Testowanie hipotez

Autorzy

Maciej Karczewski 262282
Karolina Wypych 262333

Politechnika Wrocławska

Wydział Matematyki

16 czerwca 2022r.

Spis treści

1 Wprowadzenie

Dowolne przypuszczenie dotyczące rozkładu populacji i wartości parametru rozkładu nazywamy *hipotezą statystyczną*. Dwie hipotezy komplementarne w problemie testowania hipotez to odpowiednio *hipoteza zerowa* H_0 i *hipoteza alternatywna* H_1. Procedura testowania hipotezy statystycznej polega na znalezieniu reguły pozwalającej na podjęcie decyzji, dla jakich wartości próby przyjąć hipotezę zerową, a dla jakich odrzucić ją i za prawdziwą uznać hipotezę alternatywną. Niezbędne jest przy tym posługiwanie się następującymi pojęciami:

1. C- zbiór krytyczny, czyli zbiór wartośći statystyki testowej prowadzących do odrzucenia hipotezy zerowej H_0 na korzyść H_1,
2. C' - zbiór przyjęć hipotezy H_0,
3. wartości krytyczne testu, czyli wartości brzegowe C graniczące z C'.

2 Potrzebne definicje

2.1 Błąd I rodzaju

Błąd I rodzaju to odrzucenie prawdziwej hipotezy zerowej H_0. Prawdopodobieństwo popełnienia błędu I rodzaju nazywamy poziomem istotności testu i oznaczamy α.

2.2 Błąd II rodzaju

Błąd II rodzaju jest błędem polegającym na przyjęciu hipotezy zerowej, która jest nie prawdziwa zwyczajowo oznaczanym jako β. Zmiejszamy ten błąd zwiększając poziom istotności testu α.

2.3 Moc testu

Moc testu to prawdopodobieństwo uniknięcia błędu II rodzaju i wyraża się wzorem $1 - \beta$. Im większa moc testu tym test jest lepszy jako narzędzie do różnicowania między hipotezą prawdziwą a fałszywą. Moc testu zależy od

1. liczebności próby,
2. rzeczywistej wartości badanego parametru,
3. przyjętego poziomu instotności α.

2.4 p-wartość

Najmniejszy poziom istotności α, przy którym zaobserwowane wartości statystyki testowej prowadzą do odrzucenia hipotezy zerowej nazywamy p-wartością. W praktyce gdy:

- p-wartość < 0.01 prawie zawsze odrzucamy hipotezę zerową,
- $0.01 <$ p-wartość < 0.1 decyzja zależy od preferencji,
- p-wartość > 0.1 nie odrzucamy hipotezy zerowej.

W zależności od postaci hipotezy alternatywnej, $p - warto$ść obliczamy przy pomocy odpowiednich wzorów:

1. Dla $H_1 : \theta \neq \theta_0$ p-wartość $= 2P_{H_O}(Z \geqslant |z|) = 2 - 2\Phi(|Z|)$
2. Dla $H_2 : \theta > \theta_0$ p-wartość $= P_{H_O}(Z \geqslant z) = 1 - \Phi(Z)$
3. Dla $H_3 : \theta < \theta_0$ p-wartość $= P_{H_O}(Z \leqslant z) = \Phi(Z)$

3 Testy dla wartości średniej w rodzinie rozkładów normalnych- przypadek znanej wariancji.

Weryfikując hipotezy statystyczne dotyczące wartości oczekiwanej w próbie o rozkładzie z rodziny normalnych przy znanym odchyleniu standardowym, posługujemy się statystyką testową Z, którą wyznacza się przez standaryzację średniej próbkowej $\bar{X}$.
Jeśli zachodzi hipoteza zerowa H_0 to $\bar{X} \sim N(\mu_0, \frac{\sigma}{\sqrt{n}})$. Wówczas wspomnianą wcześniej statystykę testową wyznacza się ze wzoru

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}, \tag{1}$$

gdzie:
μ_0 - wartość średnia zgodna z hipotezą zerową,
σ - znane odchylenie standardowe,
n- długość próby.
Taka statystyka zbiega do rozkładu standardowego normalnego $N(0,1)$. Zatem, gdy H_0 jest fałszywa, Z powinna przyjmować duże wartości (rozbieżność między średnią wyznaczoną z próbki $\bar{X}$ a wartością hipotezy zerowej μ_0, którą testujemy). Statystykę Z można zapisać następująco

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} + \frac{\mu - \mu_0}{\frac{\sigma}{\sqrt{n}}}, \tag{2}$$

gdzie μ jest prawdziwą wartością parametru średniej rozkładu, z którego pochodzi badana próbka, a niebieska część zbiega do rozkładu standardowego normalnego $N(0,1)$.
Ponieważ określenie „duże” jest pojęciem względnym i niewiele mówiącym, doprecyzujemy wartości statystyki testowej, dla których odrzucamy statystykę zerową. W zależności od postaci statystyki zerowej, przedział krytyczny przujmuje różną postać.

1. Dla $H_1 : \mu \neq \mu_0$ wynosi on $C = (-\infty, -z_{1-\frac{\alpha}{2}}] \cup [z_{1-\frac{\alpha}{2}}, \infty)$
2. Dla $H_2 : \mu > \mu_0$ wynosi on $C = [z_{1-\alpha}, \infty)$
3. Dla $H_3 : \mu < \mu_0$ wynosi on $C = (-\infty, -z_{1-\alpha}]$,

gdzie z_β oznacza kwantyl rzędu β rozkładu $N(0,1)$.

