

STATYSTYKA W ANALIZIE DANYCH

SPRAWOZDANIE PROJEKT 2

Temat projektu

Problem 1 – badanie testów porównujących wartości średnie prób

Problem 2 – zbadanie różnic inflacji „studenckiej” i oficjalnej

Problem 3 – estymacja

Mateusz Tobiasz

Mikołaj Bańkowski

Prowadzący

Rafał Rytel-Andrianik

9 czerwca 2024

Spis treści

1 Problem 1 – badanie testów porównujących wartości średnie prób	2
1.1 Metoda rozwiązania problemu	2
1.2 Prezentacja wyników	3
1.2.1 Pierwszy eksperyment: równe wariancje i $n=30$	3
1.2.2 Drugi eksperyment: różne wariancje i $n=30$	4
1.2.3 Trzeci eksperyment: równe wariancje i $n=15$	5
1.2.4 Czwarty eksperyment: różne wariancje i $n=15$	6
1.3 Interpretacja wyników	7
2 Problem 2 – zbadanie różnic inflacji „studenckiej” i oficjalnej	8
2.1 Metoda rozwiązania problemu	8
2.2 Prezentacja wyników	9
2.2.1 Wykresy pudełkowe inflacji	10
2.2.2 Wykres pudełkowy różnic	12
2.2.3 Test statystyczny	13
2.3 Interpretacja wyników	14
3 Problem 3 – estymacja	14
3.1 Metoda rozwiązania problemu	15
3.2 Prezentacja wyników	15
3.2.1 Testy normalności	16
3.2.2 Histogramy	17
3.3 Interpretacja wyników	17

1 Problem 1 – badanie testów porównujących wartości średnie prób

Przyjmijmy, że n -elementowe próby losowe X i Y są z rozkładów normalnych potencjalnie różniących się wartościami oczekiwanymi i wariancjami. Chcemy zweryfikować hipotezę o niezerowości tej różnicy.

a) Określić właściwy test zakładający rozkłady normalne. Wykorzystując symulacje komputerowe, sprawdzić moc tego testu w zależności od różnicy delta wartości oczekiwanych obu prób. Określić sensowne kryterium wykrywalności i na podstawie wyników symulacji wskazać jaka jest minimalna wykrywalna różnica delta.

b) Powtórzyć (a) dla testu sumy rang Wilcozona i wyciągnąć odpowiednie wnioski. W szczególności przedyskutować założenie normalności:

- Jaki jest widoczny zysk z wykorzystania testu zakładającego normalność prób losowych
- Jakie jest ryzyko (co stanie się gdy rzeczywiste rozkłady jednak nie będą normalne).

1.1 Metoda rozwiązania problemu

W problemie 1 analizujemy dwie metody testowania: test t-Studenta dla dwóch prób niezależnych, który zakłada ich normalność oraz test sumy rang Wilcozona dla dwóch prób niezależnych, który nie narzuca określonego rozkładu. Przyjmujemy, że hipoteza zerowa (H_0) oznacza równość wartości oczekiwanych obu prób, natomiast hipoteza alternatywna (H_1) zakłada, że te wartości są od siebie różne.

W celu policzenia mocy testu (tj. prawdopodobieństwa wykrycia różnicy pomiędzy wartościami oczekiwanymi) zastosowano symulacje komputerowe. Polegało to na tym, że kilkaset razy wygenerowano dwie próbki o zadanej liczności z rozkładów normalnych różniących się wartościami oczekiwanymi o pewną wartość Δ oraz z potencjalnie różnymi wariancjami. Po wygenerowaniu obu próbek w danej iteracji sprawdzano p -wartość testu, aby określić, która hipoteza została przyjęta. Zliczając, ile razy hipoteza zerowa została odrzucona (tj. p -wartość była mniejsza od zadanego poziomu istotności) i dzieląc tę liczbę przez ilość iteracji, otrzymujemy moc testu.

Powyższe działanie zastosowano dla testu t-Studenta oraz dla testu sumy rang Wilcozona. W tym celu wygenerowano wektor różnych wartości Δ i dla każdej z nich obliczono moce obu testów. Wyniki przedstawione zostały na wykresach zależności mocy testu od różnicy Δ wartości oczekiwanych prób.

Następnie określono sensowne kryterium wykrywalności na poziomie mocy testu równej 0.8. Na tej podstawie dla obu testów wskazano minimalną wykrywalną różnicę pomiędzy wartościami oczekiwanymi prób.

Na koniec przeprowadzono analizę otrzymanych rezultatów, która wskazała czy widoczny jest zysk z wykorzystania testu zakładającego normalność prób losowych. Ponadto odpowiedziano na pytanie, jakie jest potencjalnie ryzyko zastosowania testu t-Studenta dla prób, które nie pochodzą z rozkładów normalnych.

