

Statystyka Stosowana - testowanie hipotez statystycznych

Paulina Mruczek (277428), Maciej Szczerczowski (277487)

1. Wstęp

Celem niniejszego projektu jest przeprowadzenie weryfikacji hipotez statystycznych dotyczących parametrów populacji normalnych na podstawie próby losowej. Projekt obejmuje analizę dwóch przypadków:

1. Weryfikacja hipotez dotyczących średniej μ przy znanym odchyleniu standardowym $\sigma = 0,2$;
2. Weryfikacja hipotez dotyczących wariancji σ^2 przy znanej średniej $\mu = 0,2$.

W obu przypadkach przyjęty został poziom istotności $\alpha = 0,05$, a analiza została przeprowadzona dla trzech różnych wariantów hipotez: jednostronnych i dwustronnej. Dla wszystkich wariantów obliczone zostały wartości statystyk testowych, zbiory krytyczne oraz wartości p . Ocenie poddano również wpływ zmiany poziomu ufności na wyniki testów. Dodatkowo zasymulowane zostały prawdopodobieństwa popełnienia błędów I i II rodzaju oraz moc testów.

Projekt ilustruje praktyczne zastosowanie testowania hipotez w sytuacji, gdy znane są niektóre parametry populacji, a inne podlegają weryfikacji. Wnioski wyciągnięte na podstawie wyników testów mają na celu lepsze zrozumienie zależności między poziomem istotności, mocą testu, a prawdopodobieństwami błędów statystycznych.

2. Testy dla wartości średniej

Na podstawie danych z próby wyznaczono wartość statystyki testowej dla znanego odchylenia standardowego:

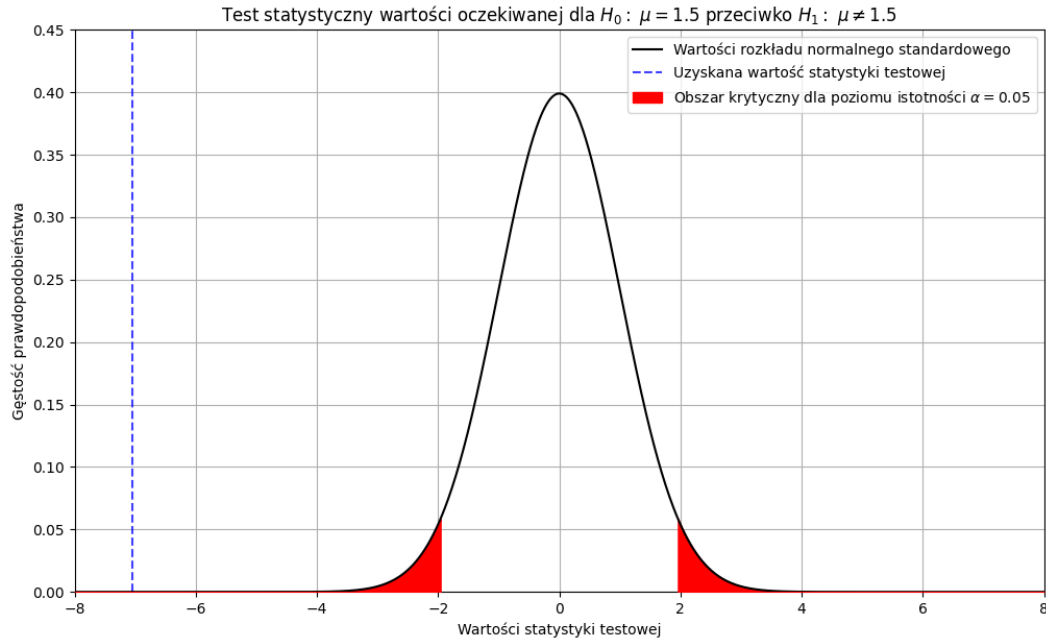
$$Z = \frac{\bar{X} - \mu_0}{\sqrt{\frac{\sigma^2}{n}}} \approx -7,04.$$

Przyjęta została hipoteza zerowa $H_0 : \mu_0 = 1,5$. W dalszej części przeanalizowano trzy warianty hipotezy alternatywnej.

2.1. $H_1 : \mu \neq 1,5$

- Obszar krytyczny $C = (-\infty; -1,96) \cup (1,96; \infty)$.
- Wartość statystyki testowej $Z \in C$.
- $p_{\text{wartość}} = 1,90 \cdot 10^{-12} < \alpha$.

Wartość statystyki testowej należy do zbioru krytycznego, dlatego odrzucamy hipotezę zerową na korzyść hipotezy alternatywnej. Średnia istotnie różni się od 1,5.