4 Zadanie 1.

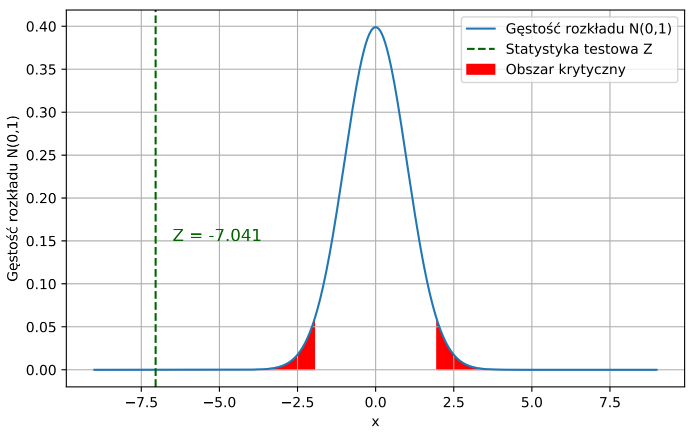
W tym zadaniu pracujemy na danych ze strony domowej dr hab. inż. Agnieszki Wyłomańskiej, będących próbą z rozkładu $N(\mu, \sigma = 0.2)$. Na poziomie istotności $\alpha = 0.05$ testujemy hipotezę zerową $\mu_0 = 1.5$ przeciwko trzem hipotezom alternatywnym:

1. $H_1 : \mu \neq 1.5$
2. $H_2 : \mu > 1.5$
3. $H_3 : \mu < 1.5$

Dla wszystkich powyższych hipotez statystyka testowa wynosi $Z = -7.041$.

4.1 $H_1 : \mu \neq 1.5$

Przedział krytyczny dla tego przypadku $C = (-\infty, -1.960] \cup [1.960, \infty)$.



Rysunek 1: Rozkład statystyki Z pod warunkiem hipotezy zerowej, wraz z obszarem krytycznym dla $\alpha = 0.05$ i wartością rzeczywistą statystyki dla hipotezy alternatywnej $\mu \neq 1.5$

Ponieważ $Z \in C$, odrzucamy hipotezę zerową H_0 na poziomie istotności $\alpha = 0.05$ i przyjmujemy hipotezę alternatywną H_1. W tej sytuacji Z ma rozkład normalny o średniej $\frac{\mu-\mu_0}{\frac{\sigma}{\sqrt{n}}}$ i odchyleniu standardowym równym 1.

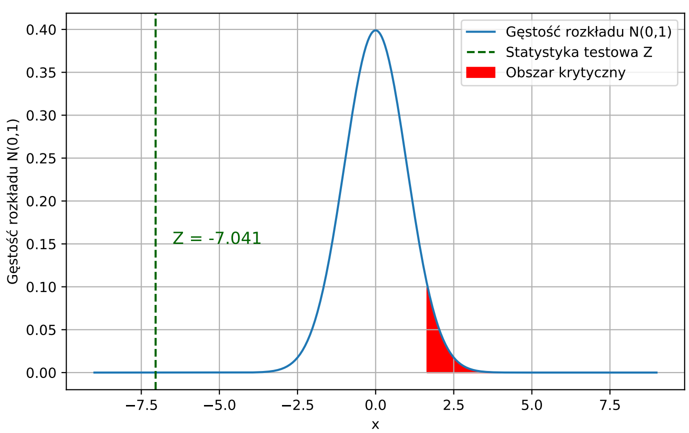
Gdybyśmy rozważali niższy poziom istotności, np $\alpha = 0.01$, zawężybyliśmy zbiór krytyczny do $C_{0.01} = (-\infty, -2.576] \cup [2.576, \infty)$.

Natomiast poprzez podniesienie poziomu istoności, np. do $\alpha = 0.1$, rozszerzylibyśmy zbiór krytyczny do $C_{0.1} = (-\infty, -1.645] \cup [1.645, \infty)$.

Ponieważ obniżanie poziomu istotności zmniejsza przedział krytyczny, domyślamy się, że dla odpowiednio małego α hipoteza H_0 zostanie przyjęta. Wystarczy wyznaczyć p-wartość, a dla każdego α takiego, że $\alpha < p$-wartość, hipoteza zerowa będzie akceptowana. Dla rozważanej hipotezy alternatywnej otrzymujemy p-wartość $= 1.902 \cdot 10^{-12}$. Co w praktyce oznacza, że zawsze będziemy odrzucać hipotezę zerową.

4.2 $H_2 : \mu > 1.5$

Tym razem obszar krytyczny równy jest $C = [1.645, \infty)$.



Rysunek 2: Rozkład statystyki Z pod warunkiem hipotezy zerowej, wraz z obszarem krytycznym dla $\alpha = 0.05$ i wartością rzeczywistą statystyki dla hipotezy alternatywnej $\mu > 1.5$

Ponieważ $Z \notin C$, przyjmujemy hipotezę zerową H_0 na poziomie istotności $\alpha = 0.05$.