1.2 Prezentacja wyników

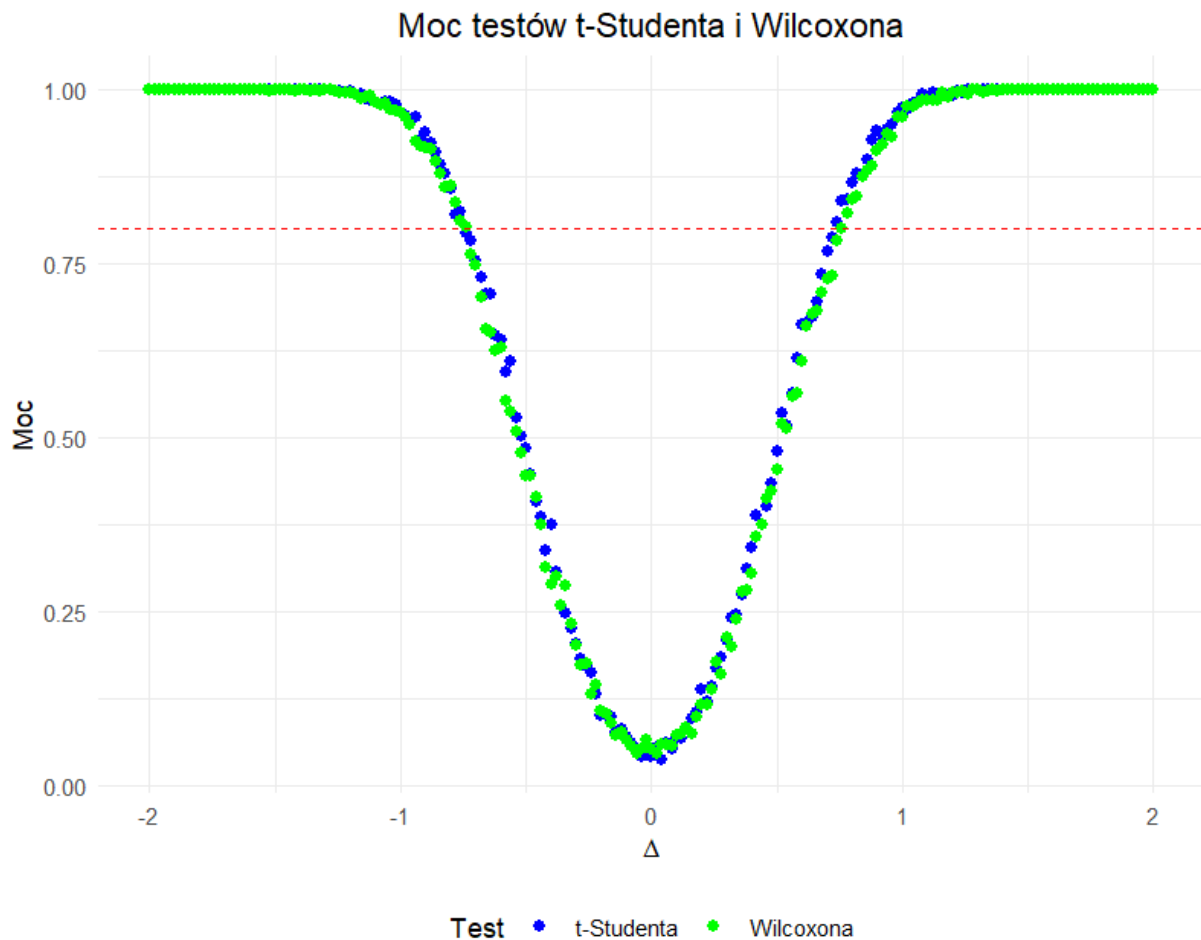
Dla wszystkich poniższych rozważań ustalono następujące wartości parametrów:

- $\alpha=0.05$ (poziom istotności obu testów)
- `num_simulations=1000` (liczba iteracji dla danej wartości Δ , w których generowane są próbki z rozkładów normalnych)
- $\mu_x=0$ (wartość oczekiwana pierwszego rozkładu normalnego – wartość oczekiwana drugiego rozkładu jest równa różnicy Δ)

Ponadto w każdym eksperymencie wygenerowano wektor wartości Δ za pomocą funkcji `seq` z przedziału $[-2, 2]$ i krokiem równym 0.02.

1.2.1 Pierwszy eksperyment: równe wariancje i $n=30$

W pierwszym eksperymencie przyjęto, że rozkłady normalne, z których generowane były próbki mają takie same wariancje – każda równa 1. Ponadto ustalono licznosc wygenerowanych prób losowych na 30. Otrzymano następujący wykres zależności mocy obu testów od różnicy pomiędzy wartościami oczekiwanymi generowanych prób:

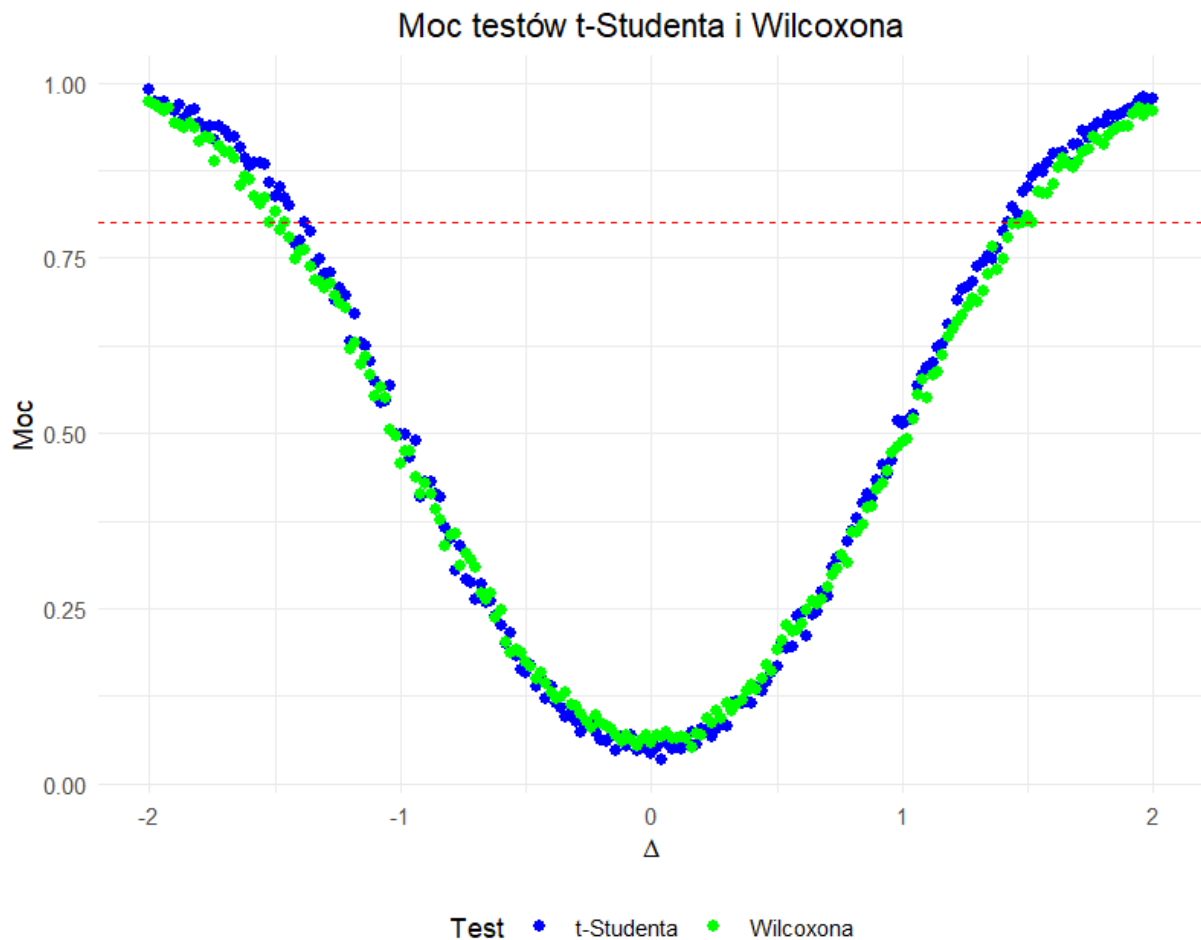


Rysunek 1: Wykres zależności mocy testu t-Studenta i testu sumy rang Wilcoxona od różnicy pomiędzy wartościami oczekiwanymi prób w eksperymencie 1. Czerwoną linią oznaczono poziom, gdzie moc testu jest równa 0.8.

Sprawdzono, że minimalna różnica Δ wykryta przez test t-Studenta z mocą co najmniej 0.8 wynosi 0.74. Dla testu sumy rang Wilcoxona różnica ta wynosi 0.76. Obie wartości podano jako wartości bezwzględne (tj. traktowane jako odległości wartości oczekiwanych prób od siebie).

1.2.2 Drugi eksperyment: różne wariancje i $n=30$

Podobnie jak w eksperymencie pierwszym ustalono licznosc prób na 30. Jednakże w tym przypadku zmieniono wariancje rozkładów normalnych, z których generowane były próbki. Jeden z nich otrzymał wariancję równą 1, a drugi wariancję równą 2.5. Ponownie wyrysowano wykres zależności mocy rozważanych testów od różnicy pomiędzy wartościami oczekiwanymi generowanych prób:

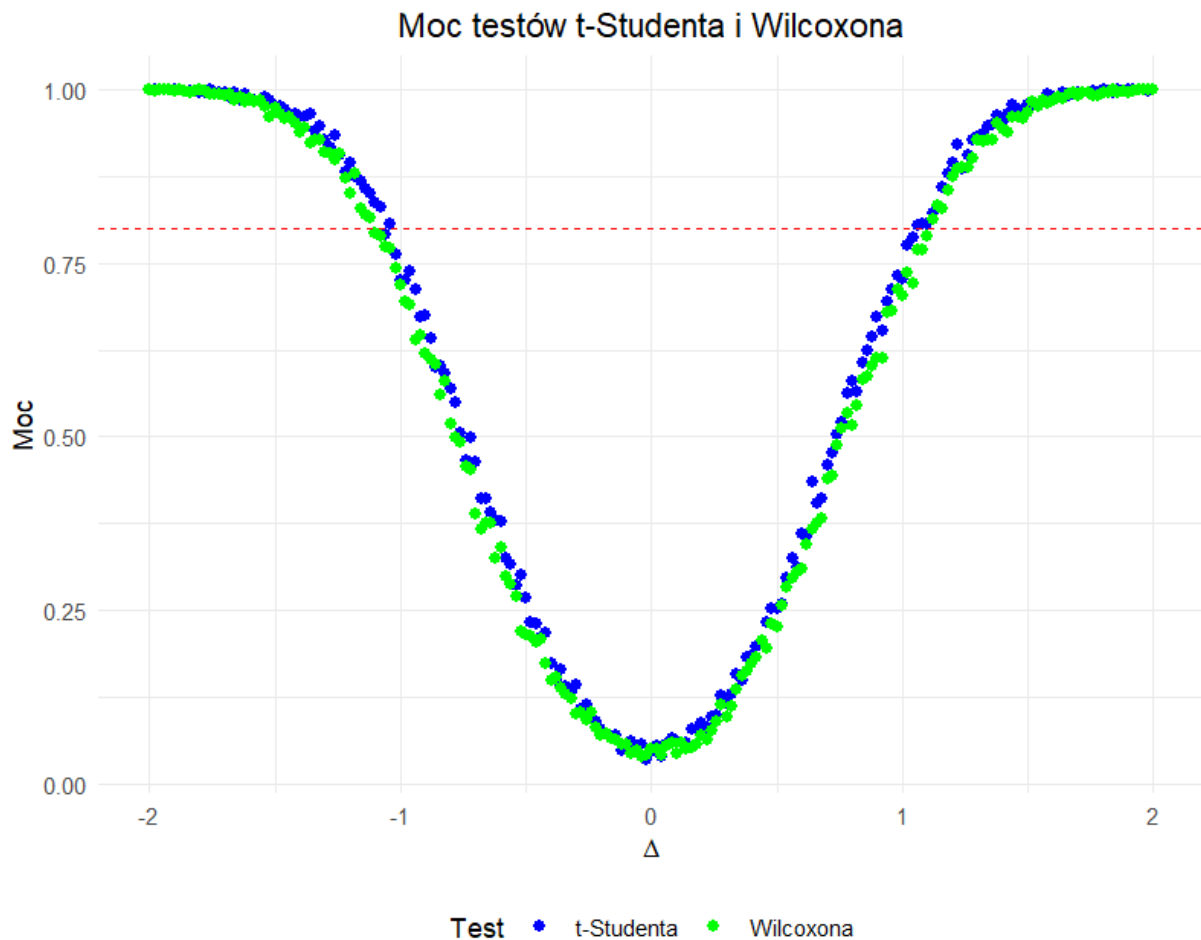


Rysunek 2: Wykres zależności mocy testu t-Studenta i testu sumy rang Wilcoxona od różnicy pomiędzy wartościami oczekiwanymi prób w eksperymencie 2. Czerwoną linią oznaczono poziom, gdzie moc testu jest równa 0.8.

Tym razem minimalna różnica Δ wykryta przez test t-Studenta z mocą co najmniej 0.8 wynosi 1.38. Dla testu sumy rang Wilcoxona różnica ta wynosi 1.48. Obie wartości podano jako wartości bezwzględne (tj. traktowane jako odległości wartości oczekiwanych prób od siebie).

1.2.3 Trzeci eksperyment: równe wariancje i $n=15$

Znów przyjęto, że rozkłady normalne, z których generowane są próby losowe mają takie same wariancje równe 1. Tym razem licznosc prób ustalono na 15. Poniżej znajduje się wykres zależności mocy obu testów od różnicy pomiędzy wartościami oczekiwanymi generowanych prób:

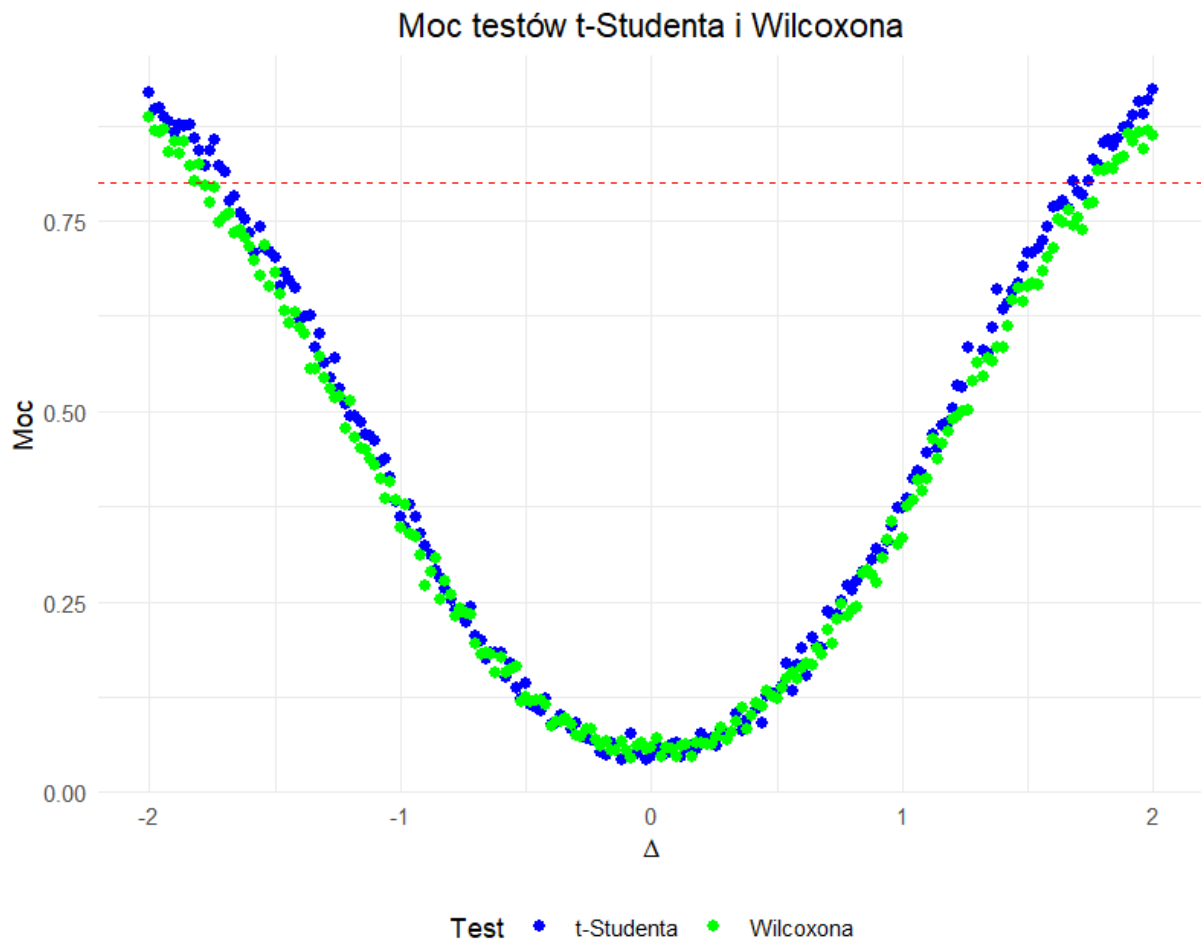


Rysunek 3: Wykres zależności mocy testu t-Studenta i testu sumy rang Wilcoxona od różnicy pomiędzy wartościami oczekiwanymi prób w eksperymencie 3. Czerwoną linią oznaczono poziom, gdzie moc testu jest równa 0.8.