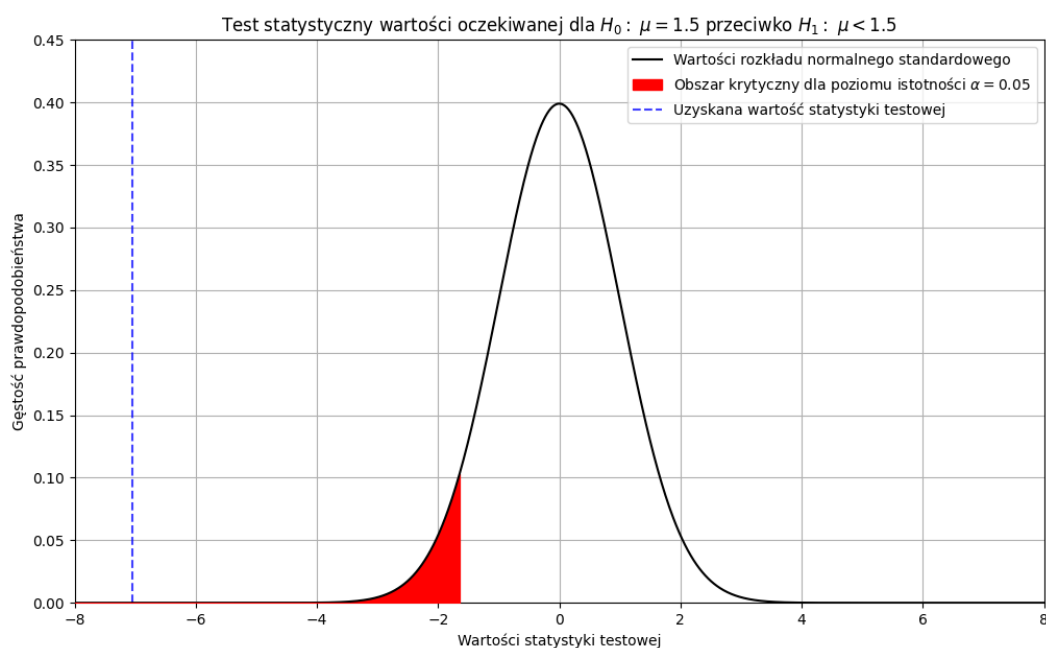


Rysunek 1: Wykres obszarów krytycznych dla hipotezy alternatywnej $\mu \neq 1,5$

2.2. $H_1 : \mu < 1,5$

- Obszar krytyczny $C = (-\infty; -1,64)$.
- Wartość statystyki testowej $Z \in C$.
- $p_{\text{wartość}} = 9,51 \cdot 10^{-13} < \alpha$.

Wartość statystyki testowej należy do zbioru krytycznego, dlatego odrzucamy hipotezę zerową na korzyść hipotezy alternatywnej. Średnia populacji jest istotnie mniejsza od 1,5.

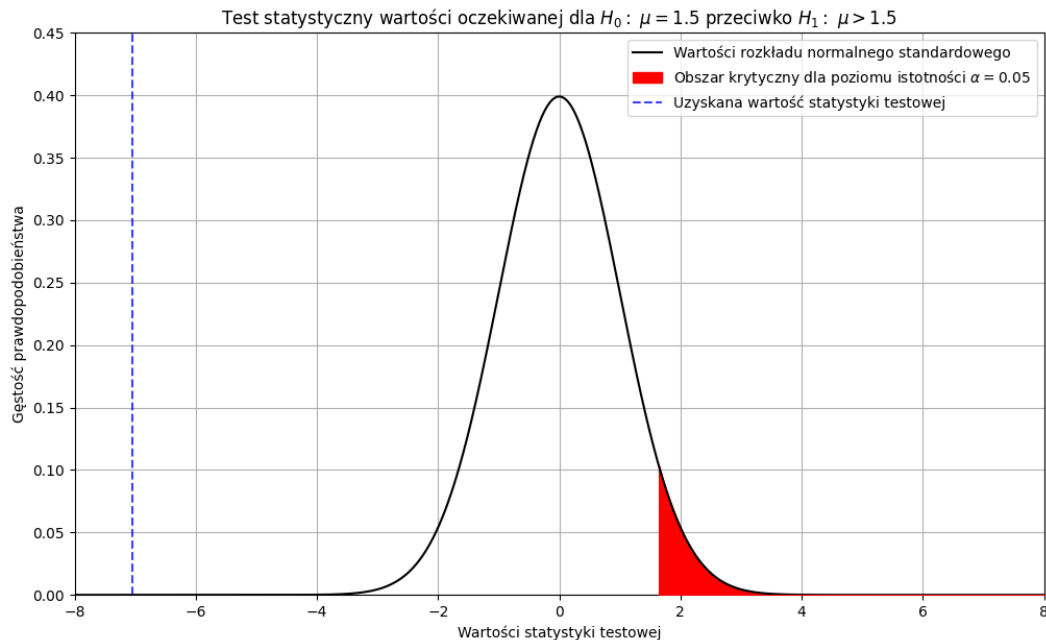


Rysunek 2: Wykres obszaru krytycznego dla hipotezy alternatywnej $\mu < 1,5$

2.3. $H_1 : \mu > 1,5$

- Obszar krytyczny $C = (1,64; \infty)$.
- Wartość statystyki testowej $Z \notin C$.
- $p_{\text{wartość}} \approx 1 > \alpha$.

Wartość statystyki testowej nie należy do zbioru krytycznego, dlatego nie mamy podstaw do odrzucenia hipotezy zerowej. Nie można stwierdzić, że średnia jest większa niż 1,5.