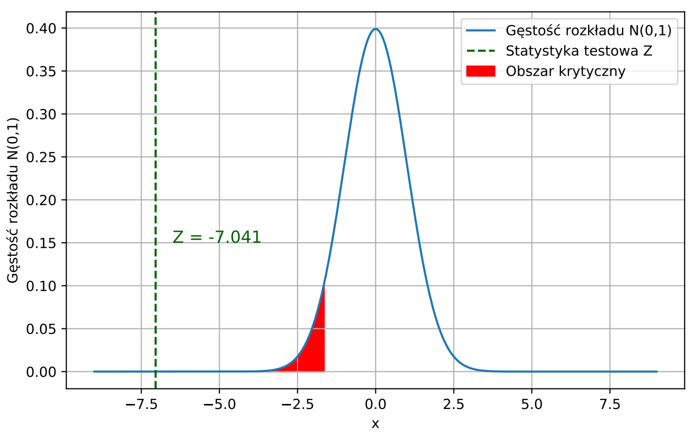
Tutaj również sprawdzamy, jak zmienią się zbiory krytyczne w zależności od przyjętego poziomu istotności. Tym razem otrzymujemy

1. $C_{0.01} = [2.326, \infty)$,
2. $C_{0.1} = [1.282, \infty)$,

więc dla każdego poziomu przyjmujemy H_0. Odrzucić ją moglibyśmy dopiero dla $\alpha >$ p-wartość, a w tym przypadku p-wartość $= 0.9999999999990488 \approx 1$.Co w praktyce oznacza, że zawsze będziemy przyjmować hipotezę zerową.

4.3 $H_3 : \mu < 1.5$

Zbiór krtytyczny testu dla tej alternatywy ma postać $C = (-\infty, -1.645]$



Rysunek 3: Rozkład statystyki Z pod warunkiem hipotezy zerowej, wraz z obszarem krytycznym dla $\alpha = 0.05$ i wartością rzeczywistą statystyki dla hipotezy alternatywnej $\mu < 1.5$

Ponieważ $Z \in C$, odrzucamy hipotezę zerową H_0 na poziomie istotności $\alpha = 0.05$ i przyjmujemy hipotezę alternatywną H_3. W tej sytuacji Z ma taki sam rozkład jak w 4.1. Wyznaczamy także zbiory krytyczne dla innych poziomów istotności. Tym razem otrzymujemy

1. $C_{0.01} = (-\infty, -2.326]$,
2. $C_{0.1} = (-\infty, -1.282]$,

więc dla każdego z powyższych poziomów odrzucamy H_0. Przyjąć ją moglibyśmy dopiero dla $\alpha <$ p-wartość $= 9.512 \cdot 10^{-13}$. Co w praktyce oznacza, że zawsze będziemy odrzucać hipotezę zerową.

Biorąc pod uwagę wyniki testów przeprowadzonych dla powyższych trzech hipotez, można wysnuć wniosek, że badana próba pochodzi z rozkładu normalnego $N(\mu, 0.2)$, gdzie $\mu < 1.5$.

5 Testy dla wariancji w rodzinie rozkładów normalnych.

Dla testowania hipotez w rodzinie rozkładów normalnych,w celu zbadania wariancji, używamy statystyki χ^2. Statystykę tę obliczmy następującym wzorem

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

gdzie,
$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(\bar{X} - x_i\right)^2$ jest estymatorem wariancji,
σ_0^2 jest wariancją występującą w hipotezie zerowej.

Jeśli hipoteza zerowa jest prawdziwa to statystka χ^2 ma rozkład χ^2 z $(n-1)$ stopniami swobody. Następnie sprawdzamy, czy nasza statystyka przyjmuje nietypowe wartości, co oznacza, że znajduje się w obszarze krytycznym, w którym odrzucamy hipotezę zerową, czyli z prawdopodobieństwem $1-\alpha$, nie jest to rozkład χ^2 z $(n-1)$ stopniami swobody, co implikuje fałszywość hipotezy zerowej.

Przdziały krytyczne w zależności od hipotezy alternatywnej:

- $H_1 : \sigma^2 \neq \sigma_0^2,$ to zbiór krytyczny ma postać $(-\infty, \chi^2_{\frac{\alpha}{2},n-1}] \cup [\chi^2_{1-\frac{\alpha}{2},n-1}, \infty)$
- $H_2 : \sigma^2 > \sigma_0^2,$ to zbiór krytyczny ma postać $[\chi^2_{1-\alpha,n-1}, \infty)$
- $H_3 : \sigma^2 < \sigma_0^2,$ to zbiór krytyczny ma postać $(-\infty, \chi^2_{\alpha,n-1}]$

gdzie $\chi^2_{\alpha,n-1}$ oznacza kwantyl rzędu α rozkładu χ^2 o $(n-1)$ stopniach swobody.

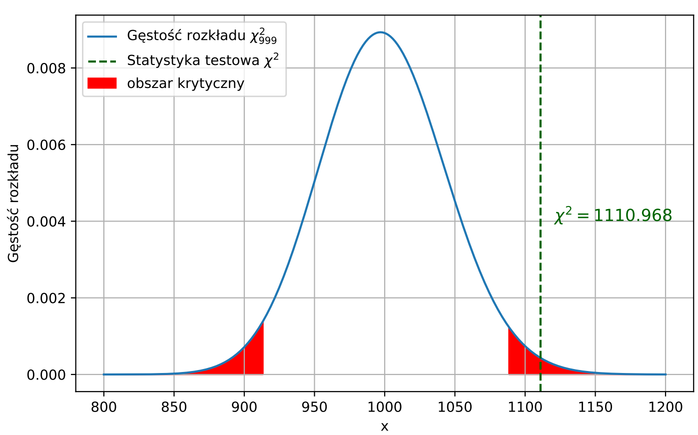
6 Zadanie 2.