W tym przypadku minimalna różnica Δ wykryta przez test t-Studenta z mocą co najmniej 0.8 wynosi 1.06. Dla testu sumy rang Wilcoxona różnica ta wynosi 1.10. Obie wartości podano jako wartości bezwzględne (tj. traktowane jako odległości wartości oczekiwanych prób od siebie).

1.2.4 Czwarty eksperyment: różne wariancje i $n=15$

Przyjęto taką samą licznosc prób jak w eksperymencie powyżej. Tym razem pierwszy rozkład normalny miał wariancję równą 1, a drugi wariancję równą 2. Otrzymano następujący wykres zależności mocy obu testów od różnicy pomiędzy wartościami oczekiwanymi generowanych prób:



Rysunek 4: Wykres zależności mocy testu t-Studenta i testu sumy rang Wilcoxona od różnicy pomiędzy wartościami oczekiwanymi prób w eksperymencie 4. Czerwoną linią oznaczono poziom, gdzie moc testu jest równa 0.8.

Zweryfikowano, że minimalna różnica Δ wykryta przez test t-Studenta z mocą co najmniej 0.8 wynosi 1.74. Dla testu sumy rang Wilcoxona różnica ta wynosi 1.82. Obie wartości podano jako wartości bezwzględne (tj. traktowane jako odległości wartości oczekiwanych prób od siebie).

1.3 Interpretacja wyników

Analizując powyższe eksperymenty, można wywnioskować, że test t-Studenta wykrywa rozbieżności pomiędzy rozkładami dla mniejszych wartości Δ niż test sumy rang Wilcoxona. Różnica ta jest niewielka w przypadku, gdy wariancje obu rozkładów są równe (eksperyment 1 i 3). Większą korzyść z wykorzystania testu t-Studenta można zaobserwować, kiedy wariancje prób nie są równe. Wtedy, chociażby w eksperymencie 2, zauważamy, że test t-Studenta wykrył różnicę pomiędzy wartościami oczekiwanymi rozkładów dla Δ mniejszej o 0.1 niż wartość, przy której różnicę wykrył test sumy rang Wilcoxona dla ustalonej mocy. Są to wyniki zgodne z intuicją, ponieważ test t-Studenta ma silniejsze założenia niż test sumy rang Wilcoxona. Sprawia to, że test ten jest dokładniejszy i jest w stanie wykryć mniejsze różnice pomiędzy wartościami

oczekiwanymi. Aczkolwiek przed wykonaniem eksperymentów spodziewano się, że zysk z wykorzystania testu t-Studenta będzie większy – szczególnie dla równych wariancji prób oba testy wykrywają podobną różnicę pomiędzy wartościami oczekiwanymi dla ustalonej mocy.

Przez silniejsze założenia test t-Studenta dla niewłaściwych próbek może dawać mylne rezultaty. Może się to stać w przypadku, gdy nie pochodzą one z rozkładów normalnych. Wtedy przez niezgodność z teorią test ten może dawać gorsze wyniki, a nawet nie spełniać oczekiwań np. poziom istotności testu może nie być równy temu zadanemu. Na takie odstępstwa od rozkładów normalnych odporny jest właśnie test sumy rang Wilcoxona.

2 Problem 2 – zbadanie różnic inflacji „studenckiej” i oficjalnej

W projekcie 1 analizowana była inflacja „własna studencka” w odniesieniu do oficjalnej. Teraz należy dokładniej przeanalizować ich różnicę (należy posłużyć się wynikami projektu 1 wybranej osoby z zespołu; inflacja własna to próba X, inflacja oficjalna to Y, czas określa nam kolejne elementy obu prób)

- narysować wykresy pudełkowe obu inflacji i wyciągnąć wnioski
- narysować wykres pudełkowy różnic $X_i - Y_i$
- sformułować odpowiednią hipotezę o tym, że jedna z tych inflacji jest większa od drugiej, przeprowadzić test i zinterpretować wynik. Proszę pamiętać, aby przy wyborze testu przemyśleć założenia – w szczególności czy należy zakładać rozkład normalny prób.

2.1 Metoda rozwiązania problemu

W celu dokładnego przeanalizowania różnicy między inflacją „studencką” a oficjalną, należy przeprowadzić szereg kroków analitycznych i statystycznych. Najpierw wczytamy dane dotyczące inflacji „studenckiej” i oficjalnej, które będą stanowiły próby X (inflacja studencka) i Y (inflacja oficjalna). Następnie narysujemy wykresy pudełkowe dla obu zbiorów danych, co pozwoli nam ocenić mediany, kwartyle oraz ewentualne wartości odstające dla obu typów inflacji.

W kolejnym kroku narysujemy wykres pudełkowy różnic między inflacją studencką a oficjalną ($X_i - Y_i$), co pozwoli na ocenę, czy istnieją istotne różnice między obiema inflacjami w poszczególnych miesiącach. Następnie sformułujemy hipotezę zerową oraz hipotezę alternatywną: Hipoteza zerowa (H_0) zakłada, że mediany inflacji studenckiej i oficjalnej są równe, podczas gdy hipoteza alternatywna (H_1) zakłada, że mediana inflacji studenckiej różni się od mediany inflacji oficjalnej.