Rysunek 3: Wykres obszaru krytycznego dla hipotezy alternatywnej $\mu > 1,5$

2.4. Zmiana parametru α

Im większy poziom istotności α , tym większy jest obszar krytyczny. Co za tym idzie, dla większych α częściej będziemy odrzucać hipotezę zerową na rzecz hipotezy alternatywnej, natomiast dla małych α rzadziej będziemy ją odrzucać. Wyniki dla przykładowych parametrów α możemy zobaczyć w poniższej tabeli.

	$\alpha = 0,1$	$\alpha = 0,05$	$\alpha = 0,01$
$H_1: \mu \neq 1,5$	$(-\infty; -1,64) \cup (1,64; \infty)$	$(-\infty; -1,96) \cup (1,96; \infty)$	$(-\infty; -2,58) \cup (2,58; \infty)$
$H_1: \mu < 1,5$	$(-\infty; -1,28)$	$(-\infty; -1,64)$	$(-\infty; -2,33)$
$H_1: \mu > 1,5$	$(1,28; \infty)$	$(2,33; \infty)$	$(1056,7; \infty)$

Tabela 1: Obszary krytyczne dla różnych hipotez alternatywnych i wartości poziomu istotności α ; na zielono przypadki, gdy przyjmujemy hipotezę zerową, a na czerwono gdy ją odrzucamy na rzecz hipotezy alternatywnej

3. Testy dla wariancji

Na podstawie danych z próby wyznaczono wartość statystyki testowej:

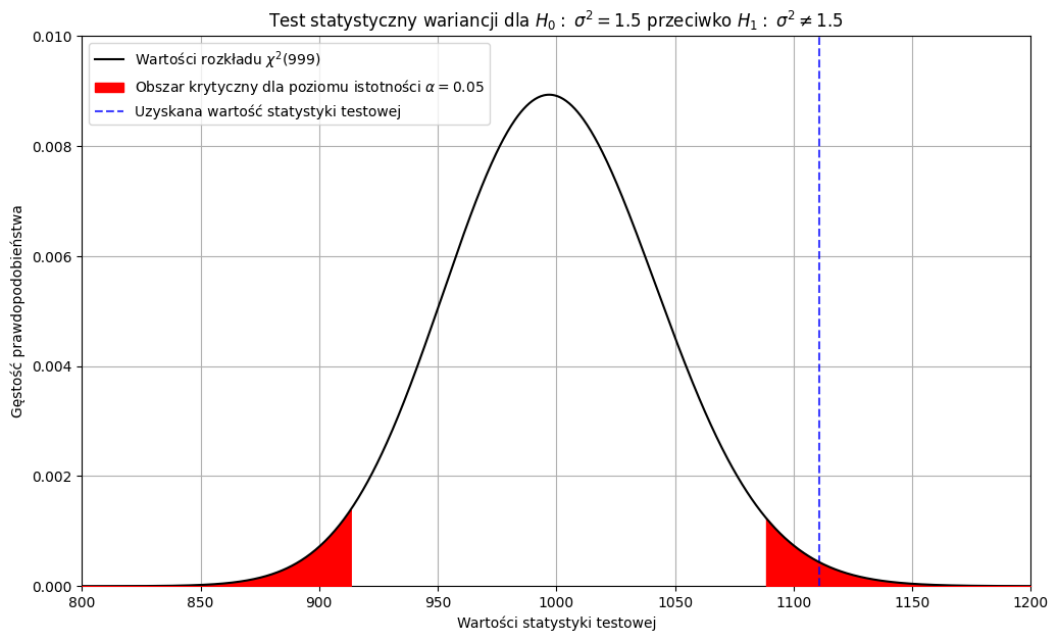
$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} \approx 1110,97.$$

Przyjęta została hipoteza zerowa $H_0 : \sigma_0^2 = 1,5$. W dalszej części przeanalizowano trzy warianty hipotezy alternatywnej.

3.1. $H_1 : \sigma^2 \neq 1,5$

- Obszar krytyczny $C = (-\infty; 913,3) \cup (1088,5; \infty)$.
- Wartość statystyki testowej $\chi^2 \in C$.
- $p_{\text{wartość}} = 0,015 < \alpha$.

Wartość statystyki testowej należy do zbioru krytycznego, dlatego odrzucamy hipotezę zerową na korzyść hipotezy alternatywnej. Wariancja znacząco różni się od 1,5.

