Testowanie hipotez w rodzinie rozkładów normalnych

Adam Wrzesiński, Joanna Kołaczek

18.07.2022

Spis treści

1 Przypadek nieznanej średniej **3**

1.1 $H_1 : \mu \neq \mu_0$ 4

1.2 $H_1 : \mu > \mu_0$ 5

1.3 $H_1 : \mu < \mu_0$ 6

2 Przypadek nieznanej wariancji **7**

2.1 $H_1 : \sigma^2 \neq \sigma_0^2$ 8

2.2 $H_1 : \sigma^2 > \sigma_0^2$ 9

2.3 $H_1 : \sigma^2 < \sigma_0^2$ 10

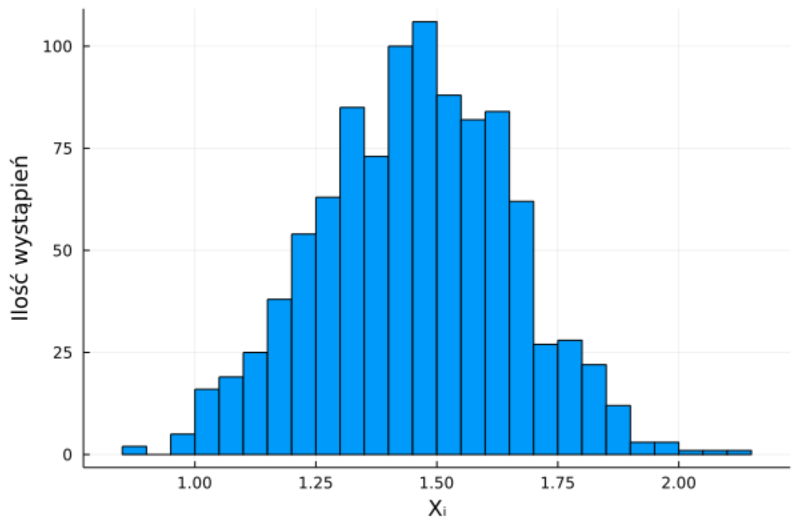
3 Błędy statystyczne **11**

Wstęp

Niniejszy raport powstał na potrzeby realizacji laboratorium ze Statystyki Stosowanej, prowadzonych przez dr inż. Aleksandrę Grzesiek, do wykładu dr hab. inż. Krzysztofa Burneckiego. Będziemy testować hipotezy statystyczne dla wartości średniej i wariancji w rodzinie rozkładów normalnych. W tym celu zobrazujemy obszary krytyczne, wyznaczymy p-wartości oraz prawdopodobieństwa wystąpienia błędów I i II rodzaju. Życzymy miłej lektury.

1 Przypadek nieznanej średniej

Dysponujemy próbą danych z rozkładu $N(\mu,0.2)$. Jej rozkład wizualizujemy za pomocą histogramu (rys. 1).



Rysunek 1: Histogram

Nasza hipoteza zerowa to $H_0 : \mu_0 = 1.5$. Statystykę testową wyznaczamy ze wzoru (1.1).

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

Jeżeli H_0 jest prawdziwa, to $Z \sim N(0,1)$. Aby obliczyć średnią $\bar{X}$ skorzystamy ze wzoru na średną (1.2).

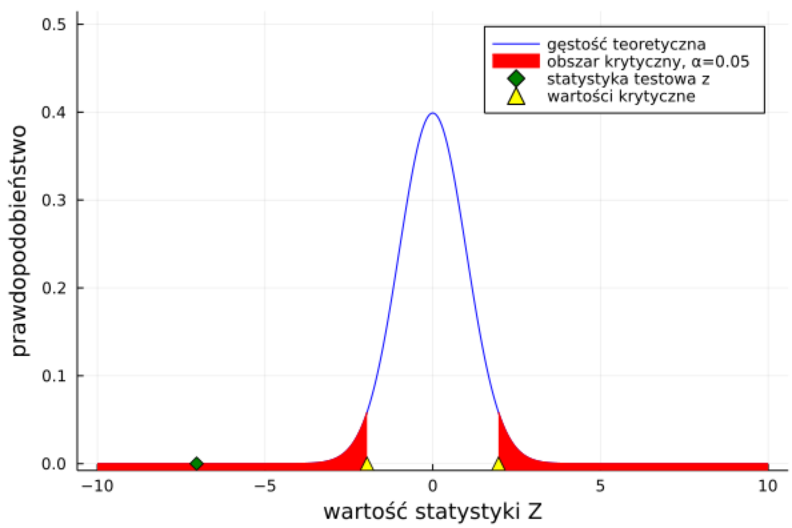
$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

W naszym przypadku $\bar{X} = 1.455$, $z = -7.041$. Na poziomie istotności $\alpha = 0.05$ zweryfikujemy kolejne hipotezy alternatywne H_1.