W tym zadaniu pracujemy na danych ze strony domowej dr hab. inż. Agnieszki Wyłomańskiej. Badamy hipotezę zerową $\sigma^2 = 1.5$ na poziomie istotności $\alpha = 0.05$. Wiemy, że dane pochodzą z rozkładu $N(0.2, \sigma^2)$ o nieznanej sigmie. Do dyspozycji mamy 1000 obserwacji. Niezależnie od hipotezy alternatywnej korzystamy z statystyki χ^2, która dla naszych danych wynosi 1110.968. Do oceny czy akceptujemy czy odrzucamy hipotezę zerową musimy zdefiniować hipotezę alternatywną, która ma następujące postacie

1. $H_1 : \sigma^2 \neq 1.5$
2. $H_2 : \sigma^2 > 1.5$
3. $H_3 : \sigma^2 < 1.5$

6.1 $H_1 : \sigma^2 \neq 1.5$

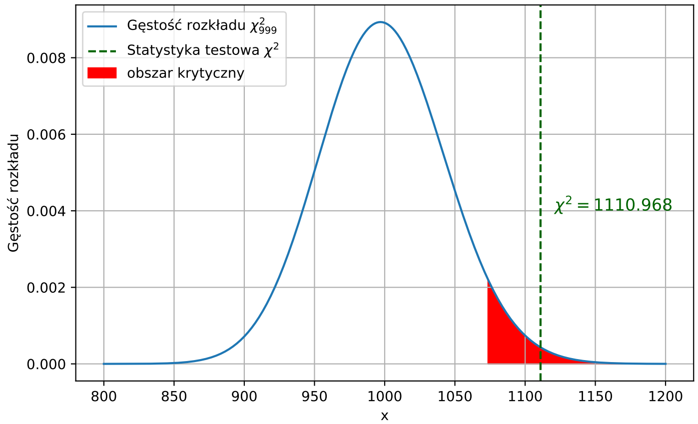
W tym przypadku przedział akceptacji hipotezy zerowej dla $\alpha = 0.05$ ma postać $(\chi^2_{0.025,999}; \chi^2_{0.975,999}) = [913.301; 1088.487]$, czyli nasza statystyka wynosząca 1110.968 nie mieści się w zbiorze akceptacji hipotezy zerowej. Oznacza to, że jest w zbiorze krytycznym, więc musimy ją odrzucić. Obliczamy p-wartość która wynosi 0.015, czyli dla $\alpha > 0.015$ będziemy odrzucać hipotezę zerową, natomiast dla $\alpha < 0.015$ będziemy ją przyjmować.



Rysunek 4: Rozkład statystyki χ^2 pod warunkiem hipotezy zerowej, wraz z obszarem krytycznym dla $\alpha = 0.05$ jak i wartością rzeczywistą statystyki dla hipotezy alternatywnej $\sigma^2 \neq 1.5$

6.2 $H_2 : \sigma^2 > 1.5$

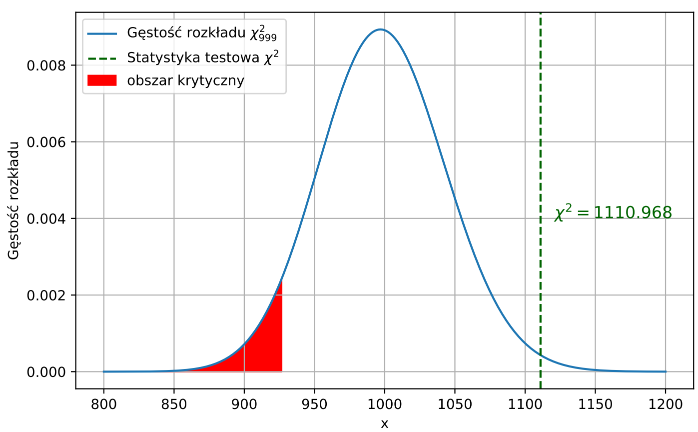
Tym razem przedział krytyczny dla $\alpha = 0.05$ ma postać $[\chi^2_{0.95,999};\infty) = [1073.643;\infty)$, czyli nasza statystyka wynosząca 1110.968 należy do zbioru krytycznego, co oznacza, że hipotezę zerową musimy odrzucić. Obliczamy p-wartość, która wynosi 0.0075, czyli dla $\alpha > 0.0075$ będziemy odrzucać hipotezę zerową, natomiast dla $\alpha < 0.0075$ będziemy ją przyjmować.



