigma^2 \neq \sigma_0^2$	$\sigma^2 > \sigma_0^2$	$\sigma^2 < \sigma_0^2$
1,47	0,069	–	0,109
1,48	0,049	–	0,079
1,49	0,053	–	0,067
1,51	0,078	0,08	–
1,52	0,049	0,075	–
1,53	0,073	0,109	–

Tabela 4: Wartości mocy testu dla podanych w 2. zadaniu hipotez alternatywnych, wyznaczone symulacyjnie dla różnych wartości σ^2 .

6. Podsumowanie

Po przeprowadzeniu testów statystycznych dla hipotez dotyczących parametru μ w pierwszym zadaniu, wnioskujemy, że najprawdopodobniej μ jest mniejsze od 1,5. W przypadku próby z zadania drugiego testy wyraźnie wskazują, że parametr σ^2 jest większy od 1,5.

Następnie wyznaczyliśmy symulacyjnie prawdopodobieństwo popełnienia błędu I i II rodzaju dla badanych hipotez. Jak możemy zauważyć na wykresach pudełkowych, wartości prawdopodobieństwa błędu I rodzaju oscylują wokół $\alpha = 0,05$. W przypadku wyników symulacji dla błędu II rodzaju, tak jak byśmy przypuszczali, wraz ze zwiększaniem różnicy pomiędzy prawdziwym parametrem, a zakładanym w H_0 , maleje wartość prawdopodobieństwa i wzrasta moc testu.

Źródła

- [1] http://prac.im.pwr.edu.pl/~wyloman/statystyka_stosowana_2122/lista7.pdf
- [2] http://prac.im.pwr.edu.pl/~wyloman/statystyka_stosowana_2122/lista7_zad1.txt
- [3] http://prac.im.pwr.edu.pl/~wyloman/statystyka_stosowana_2122/lista7_zad2.txt