Ze względu na charakter danych, które są próbami sparowanymi i mogą nie mieć normalnego rozkładu, zastosujemy test Wilcoxona dla prób sparowanych, ponieważ jest to test nieparametryczny i nie wymaga założenia normalności danych. Alternatywnie, jeśli dane będą miały rozkład normalny, możemy zastosować test t-Studenta dla prób sparowanych. Przeprowadzimy wybrany test statystyczny i obliczymy p-wartość, aby sprawdzić, czy różnice między inflacjami są statystycznie istotne.

Na koniec zinterpretujemy wyniki testu, oceniając ich istotność statystyczną. Wyciągniemy wnioski dotyczące różnic między inflacją studencką a oficjalną i skomentujemy, czy wyniki potwierdzają hipotezę zerową, czy ją odrzucają.

2.2 Prezentacja wyników

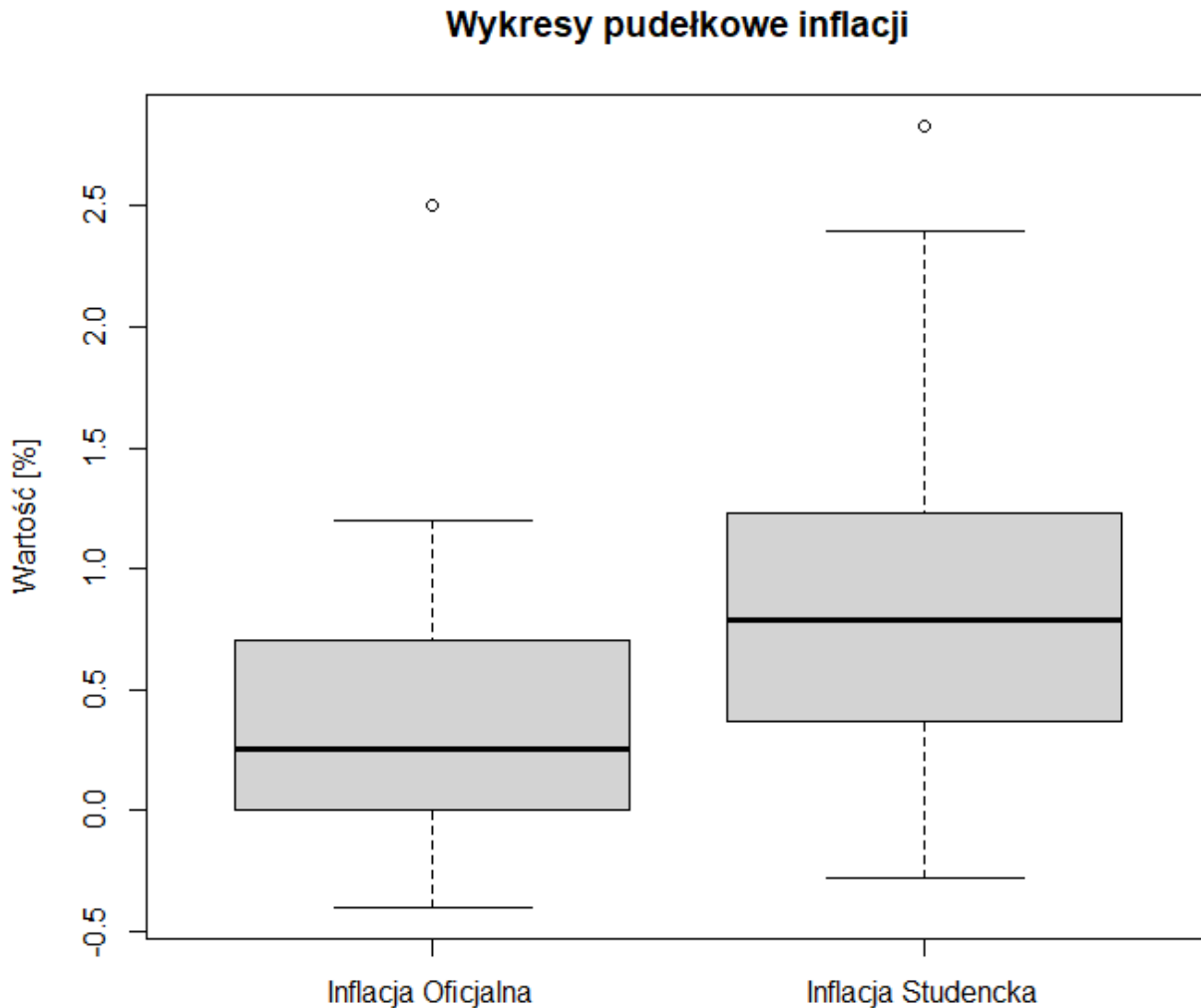
W tym dziale zamieszczono wyniki przeprowadzonej, analizy na podstawie, których zostaną wyciągnięte odpowiednie wnioski.

Wykres pudełkowy pokazuje, że inflacja studencka jest generalnie wyższa i bardziej zmienna w porównaniu do inflacji oficjalnej. Większy rozstęp międzykwartyłowy i szerszy zakres wartości wskazują na większe zróżnicowanie wartości inflacji wśród studentów. Obecność outlierów w obu przypadkach sugeruje pewne ekstremalne wartości inflacji, które mogą mieć znaczący wpływ na analizę statystyczną.