1.1 $H_1 : \mu \neq \mu_0$

Zaczynamy od zadania sobie pytania - "Czy rzeczywista średnia jest różna od zakładanej?". Wobec tego średnia może być mniejsza lub większa od podważanej. Ponieważ gęstość rozkładu standardowego jest funkcją parzystą, to $Z_\alpha = -Z_{1-\alpha}$. Korzystając z tej równości stwierdzamy, że pole pod wykresem na obszarze krytycznym również dzieli się na dwie połowy. Stąd wnioskujemy, że p-wartość jest dwukrotnością pola na dowolej połowie obszaru krytycznego wyznaczanego przez statystykę z. Podsumowując:

- Obszar krytyczny $c = \{x : x \leq -Z_{1-\alpha/2} \vee x \geq Z_{1-\alpha/2}\}$.
- p-wartość $p = 2P_{H_0}(Z \geq |z|)$



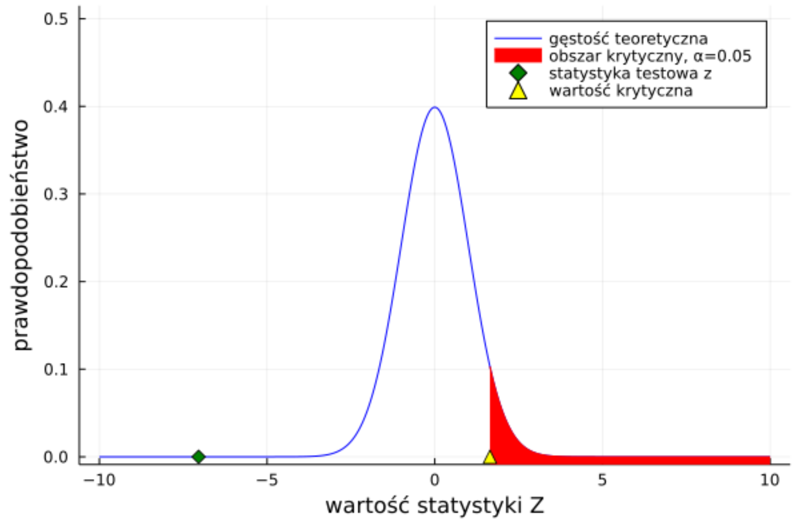
Rysunek 2: Obszar krytyczny na tle gęstości Z

W naszym przypadku $c = (-\infty, -1.96) \cup (1.96, \infty)$ oraz $p = 1.9 \cdot 10^{-12}$.
Zaznaczamy na wykresie obszar krytyczny c oraz statystykę z (rys. 2). Ponieważ statystyka z znalazła się w obszarze krytycznym możemy przyjąć naszą hipotezę alternatywną $H_1 : \mu \neq 1.5$. Ponadto p-wartość jest bardzo mała - prawie zawsze odrzucimy hipotezę zerową.

1.2 $H_1 : \mu > \mu_0$

Zadajemy pytanie - "Czy rzeczywista średnia jest większa od zakładanej?". Spodziewamy się więc jednostronnego obszaru krytycznego. Upraszcczają się zatem postacie tego obszaru oraz p-wartości:

- Obszar krytyczny $c = \{x : x \geq z_{1-\alpha}\}$.
- p-wartość $p = P_{H_0}(Z \geq z)$