Rysunek 5: Rozkład statystyki χ^2 pod warunkiem hipotezy zerowej, wraz z obszarem krytycznym dla $\alpha = 0.05$ jak i wartością rzeczywistą statystyki dla hipotezy alternatywnej $\sigma^2 > 1.5$

6.3 $H_3 : \sigma^2 < 1.5$

W tym przypadku przedział krytyczny dla $\alpha = 0.05$ ma postać $(\infty;\chi^2_{0.05,999}] = (-\infty; 926.631]$ czyli nasza statystyka wynosząca 1110.968 nie należy do zbioru krytycznego, więc hipotezę zerową przyjmujemy. Obliczamy p-wartość która wynosi 0.992 co oznacza, że dla $\alpha > 0.992$ będziemy odrzucać hipotezę zerową, natomiast dla $\alpha < 0.992$ będziemy ją przyjmować. Patrząc na to, że nasza p-wartość jest bardzo duża, możemy wykluczyć hipotezę alternatywną.



Rysunek 6: Rozkład statystyki χ^2 pod warunkiem hipotezy zerowej, wraz z obszarem krytycznym dla $\alpha = 0.05$ jak i wartością rzeczywistą statystyki dla hipotezy alternatywnej $\sigma^2 < 1.5$

Biorąc pod uwagę wyniki testów przeprowadzonych dla powyższych trzech hipotez, można wysnuć wniosek, że badana próba pochodzi z rozkładu normalnego $N(0.2, \sigma)$, gdzie $\sigma > 1.5$.

7 Zadanie 3.

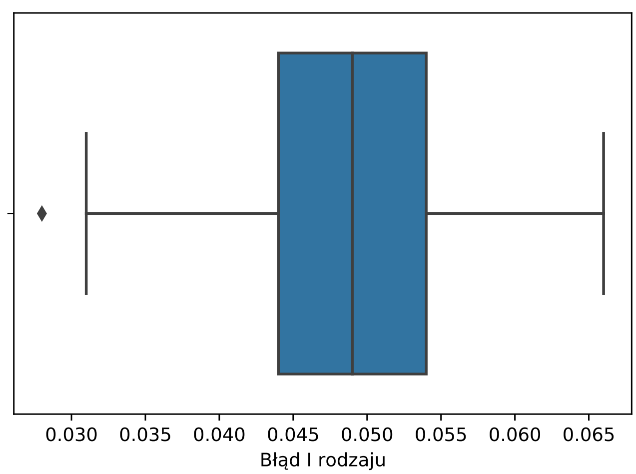
7.1 Błąd I rodzaju

Aby wyznaczyć symulacyjnie błąd I rodzaju musimy wygenerować prostą próbę losową z rozkładu normalnego o parametrach zgodnych z H_0 ($\mu = 1.5$ oraz $\sigma = 0.2$) i sprawdzić, ile razy odrzucimy hipotezę zerową.
Algorytm:

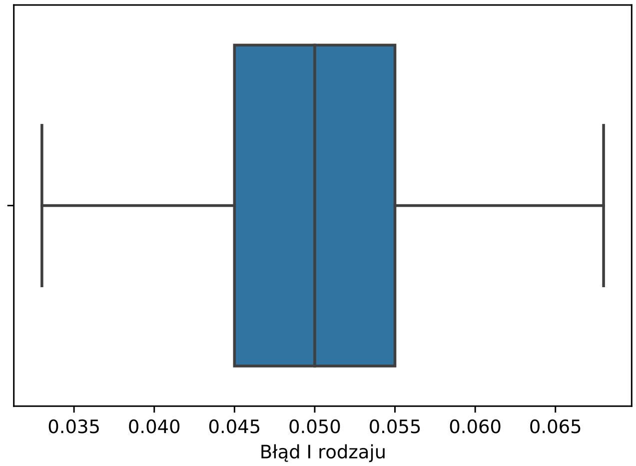
1. Ustalamy $\alpha = 0.05, n = 1000$
2. Generujemy $X_1, \ldots, X_n$ - prostą próbę losową z rozkładu $N(\mu, \sigma)$ (parametry zgodne z H_0)
3. Wyznaczamy wartość statystyki testowej Z (lub χ^2 w Zadaniu 2.)
4. Wyznaczamy obszar krytyczny (jego postać będzie zależała od postaci hipotezy alternatywnej, czyli dla każdego z podpunktów Zadania 1 oraz Zadania 2 będziemy tutaj mieć inny obszar)
5. Sprawdzamy, czy statystyka Z (lub χ^2 w drugim zadaniu) jest w obszarze krytycznym
6. Powtarzamy kroki $2. - 5.$ $N = 1000$ razy i zliczamy, ile razy statystyka testowa jest w obszarze krytycznym
7. $\{Z$ (lub χ^2) w obszarze krytycznym$\}/N$ daje w przybliżeniu błąd I rodzaju