Wykres pudełkowy różnic ($X_i - Y_i$) pokazuje, że w większości przypadków inflacja studencka przewyższała inflację oficjalną. Zmienność różnic jest umiarkowana, a obecność wartości odstających wskazuje na pewne ekstremalne przypadki, gdzie różnice były znacznie większe lub mniejsze niż większość obserwacji. Mediana różnic wyraźnie powyżej 0 potwierdza, że generalnie inflacja studencka była wyższa niż oficjalna.

Wynik testu statystycznego Wilcoxona wskazuje na statystycznie istotną różnicę między inflacjami ($p\text{-value} < 0.05$).

2.2.1 Wykresy pudełkowe inflacji



Rysunek 5: Wykresy pudełkowe inflacji

Mediana:

- Mediana inflacji oficjalnej wynosi około 0.25, co oznacza, że połowa obserwacji inflacji oficjalnej jest mniejsza od tej wartości.
- Mediana inflacji studenckiej wynosi około 0.75, co oznacza, że połowa obserwacji inflacji studenckiej jest większa od tej wartości.

Rozstęp międzykwartyłowy (IQR):

- IQR dla inflacji oficjalnej jest mniejszy niż dla inflacji studenckiej, co wskazuje na mniejszą zmienność w wartościach inflacji oficjalnej w porównaniu do inflacji studenckiej.

Zakres danych:

- Inflacja oficjalna ma mniejszy zakres wartości w porównaniu do inflacji studenckiej. Najniższe wartości inflacji oficjalnej sięgają około -0.5, a najwyższe nie przekraczają wartości 1.25.
- Inflacja studencka ma szerszy zakres wartości w porównaniu do inflacji oficjalnej. Najniższe wartości inflacji studenckiej sięgają około -0.25, a najwyższe nie przekraczają wartości 2.5.

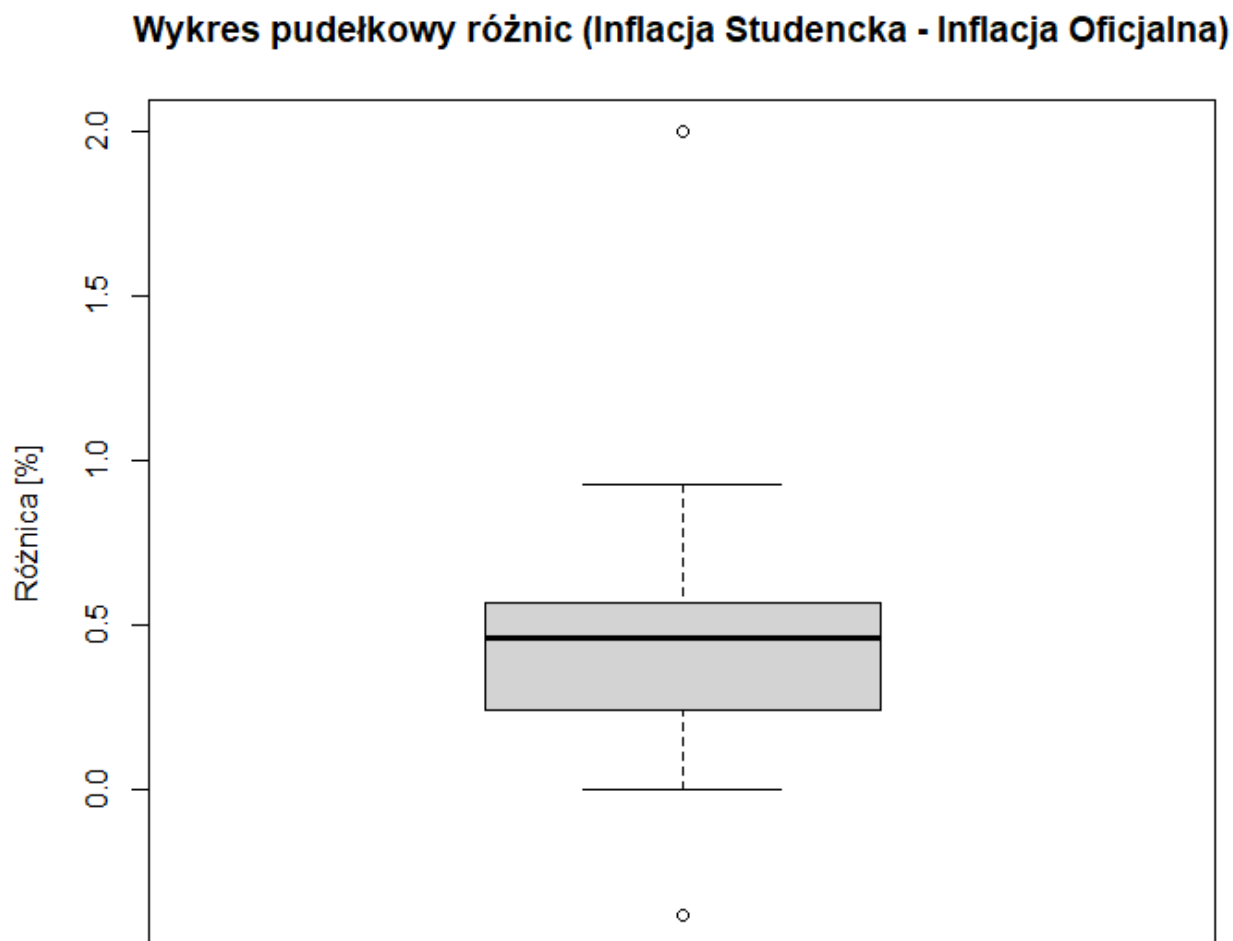
Wartości odstające:

- W obu zestawach danych widać punkty odstające, które są zaznaczone jako kółka. W przypadku inflacji oficjalnej są to powyżej 1.25, natomiast dla inflacji studenckiej powyżej 2.5.