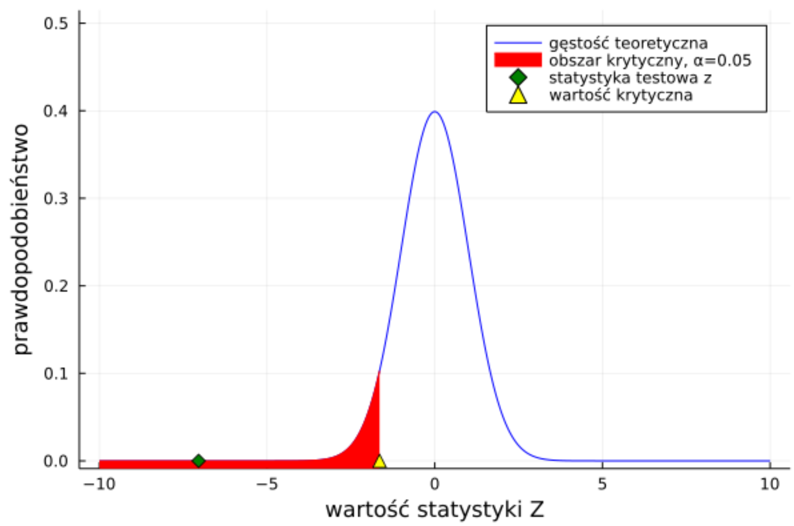
Rysunek 3: Obszar krytyczny na tle gęstości Z

W naszym przypadku $c = (1.64, \infty)$ oraz $p = 0.999$.
Zaznaczamy na wykresie [3] c oraz statystykę z. Ponieważ statystyka z znalazła się poza obszarem krytycznym odrzucamy hipotezę alternatywną, duża p-wartość dodatkowo umacnia nas w przekonaniu, że nie możemy odrzucić hipotezy zerowej.

1.3 $H_1 : \mu < \mu_0$

Ponownie zaczynamy od znalezienia postaci obszaru krytycznego oraz p-wartości. Tym razem interesuje nas przedział na lewo od odpowiednich wartości krytycznych:

- Obszar krytyczny $c = \{x : x \leq -z_{1-\alpha}\}$.
- p-wartość $p = P_{H_0}(Z \leq z)$



Rysunek 4: Obszar krytyczny na tle gęstości Z

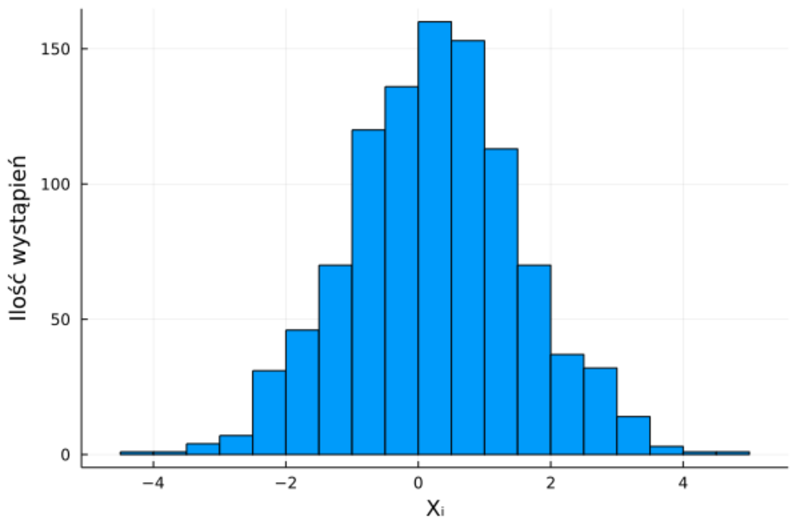
W naszym przypadku $c = (-\infty, -1.64)$ oraz $p = 9.51 \cdot 10^{-13}$.
Sprawdźmy, czy statystyka z znalazła się w obszarze krytycznym c (rys.4). Ponieważ znalazła się ona w obszarze krytycznym odrzucamy hipotezę zerową na korzyść hipotezy alternatywnej. $H_1 : \mu < 1.5$. Ponadto p-wartość jest bardzo mała - prawie zawsze odrzucimy hipotezę zerową.

Wniosek

Odrzucamy hipotezę zerową. Z wyjątkowo dużą śmialością stwierdzamy, że rzeczywista średnia jest mniejsza od zakładanej na poziomie istotności 5%.

2 Przypadek nieznanej wariancji

Dysponujemy próbą danych z rozkładu $N(0.2, \sigma^2)$, której rozkład możemy zobaczyć na histogramie (rys. 5).