7.1.1 Część do zadania 1.

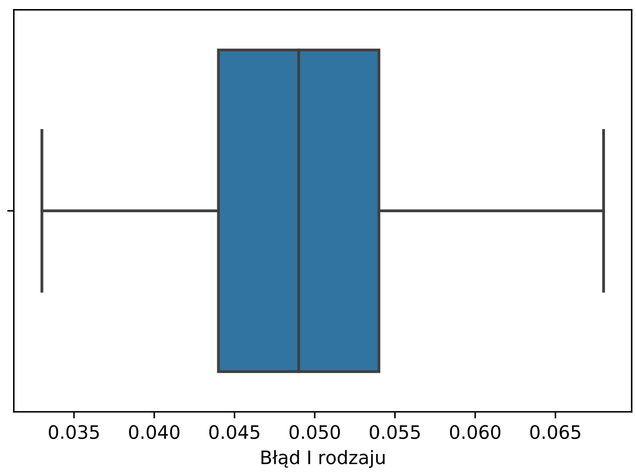
Wyniki otrzymane przez zaimplemetnowanie zaprezentowanego wyżej algorytmu zaprezentujemy za pomocą wykresów pudełkowych i tabelek. Dla $\alpha = 0.05$ powtórzymy algorytm wszystkie kroki $M = 100$ razy i otrzymane rezultaty przestawimy na wykresie. Natomiast w tabeli zamieszczone zostaną wartości również dla poziomów istotności $\alpha = 0.01$ i $\alpha = 0.05$



Rysunek 7: Wykres pudełkowy przedstawiający wartość błędu I rodzaju dla wartości średniej dla hipotezy alternatywnej H_1 i poziomu istotności $\alpha = 0.05$



Rysunek 8: Wykres pudełkowy przedstawiający wartość błędu I rodzaju dla wartości średniej dla hipotezy alternatywnej H_2 i poziomu istotności $\alpha = 0.05$



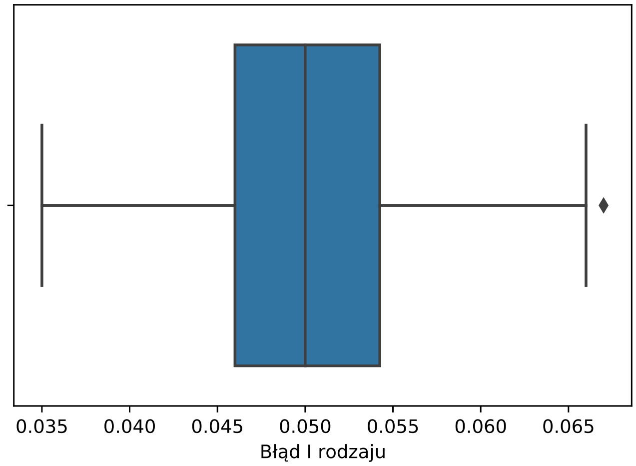
Rysunek 9: Wykres pudełkowy przedstawiający wartość błędu I rodzaju dla wartości średniej dla hipotezy alternatywnej H_3 i poziomu istotności $\alpha = 0.05$

	$H_1 : \mu \neq 1.5$	$H_2 : \mu > 1.5$	$H_3 : \mu < 1.5$
$\alpha = 0.1$	0.099	0.099	0.100
$\alpha = 0.05$	0.049	0.050	0.049
$\alpha = 0.01$	0.010	0.011	0.010

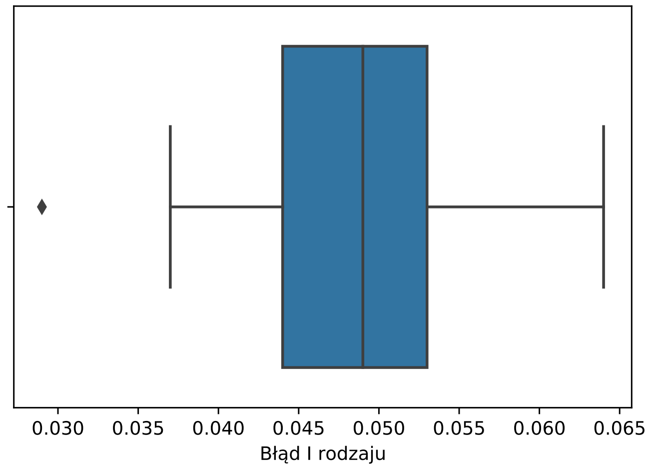
Tabela 1: Wartości błędu I rodzaju dla testów średniej przy różnych poziomach istotności α w zależności od postaci hipotezy alternatywnej.

7.1.2 Część do zadania 2.

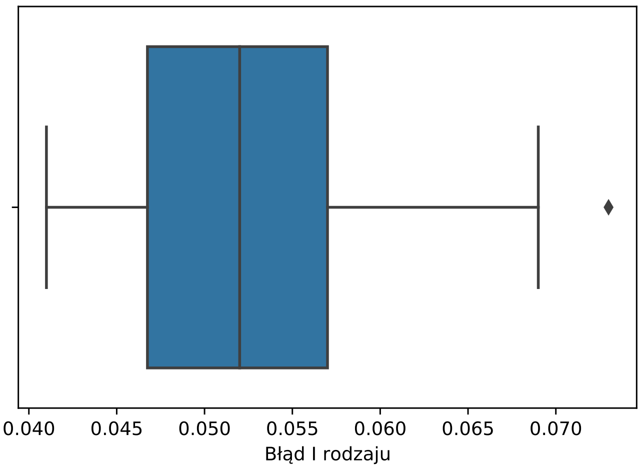
Dla testów wariancji rezultaty zostaną przedstawione analogicznie do tych dotyczących wartości średniej.