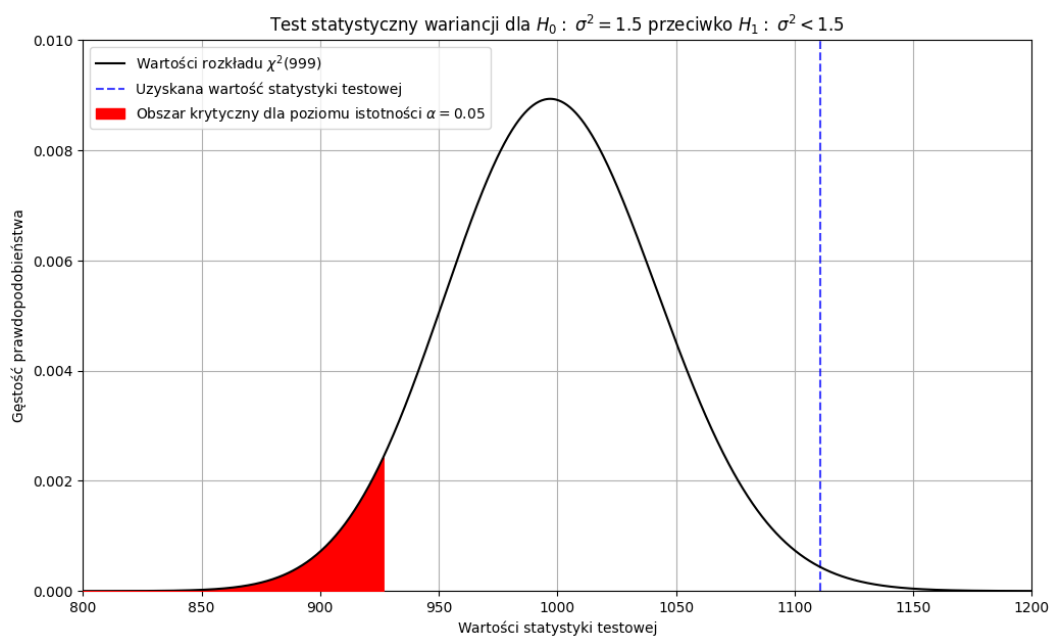


Rysunek 4: Wykres obszaru krytycznego dla hipotezy alternatywnej $\sigma^2 \neq 1,5$

3.2. $H_1 : \sigma^2 < 1,5$

- Obszar krytyczny $C = (-\infty; 926,6)$.
- Wartość statystyki testowej $\chi^2 \notin C$.
- $p_{\text{wartość}} = 0,992 > \alpha$.

Wartość statystyki testowej nie należy do zbioru krytycznego, dlatego nie mamy podstaw do odrzucenia hipotezy zerowej. Nie można stwierdzić, że wariancja jest mniejsza od 1,5.

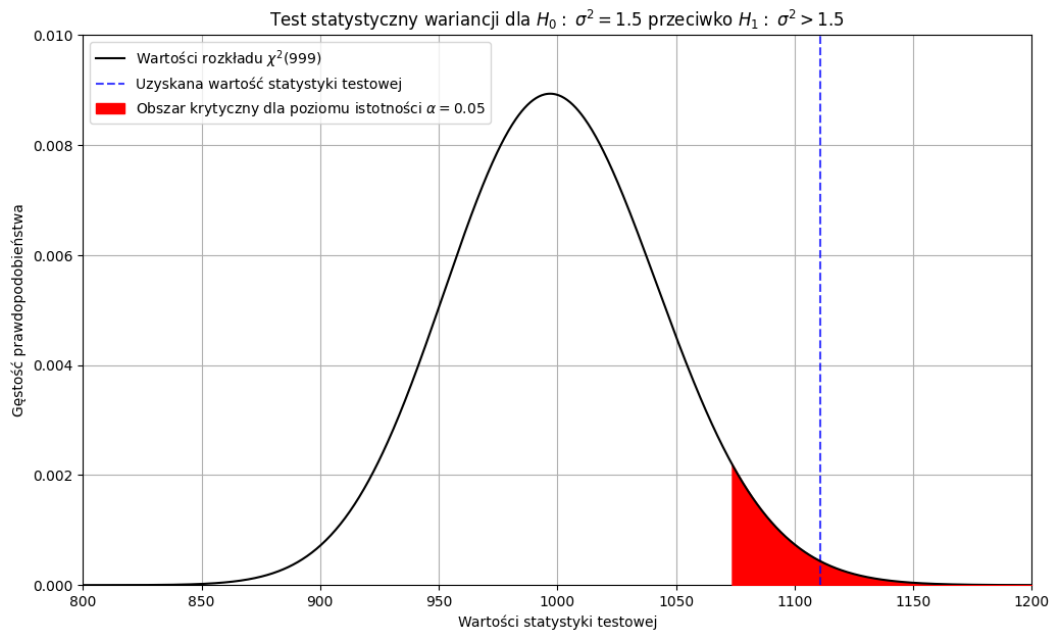


Rysunek 5: Wykres obszaru krytycznego dla hipotezy alternatywnej $\sigma^2 < 1,5$

3.3. $H_1 : \sigma^2 > 1,5$

- Obszar krytyczny $C = (1073,6; \infty)$.
- Wartość statystyki testowej $\chi^2 \in C$.
- $p_{\text{wartość}} = 0,008 < \alpha$.

Wartość statystyki testowej należy do zbioru krytycznego, dlatego odrzucamy hipotezę zerową na korzyść hipotezy alternatywnej. Średnia populacji jest większa od 1,5.



Rysunek 6: Wykres obszaru krytycznego dla hipotezy alternatywnej $\sigma^2 > 1,5$

3.4. Zmiana parametru α

Tak jak w poprzednim zadaniu im większy poziom istotności α , tym większy jest obszar krytyczny. Co za tym idzie, dla większych α częściej będziemy odrzucać hipotezę zerową na rzecz hipotezy alternatywnej, natomiast dla małych α rzadziej będziemy ją odrzucać. Wyniki dla przykładowych parametrów α możemy zobaczyć w poniższej tabeli.