Rysunek 5: Histogram

Nasza hipoteza zerowa to $H_0 : \sigma_0^2 = 1.5$. Statystykę testową wyznaczamy ze wzoru (2.1).

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2},$$

gdzie n i S to odpowiednio długość oraz wariancja badanej próby. Jeżeli H_0 jest prawdziwa, to $Z \sim \chi^2$z $n-1$ stopniami swobody. Przypomnijmy wzór na wariancję z próbki (2.2).

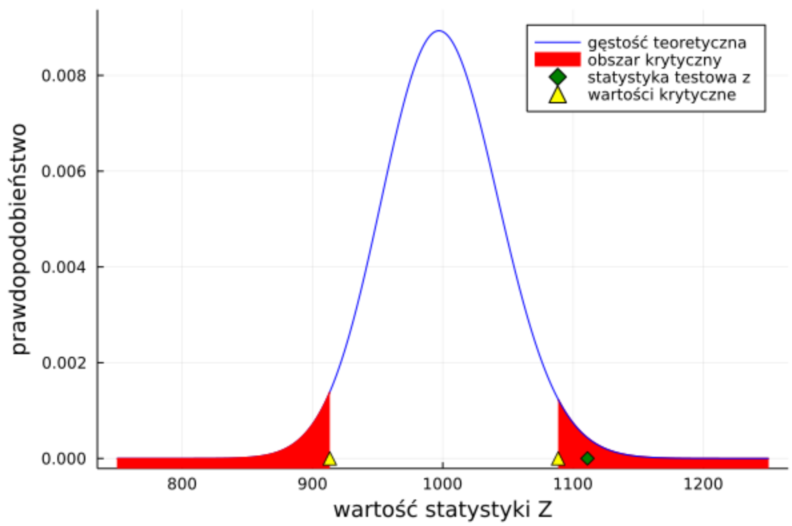
$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{X})^2.$$

Po podstawieniu do wzorów otrzymujemy $z = 1110.968$ oraz $S^2 = 1.668$. Zweryfikujemy kolejne hipotezy na poziomie istotności $\alpha = 0.05$.

2.1 $H_1 : \sigma^2 \neq \sigma_0^2$

Nie zmenia się za dużo w porównaniu rozdziałem poprzednim. Zamiast standardowego, interesuje nas rozkład χ^2 z $n-1$ stopniami swobody. W związku z tym nie posiadamy już symetrycznej gęstości. Wobec tego największe zmiany zachodzą dla p-wartości. Tym razem jest to dwukrtoność mniejszego pole na lewo lub prawo od niej. Intuicyjnie - jeśli weźmiemy większe pole, to skończymy z prawdopodobieństwem większym od 1, w związku z czym dalsze rozważania byłyby bezsensowne. Podsumowując:

- Obszar krytyczny $c = \{x : x^2 \leq \chi^2_{\alpha/2,n-1} \vee x^2 \geq \chi^2_{1-\alpha/2,n-1}\}$.
- p-wartość $p = 2min(P_{H_0}(Z \leq z),P_{H_0}(Z \geq z))$



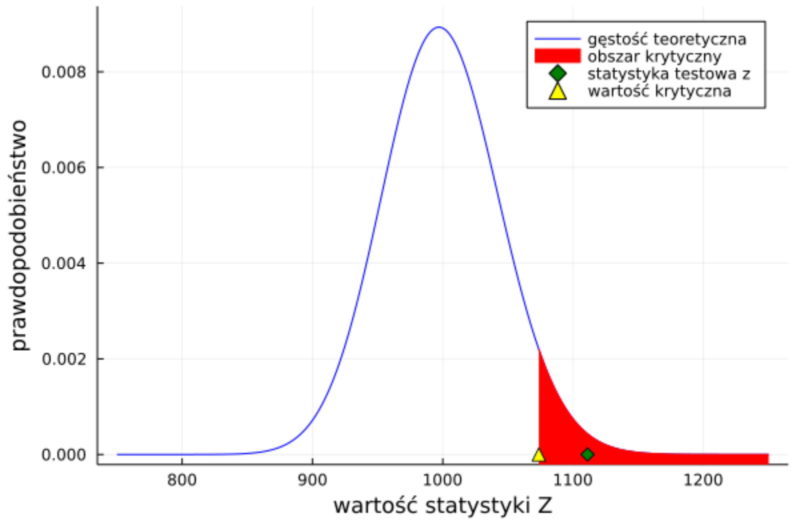
Rysunek 6: Obszar krytyczny na tle gęstości χ^2