Rysunek 10: Wykres pudełkowy przedstawiający wartość błędu I rodzaju dla wariancji dla hipotezy alternatywnej H_1 i poziomu istotności $\alpha = 0.05$



Rysunek 11: Wykres pudełkowy przedstawiający wartość błędu I rodzaju dla wariancji dla hipotezy alternatywnej H_2 i poziomu istotności $\alpha = 0.05$



Rysunek 12: Wykres pudełkowy przedstawiający wartość błędu I rodzaju dla wariancji dla hipotezy alternatywnej H_3 i poziomu istotności $\alpha = 0.05$

	$H_1 : \sigma \neq 1.5$	$H_2 : \sigma > 1.5$	$H_3 : \sigma < 1.5$
$\alpha = 0.1$	0.100	0.096	0.103
$\alpha = 0.05$	0.050	0.048	0.052
$\alpha = 0.01$	0.010	0.008	0.010

Tabela 2: Wartości błędu I rodzaju dla testów wariancji przy różnych poziomach istotności α w zależności od postaci hipotezy alternatywnej.

7.1.3 Wnioski

Patrząc na otrzymane wyniki, można stwierdzić, że zaproponowana symulacja pozwoliła poprawnie wyznaczyć błąd I rodzaju dla testów zarówno dla wartości średniej jak i wariancji. W obu przypadkach otrzymane rezultaty są bardzo zbliżone do rozważanej wartości poziomu istotności α. Z Tabeli 1 i Tabeli 2 widać, że wraz ze wzrostem parametru α wzrasta błąd I rodzaju, czyli mamy do czynienia z większym prawdopodobieństwem odrzucenia hipotezy zerowej H_0, gdy jest ona prawdziwa.

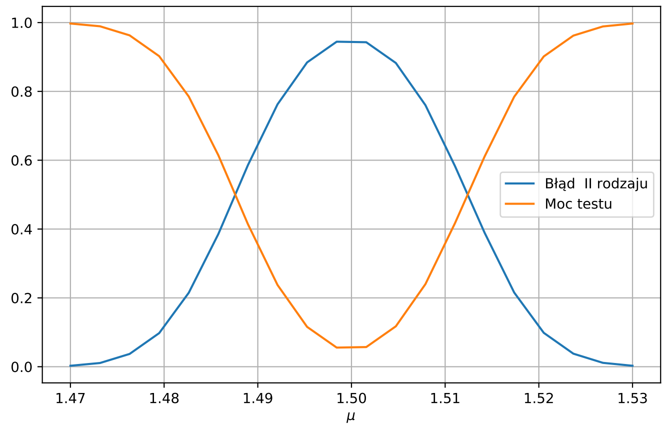
7.2 Błąd II rodzaju

Aby symulacyjnie wyznaczyć błąd II rodzaju musimy wygenerować prostą próbę losową z rozkładu normalnego o parametrach zgodnych z H_1 (ale blisko tych z H_0) i sprawdzić, ile razy przyjmujemy hipotezę zerową.
Algorytm:

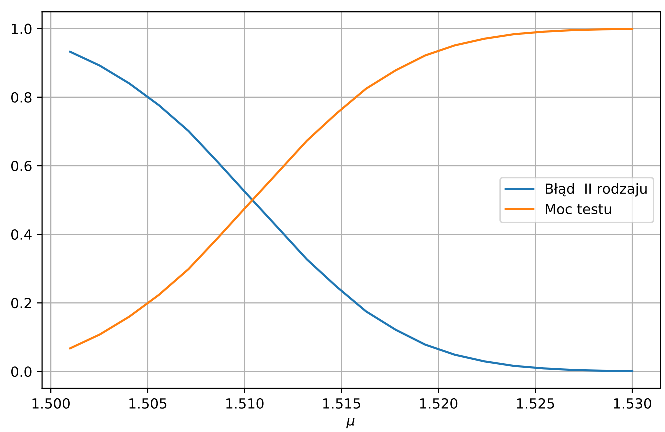
1. Ustalamy $\alpha = 0.05, \mu =$ (wartość zgodna z H_1), $\sigma = 0.2, n = 1000$
2. Generujemy $X_1, \ldots, X_n$ - prostą próbę losową z rozkładu $N(\mu, \sigma)$
3. Wyznaczamy wartość statystyki testowej Z (lub χ^2 w Zadaniu 2.)
4. Wyznaczamy obszar krytyczny (jego postać będzie zależała od postaci hipotezy alternatywnej, czyli dla każdego z podpunktów Zadania 1 oraz Zadania 2 będziemy tutaj mieć inny obszar)
5. Sprawdzamy, czy statystyka Z (lub χ^2 w drugim zadaniu) jest poza obszarem krytycznym
6. Powtarzamy kroki $2.-5.$ $N = 1000$ razy i zliczamy, ile razy statystyka testowa jest poza obszarem krytycznym
7. $\{Z$ (lub χ^2) poza obszarem krytycznym$\}/N$ daje w przybliżeniu błąd II rodzaju

7.2.1 Część do zadania 1.

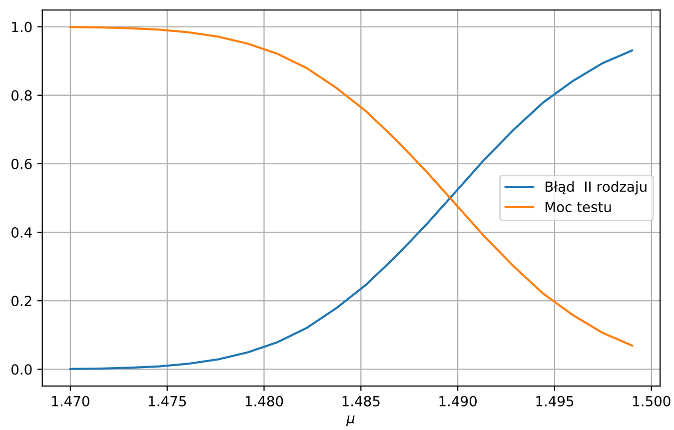
W celu zaprezentowania wyników dla błędu II rodzaju dla testów średniej rozważamy alternatywne μ i przygotowujemy wykres, który na osi x ma wartość μ a na osi y wartość błędu II rodzaju. Dodatkowo na wykresie zaznaczymy moc testu.