	$\alpha = 0,1$	$\alpha = 0,05$	$\alpha = 0,01$
$H_1 : \sigma^2 \neq 1,5$	$(-\infty; 926,6) \cup (1073,6; \infty)$	$(-\infty; 913,3) \cup (1088,5; \infty)$	$(-\infty; 887,6) \cup (1117,9; \infty)$
$H_1 : \sigma^2 < 1,5$	$(-\infty; 942,2)$	$(-\infty; 926,6)$	$(-\infty; 898)$
$H_1 : \sigma^2 > 1,5$	$(1056,7; \infty)$	$(1073,6; \infty)$	$(1105,9; \infty)$

Tabela 2: Obszary krytyczne dla różnych hipotez alternatywnych i wartości poziomu istotności α ; na zielono przypadki, gdy przyjmujemy hipotezę zerową, a na czerwono gdy ją odrzucamy na rzecz hipotezy alternatywnej

4. Błędy statystyczne

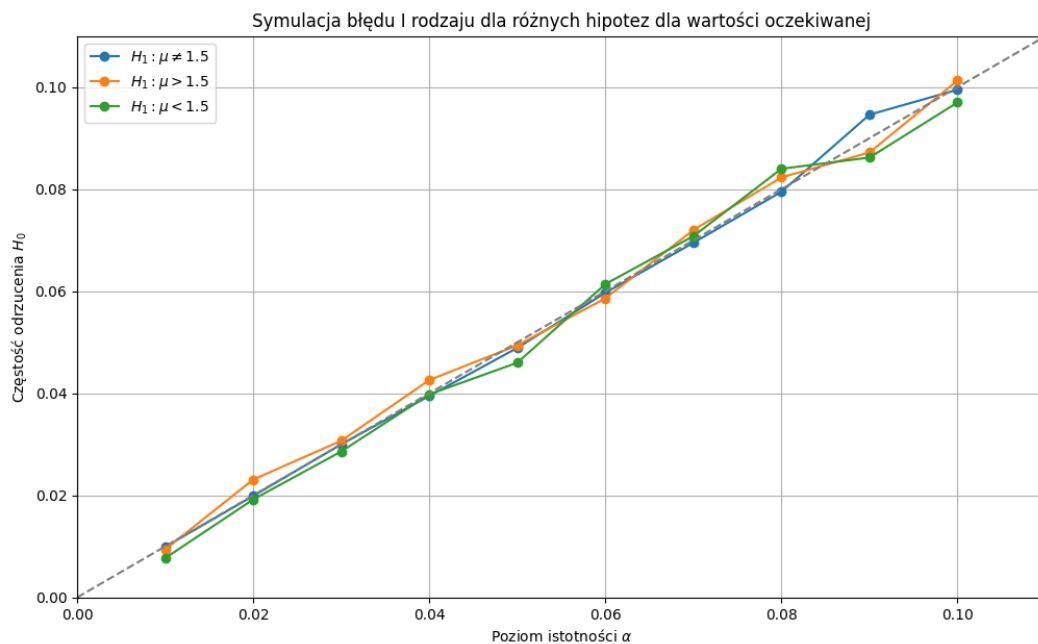
W statystycznym wnioskowaniu istotnym elementem analizy jest ocena ryzyka popełnienia błędów przy podejmowaniu decyzji na podstawie próby. W dalszej części projektu przeprowadzono symulacje mające na celu oszacowanie empirycznych wartości błędów I i II rodzaju oraz mocy testu dla różnych scenariuszy testowania hipotez dotyczących wartości średniej oraz wariancji rozkładu normalnego.

4.1. Błąd I rodzaju

Błąd I rodzaju polega na odrzuceniu hipotezy zerowej H_0 w sytuacji, gdy jest ona prawdziwa. Prawdopodobieństwo popełnienia tego błędu równe jest przyjętemu poziomowi istotności α . W celu sprawdzenia, czy rzeczywiście zachodzi taka zależność, przeprowadzono symulację Monte Carlo.

4.1.1. Test dla wartości średniej

Dla każdej z hipotez alternatywnych przeprowadzono $N = 10000$ niezależnych prób o liczności $n = 1000$ z rozkładu normalnego $\mathcal{N}(1, 5; 0, 2^2)$.