W naszym przypadku $c = [0, 913.3) \cup (1088.487, \infty)$ oraz $p = 0.015$.
Zaznaczamy na wykresie [6] c oraz statystykę z. Ponieważ statystyka z znalazła się w obszarze krytycznym możemy przyjąć naszą hipotezę alternatywną $H_1 : \mu \neq 1.5$. Ponadto p-wartość jest bardzo mała - prawie zawsze odrzucimy hipotezę zerową.

2.2 $H_1 : \sigma^2 > \sigma_0^2$

Obszary krytyczne oraz p-wartość konstruujemy analogicznie względem wcześniejszych rozważań.

- Obszar krytyczny $c = \{x : x^2 \geq \chi^2_{1-\alpha,n-1}\}$.
- p-wartość $p = P_{H_0}(Z \geq z)$

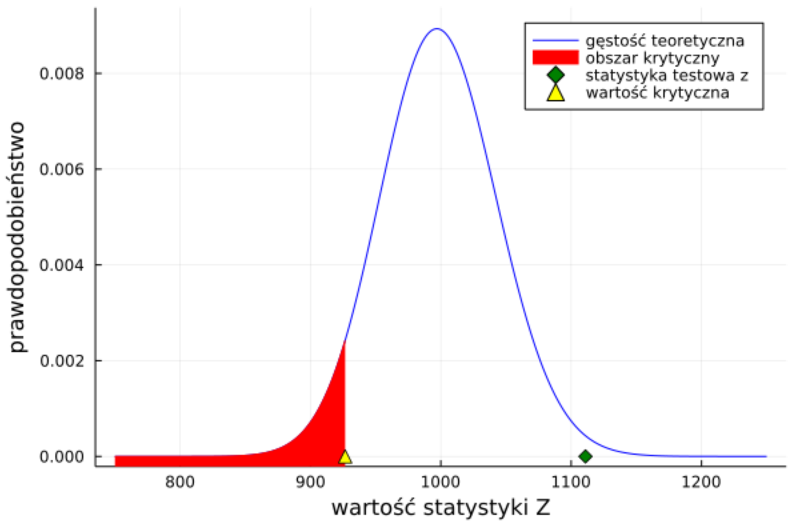


Rysunek 7: Obszar krytyczny na tle gęstości χ^2

W naszym przypadku $c = [1073.643, \infty)$ oraz 0.008.
Ponieważ statystyka z znalazła się w obszarze krytycznym (rys. 7) odrzucamy hipotezę zerową na korzyść hipotezy alternatywnej. Ponadto p-wartość jest bardzo mała - prawie zawsze odrzucimy hipotezę zerową.

2.3 $H_1 : \sigma^2 < \sigma_0^2$

- Obszar krytyczny $c = \{x : x^2 \leq \chi^2_{\alpha,n-1}\}$.
- p-wartość $p = P_{H_0}(Z \leq z)$



Rysunek 8: Obszar krytyczny na tle gęstości χ^2

W naszym przypadku $c = [0,\, 926.631)$ oraz $p = 0.992$.
Zaznaczamy na wykresie [8] c oraz statystykę z. Ponieważ statystyka z znalazła się poza obszarem krytycznym odrzucamy hipotezę alternatywną, duża p-wartość również wskazuje na to, że nie możemy odrzucić hipotezy zerowej.

Co stanie się kiedy zwiększymy bądź zmniejszymy poziom istotności?

W przypadku zwiększenia poziomu istotności α, zwiększy się nasz obszar krytyczny - test stanie się mniej wiarygodny. Zmniejszając poziom istotności, zmniejszamy obszar krytyczny - test jest bardziej wiarygodny. Można zatem powiedzieć, że najbardziej korzystne będzie wybieranie jak najmniejszych α tak aby uniknąć ryzyka odrzucenia prawdziwej hipotezy zerowej. Niestety w praktyce często możemy przeprowadzić tylko ograniczoną ilość badań, eksperymentów, pomiarów, dlatego wybór bardzo małego poziomu istotności nie jest opłacalny. Zwykle przyjmujemy $\alpha = 0.05$ (tzn. jesteśmy skłonni popełnić jeden błąd na 20 przypadków). P-wartość nie zależy od poziomu istotności.