Rysunek 13: Błąd II rodzaju i moc testu dla różnych μ dla $\alpha = 0.05$, dla tezy alternatywnej $\mu \neq 1.5$



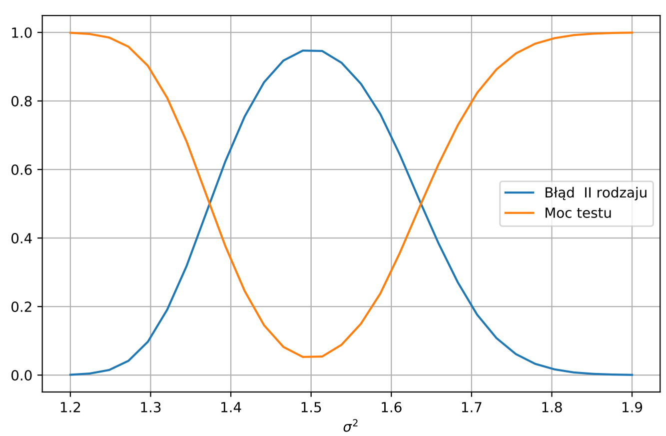
Rysunek 14: Błąd II rodzaju i moc testu dla różny μ dla $\alpha = 0.05$, dla tezy alternatywnej $\mu > 1.5$



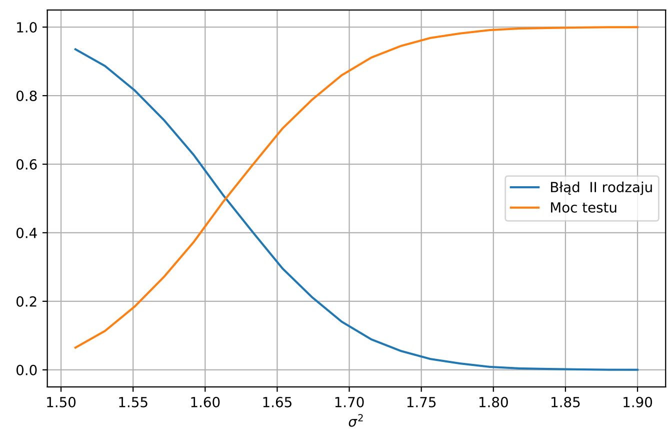
Rysunek 15: Błąd II rodzaju i moc testu dla różny μ dla $\alpha = 0.05$, dla tezy alternatywnej $\mu < 1.5$

7.2.2 Część do zadania 2.

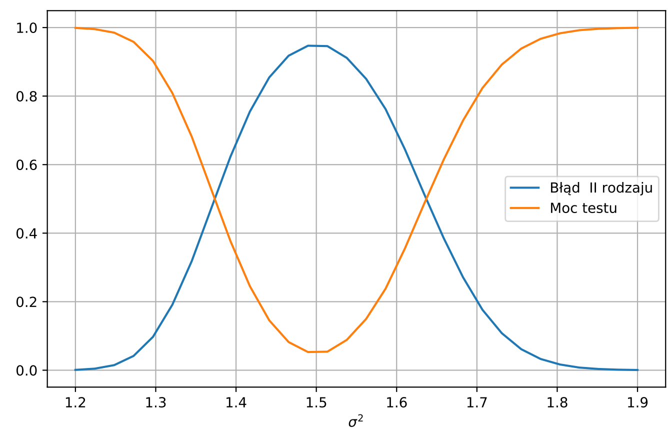
Wyniki dla testów wariancji zaprezentujemy analogicznie do wartości średniej.



Rysunek 16: Błąd II rodzaju i moc testu dla różny σ^2 dla $\alpha = 0.05$, dla tezy alternatywnej $\sigma^2 \neq 1.5$



Rysunek 17: Błąd II rodzaju i moc testu dla różny σ^2 dla $\alpha = 0.05$, dla tezy alternatywnej $\sigma^2 > 1.5$



Rysunek 18: Błąd II rodzaju i moc testu dla różny σ^2 dla $\alpha = 0.05$, dla tezy alternatywnej $\sigma^2 < 1.5$

7.2.3 Wnioski

Analizując wykresy z błędem II rodzaju i mocą testu dochodzimy do wniosku, że bład II rodzaju maleje wraz ze wzrostem różnicy pomiędzy parametrem rzeczywistym a parametrem występującym w hipotezie zerowej. Zauważamy dodatkowo, że moc testu zachowuje się przeciwnie i dodatkowo jak dodamy ją do błędu II rodzaju to dostaniemy 1, co zgadza się z teorią.