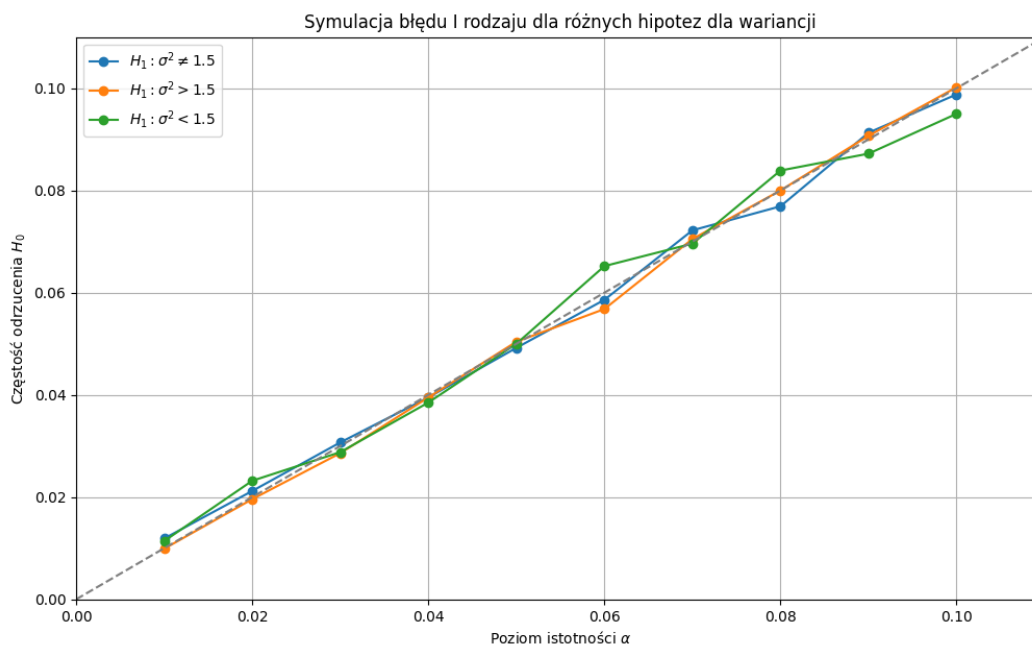
Rysunek 7: Wykres prawdopodobieństwa błędu I rodzaju dla wartości średniej

Na wykresie przedstawiono zależność między wartością poziomu istotności α a

rzeczywistą częstością odrzuceń H_0 w symulacjach. Widać, że dla każdego rodzaju hipotezy krzywe przebiegają bardzo blisko prostej $y = x$, co potwierdza zgodność z teoretycznymi założeniami.

4.1.2. Test dla wariancji

Analogiczne podejście można zastosować do testowania hipotez dotyczących wariancji. W tym przypadku test opiera się na rozkładzie chi-kwadrat. Tutaj również dla każdej z hipotez alternatywnych przeprowadzono $N = 10000$ niezależnych prób o liczności $n = 1000$, jednak tym razem z rozkładu normalnego $\mathcal{N}(0, 2; 1, 5)$.



Rysunek 8: Wykres prawdopodobieństwa błęd I rodzaju dla wariancji

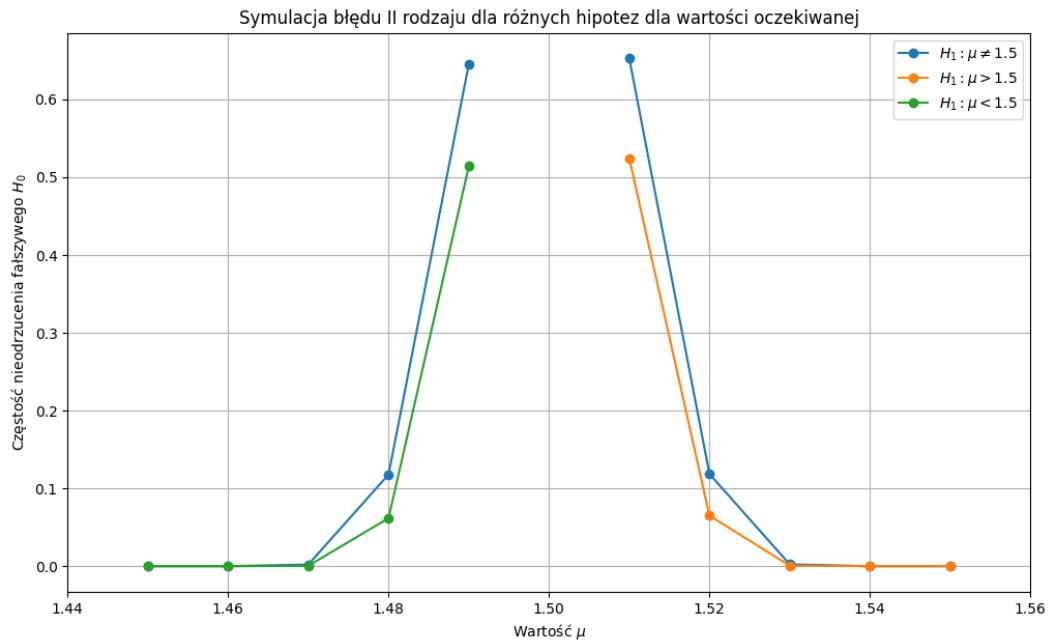
Także na tym wykresie widać liniową zależność między wartością poziomu istotności α a rzeczywistą częstością odrzuceń H_0 w symulacjach. Zgodność wyników z wartościami teoretycznymi wskazuje na poprawność implementacji i trafność przyjętych założeń.

4.2. Błąd II rodzaju

Błąd II rodzaju polega na nieodrzućeniu hipotezy zerowej H_0 w sytuacji, gdy jest ona fałszywa. Prawdopodobieństwo popełnienia tego błęd równie jest 1 - moc testu. W celu oszacowania prawdopodobieństwa błęd II rodzaju przeprowadzono symulację Monte Carlo.