3 Błędy statystyczne

Błąd I rodzaju - prawdopodobieństwo odrzucenia hipotezy zerowej, gdy ta jest prawdziwa. Jego teoretyczna wartość jest równa pozimowi istotności alpha.

Błąd II rodzaju - prawdopodobieństwo przyjęcia fałszywej hipotezy zerowej i odrzucenia prawdziwej hipotezy alternatywnej.

Moc testu - prawdopodobieństwo uniknięcia błędu drugiego rodzaju.

Przypadek dla nieznanej średniej

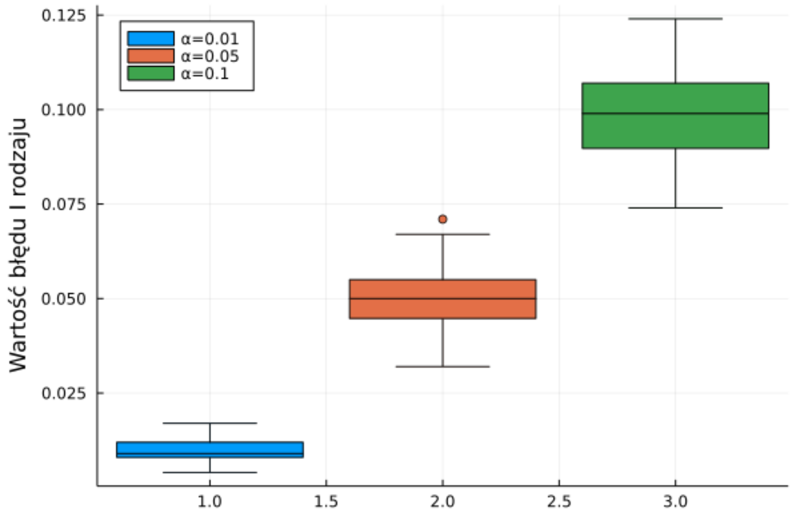
Symulacyjnie obliczymy prawdopodobieństwa popełnienia błędów I oraz II rodzaju. Wyznaczymy ponadto moc testu. Do pokazania zależności wykorzystamy wykresy pudełkowe, a rozważania będziemy podsumowywać za pomocą tabel.

Błąd I rodzaju

Algorytm:

1. Ustal poziom ufności α, długość n próby losowej X oraz postać hipotezy alternatywnej.
2. Wstaw s = 0.
3. Generuj próbę losową X z rozkładu $N(\mu_0, \sigma)$
4. Wyznacz wartość statystyki testowej Z.
5. Wyznacz odpowiedni obszar krytyczny.
6. Sprawdź, czy statystyka Z znalazła się w obszarze krytycznym. Jeśli tak, to zwiększ s o jeden.
7. Powtórz 3-6 N razy.
8. Zwróć $P_I = \frac{s}{N}$.

W naszym przypadku $\alpha \in \{0.01, 0.05, 0.1\}$, $n = 1000$, hipotezy takie jak omawiane w dwóch poprzednich rozdziałach, $N = 1000$. W celu wykonania boxplotów oraz tabel wykonujemy po 100 symulacji dla każdego przypadku.



Rysunek 9: Błąd I rodzaju w zależności od poziomu istotności α

α	Błąd I rodzaju dla średniej
0.01	0.099
0.05	0.049
0.1	0.94

Tabela 1: Średnia błędu I rodzaju dla 100 symulacji

Przeprowadzamy symulacje dla prób z rozkładu $N(1.5, 0.2)$ ($H_0 : \mu_0 = 1.5$ jest prawdziwa). Widzimy (rys. 9, tab. 1) że częstość symulacji, w których popełniamy błąd I rodzaju dąży do jego teoretycznej wartości (poziomowi istotności testu). Oznacza to, że algorytm wykorzystany do testowania hipotez działa poprawnie. Popełniamy więc $\frac{1}{\alpha}$ błędów na 100 symulacji.