4.2.1. Test dla wartości średniej

Dla każdej z hipotez alternatywnych przeprowadzono $N = 10000$ niezależnych prób o liczności $n = 1000$ z rozkładu normalnego $\mathcal{N}(\mu; 0, 2^2)$ dla różnych wartości μ zbliżonych do 1,5, przy poziomie istotności $\alpha = 0,05$

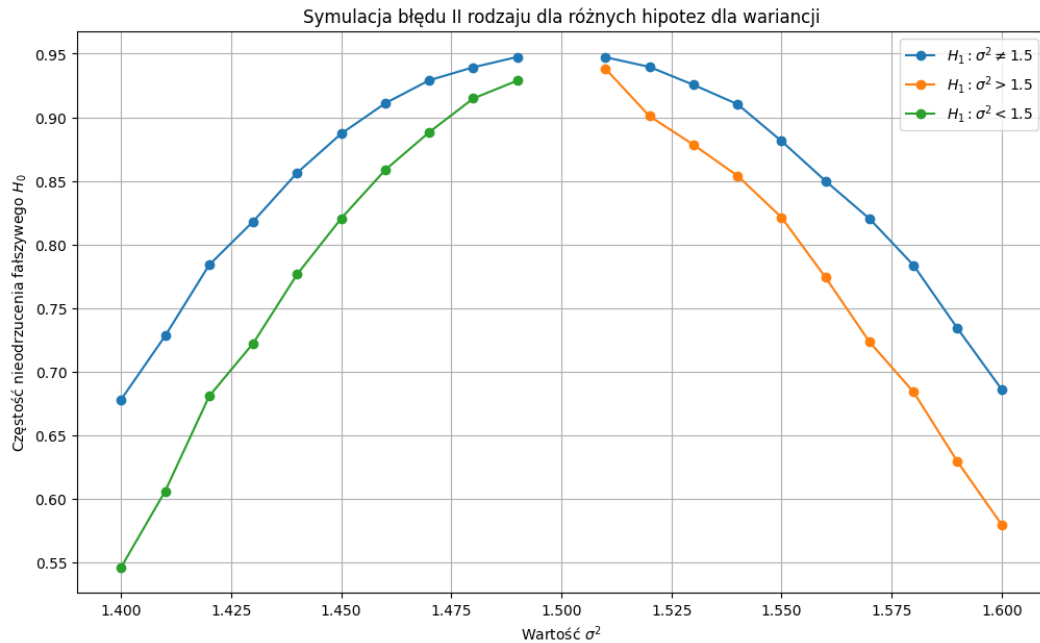


Rysunek 9: Wykres prawdopodobieństwa błęd II rodzaju dla wartości oczekiwanej

Jak widzimy, prawdopodobieństwo popełnienia błęd drugiego rodzaju jest bliskie zero już dla wartości różniących się o zaledwie 0,02 od zakładanej wartości $\mu = 1,5$.

4.2.2. Test dla wariancji

Analogiczne podejście można zastosować do testowania hipotez dotyczących wariancji. W tym przypadku test opiera się na rozkładzie chi-kwadrat. Tutaj również dla każdej z hipotez alternatywnych przeprowadzono $N = 10000$ niezależnych prób o liczności $n = 1000$, jednak tym razem z rozkładu normalnego $\mathcal{N}(0, 2; \sigma^2)$ dla różnych wartości σ^2 zbliżonych do 1,5, przy poziomie istotności $\alpha = 0,05$.



Rysunek 10: Wykres prawdopodobieństwa błędu II rodzaju dla wariancji

Widzimy, że prawdopodobieństwo popełnienia błędu drugiego rodzaju wraz z oddalaniem się od założonej wartości $\sigma^2 = 1,5$ spada dość powoli. Jest to spadek o wiele wolniejszy niż dla wartości oczekiwanej i nawet gdy oddalimy się o 0,1 od zakładanej wartości $\sigma^2 = 1,5$, prawdopodobieństwo popełnienia błędu drugiego rodzaju jest większe od $1/2$.

4.3. Moc testu

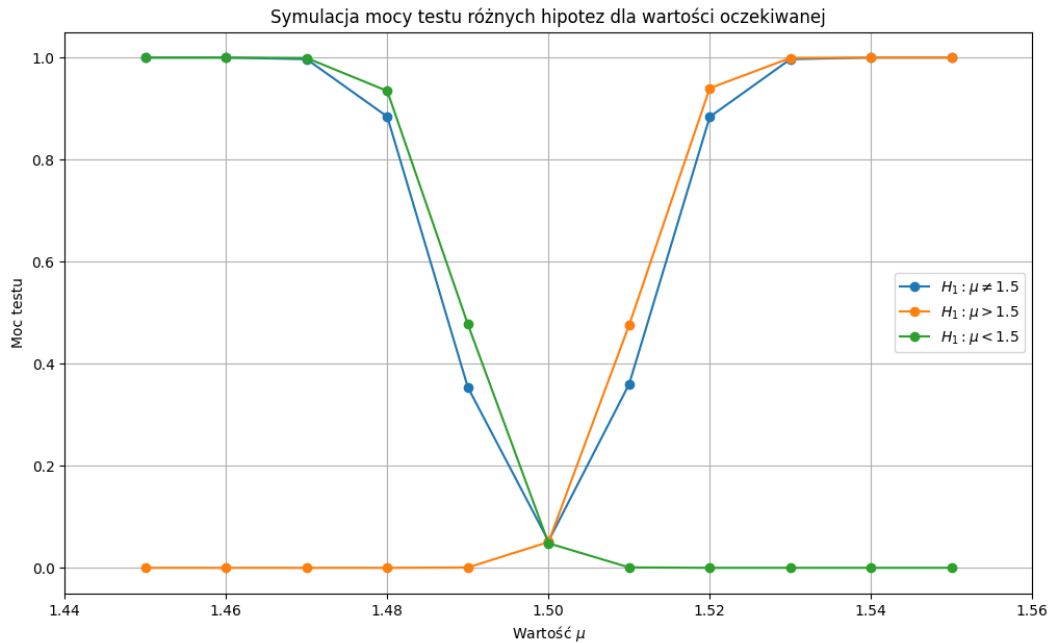
Moc testu to prawdopodobieństwo odrzucenia fałszywej hipotezy zerowej dla zadanej wartości badanego parametru. Prawdziwa jest własność:

moc testu = 1 - prawdopodobieństwo błędu II rodzaju.

W celu ustalenia mocy testu dla różnych wartości hipotez alternatywnych przeprowadzono symulacje Monte Carlo.

4.3.1. Test dla wartości średniej

Dla każdej z hipotez alternatywnych przeprowadzono $N = 10000$ niezależnych prób o licznosci $n = 1000$ z rozkładu normalnego $\mathcal{N}(\mu; 0, 2^2)$ dla różnych wartości μ zbliżonych do 1,5, przy poziomie istotności $\alpha = 0,05$

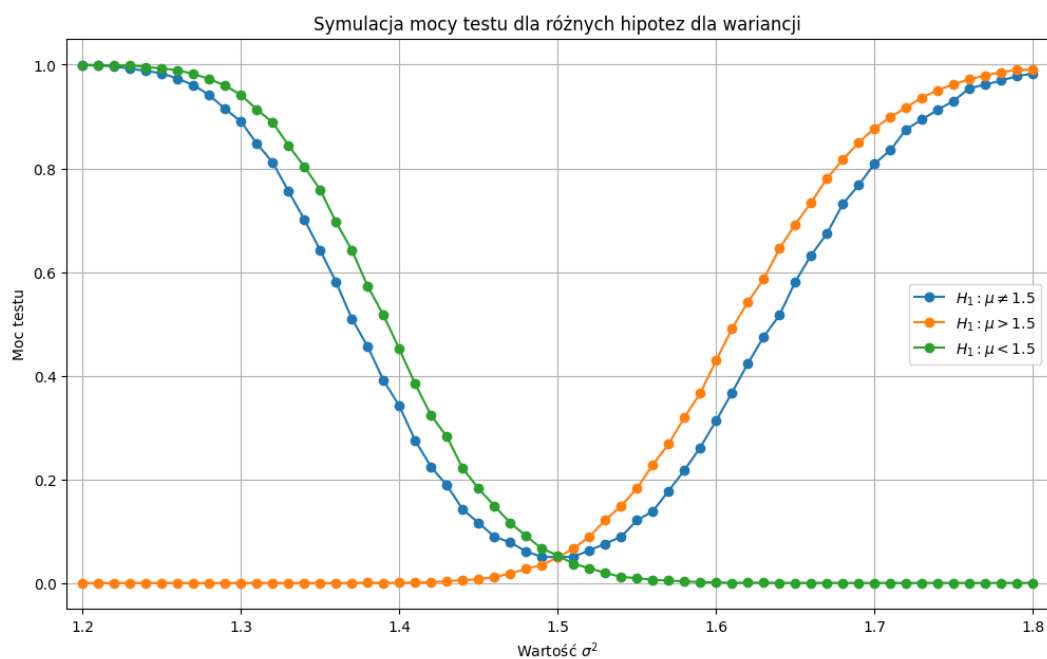


Rysunek 11: Wykres mocy testu dla wartości oczekiwanej

Jak widzimy, moc testu z dwustronną hipotezą alternatywną rośnie, gdy z zadanym parametrem μ oddalamy się od wartości 1,5 w dowolnym kierunku. Bardzo szybko wraz z oddalaniem się osiąga ona wartości zbliżone do 1. Dla hipotez jednostronnych w jednym kierunku moc szybko osiąga wartość 1, a w drugim 0. Można stwierdzić, że ogólnie przeprowadzony test ma dużą moc.

4.3.2. Test dla wartości oczekiwanej

Analogiczne podejście można zastosować do testowania hipotez dotyczących wariancji. W tym przypadku test opiera się na rozkładzie chi-kwadrat. Tutaj również dla każdej z hipotez alternatywnych przeprowadzono $N = 10000$ niezależnych prób o liczności $n = 1000$, jednak tym razem z rozkładu normalnego $\mathcal{N}(0, 2; \sigma^2)$ dla różnych wartości σ^2 zbliżonych do 1,5, przy poziomie istotności $\alpha = 0,05$.



Rysunek 12: Wykres mocy testu dla wartości oczekiwanej

Sytuacja jest podobna do wcześniejszej, jednak moc testu dużo wolniej się zwiększa przy oddalaniu się od $\sigma^2 = 1, 5$. Można wnioskować, że testy na wariancję mają ogólnie mniejszą moc niż testy na średnią.