Błąd II rodzaju

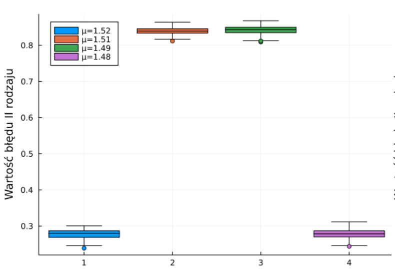
Algorytm:

1. Ustal poziom ufności α, średnią μ, odchylenie standardowe σ, długość n próby X oraz postać hipotezy alternatywnej.
2. Wstaw s = 0.
3. Generuj X z rozkładu $N(\mu, \sigma)$.
4. Wyznacz wartość statystyki testowej Z.
5. Wyznacz odpowiedni obszar krytyczny.
6. Sprawdź czy statystyka Z jest poza obszarem krytycznym. Jeśli tak, to zwiększ s o jeden.
7. Powtórz kroki 3-6 N razy.
8. Zwróć $P_{II} = \frac{s}{N}$.

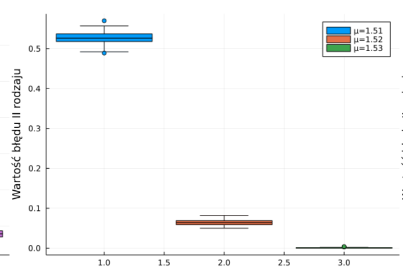
Ustalamy $\alpha = 0.05$, $\mu_0 = 1.5$, $\sigma = 0.2$, $n = 1000$ oraz następujące dane:

- dla $H_1 : \mu \neq \mu_0$, $\mu \in \{1.52; 1.51; 1.49; 1.48\}$,
- dla $H_1 : \mu > \mu_0$, $\mu \in \{1.51; 1.52; 1.53\}$,
- dla $H_1 : \mu < \mu_0$, $\mu \in \{1.49; 1.48; 1.47\}$.

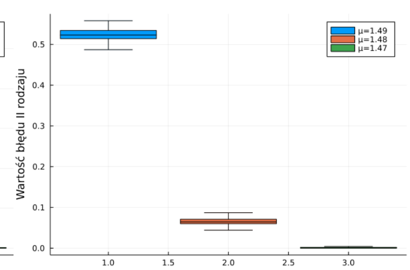
Zauważmy (rys. 10, 11, 12, tab. 2), że im bliżej μ_0 znajduje się μ tym większe prawdopodobieństwo popełnienia błędu II rodzaju, co skutkuje mniejszą mocą testu. Dla naszych danych gdy $|\mu - \mu_0| \geq 0.02$ test zwraca wyniki, które możemy uznać za rzetelne.



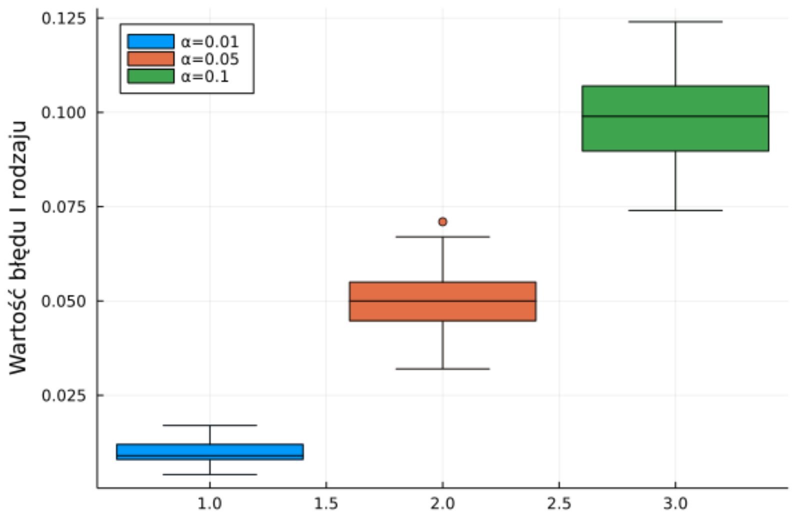
Rysunek 10: $H_1 : \mu \neq 1.5$



Rysunek 11: $H_1 : \mu > 1.5$



Rysunek 12: $H_1 : \mu < 1.5$



Rysunek 13: Błąd I rodzaju w zależności od poziomu ufności α

H_1	μ	Błąd II rodzaju dla średniej	Moc testu
$\mu \neq 1.5$	1.52	0.280	0.720
	1.51	0.839	0.161
	1.49	0.843	0.157
	1.48	0.279	0.721
$\mu > 1.5$	1.51	0.527	0.473
	1.52	0.064	0.936
	1.53	0.001	0.999
$\mu < 1.5$	1.49	0.523	0.477
	1.48	0.065	0.935
	1.47	0.001	0.999

Tabela 2: Średnia błędu II rodzaju dla 100 symulacji

Przypadek dla nieznanej wariancji

Przeprowadzimy dokładnie takie same symulacje dla tego drugiego przypadku. Zauważmy, że algorytmy są nieomal identyczne jak w przypadku nieznanej średniej. Zmienia się w zasadzie jedynie statystyka testowa i rozkład generowanej próby losowej. Dane oraz liczbę symulacji przyjmujemy takie same (zamieniamy miejscami średnią i wariancję). Dodatkowo mamy na uwadze, że tym razem drugim parametrem rozkładu normalnego jest wariancja (a nie jak w poprzednim przypadku odchylenie standardowe).

Błąd I rodzaju

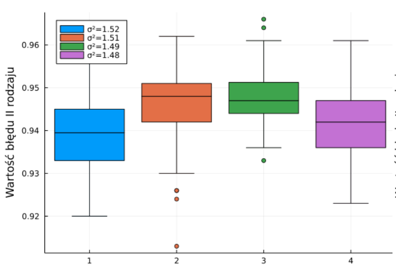
α	Błąd I rodzaju dla wariancji
0.01	0.01
0.05	0.049
0.1	0.094

Tabela 3: Średnia błędu I rodzaju dla 100 symulacji

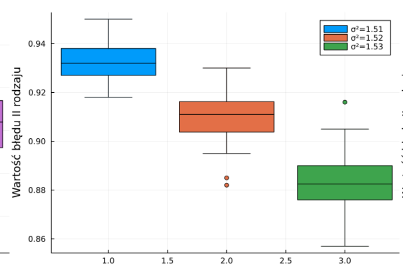
Po zwizualizowaniu wyników symulacji (rys. 13, tab. 3) , widzimy że odrzucamy prawdziwą hipotezę zerową w liczbie przypadków bliskiej określonemu teoretycznie poziomowi istotności testu. Oznacza to, że algorytm wykorzystany do testowania hipotez działa poprawnie. Ponownie, popełniamy $\frac{1}{\alpha}$ błędów na 100 symulacji.

Błąd II rodzaju

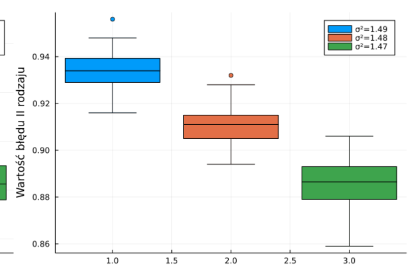
To samo robimy dla symulacji obliczeń błędu drugiego rodzaju. Zauważmy (rys. 14, 15, 16, tab.4), że im bliżej σ_0^2 znajduje się σ^2 tym większe prawdopodobieństwo popełnienia błędu II rodzaju, co skutkuje mniejszą mocą testu. Dla naszych σ^2 moc testu jest bardzo mała - istnieje duża szansa, że odrzucimy prawdziwą hipotezę alternatywną na korzyść fałszywej hipotezy zerowej.



Rysunek 14: $H_1 : \sigma^2 \neq 1.5$



Rysunek 15: $H_1 : \sigma^2 > 1.5$



Rysunek 16: $H_1 : \sigma^2 < 1.5$

H_1	σ	Błąd II rodzaju dla watiancji	moc
	1.52	0.939	0.061
	1.51	0.947	0.053
	1.49	0.947	0.053
$\sigma \neq 1.5$	1.48	0.942	0.058
	1.51	0.932	0.068
	1.52	0.911	0.089
$\sigma > 1.5$	1.53	0.883	0.117
	1.49	0.934	0.066
	1.48	0.911	0.089
$\sigma < 1.5$	1.47	0.887	0.113

Tabela 4: Mediana błędu II rodzaju dla 100 symulacji