Średnie:

- Średnie wartości dla obu inflacji (choć nie są bezpośrednio przedstawione na wykresie) mogą być interpretowane w kontekście położenia mediany i rozkładu wartości. Można przypuszczać, że średnia inflacji studenckiej jest wyższa niż inflacji oficjalnej, co wynika z wyższego położenia pudełka dla inflacji studenckiej.

2.2.2 Wykres pudełkowy różnic



Rysunek 6: Wykres pudełkowy różnic

Mediana:

- Mediana różnic inflacji studenckiej i oficjalnej wynosi około 0.4, co oznacza, że w większości przypadków inflacja studencka była wyższa niż inflacja oficjalna.

Rozstęp międzykwartyłowy (IQR):

- IQR dla różnic wynosi około 0.3, co wskazuje na umiarkowaną zmienność różnic między inflacją studencką a oficjalną.

Zakres danych:

- Różnice wartości inflacji studenckiej i oficjalnej mają zakres od około 0 do około 1. Oznacza to, że w niektórych przypadkach inflacja studencka była nieco niższa niż oficjalna, ale w innych przypadkach była znacznie wyższa.

Wartości odstające:

- Na wykresie widoczne są wartości odstające, które są zaznaczone jako kółka. Wartości te znajdują się poniżej 0 i powyżej 1, co sugeruje, że w niektórych przypadkach różnice inflacji studenckiej i oficjalnej były znacząco inne niż większość obserwacji

Średnie:

- Choć średnia różnic nie jest bezpośrednio przedstawiona na wykresie, można przypuszczać, że jest ona powyżej 0, co wynika z położenia mediany. To sugeruje, że średnio inflacja studencka jest wyższa niż oficjalna.

2.2.3 Test statystyczny

W celu zbadania różnic między inflacją "studencką" a oficjalną przyjmujemy następujące hipotezy:

- Hipoteza zerowa (H_0): Nie ma istotnej różnicy między inflacją studencką a oficjalną. Formalnie, zakładamy, że średnia (lub mediana) różnica między inflacją studencką a oficjalną jest równa zero.
- Hipoteza alternatywna (H_1): Istnieje istotna różnica między inflacją studencką a oficjalną. Formalnie, zakładamy, że średnia (lub mediana) różnica między inflacją studencką a oficjalną jest różna od zera.

Uzasadnienie wyboru testu, będzie polegało na przeprowadzeniu test Shapiro-Wilka dla normalności rozkładu różnic.

Test t-Studenta: dla prób sparowanych jest stosowany, gdy próbki są zależne (tzn. każda para pomiarów pochodzi od tego samego obiektu, np. te same miesiące dla różnych typów inflacji) i dane mają rozkład normalny. W przypadku tego testu zakładamy, że różnice między parami są normalnie rozkładowane. Jeśli p-wartości testu Shapiro-Wilka są większe niż 0.05, możemy przyjąć, że dane są normalnie rozkładowane, i zastosować test t-Studenta.

Test Wilcoxona: jest nieparametrycznym odpowiednikiem testu t-Studenta dla prób sparowanych. Jest stosowany, gdy próbki są zależne, ale nie możemy założyć, że różnice między parami są normalnie rozkładowane. Test ten porównuje mediany różnic między parami, a nie ich średnie, co czyni go bardziej odpornym na odstępstwa od normalności oraz obecność wartości odstających. Jeśli dane nie spełniają założenia normalności (tj. p-wartość testu Shapiro-Wilka jest mniejsza niż 0.05 dla którejkolwiek z próbek), stosujemy test Wilcoxona dla prób sparowanych.

Statystyka	Wartość
W	0.85337
p-wartość	0.009574

Rysunek 7: Test normalności

Wynik testu normalności wykazał, że należy przeprowadzić test Wilcoxona dla prób sparowanych. Po przeprowadzeniu odpowiedniego testu statystycznego otrzymujemy p-wartość, która informuje nas, czy możemy odrzucić hipotezę zerową na rzecz hipotezy alternatywnej:

- Jeśli p-wartość < 0.05 : Odrzucamy hipotezę zerową na poziomie istotności 0.05, co oznacza, że istnieje statystycznie istotna różnica między inflacją studencką a oficjalną.
- Jeśli p-wartość ≥ 0.05 : Nie mamy wystarczających dowodów, aby odrzucić hipotezę zerową na poziomie istotności 0.05. Oznacza to, że nie możemy stwierdzić, że istnieje statystycznie istotna różnica między inflacją studencką a oficjalną.

Test	Data	Statystyka (V)	p-wartość	Hipoteza alternatywna
Wilcoxon signed rank test	inflacja_studencka and inflacja_oficjalna	146	0.001088	true location shift is not equal to 0

Rysunek 8: Wynik testu statystycznego

Ponieważ p-wartość wynosi 0.001088, co jest znacznie mniejsze niż 0.05, możemy odrzucić hipotezę zerową na poziomie istotności 0.05. To oznacza, że istnieje statystycznie istotna różnica między inflacją studencką a oficjalną.

2.3 Interpretacja wyników

Na podstawie przeprowadzonej analizy, wykresów pudełkowych oraz wyników testu statystycznego, możemy jednoznacznie stwierdzić istnienie istotnej różnicy między inflacją "studencką" a oficjalną. Nasza hipoteza zerowa, zakładająca równość median inflacji, została odrzucona na rzecz hipotezy alternatywnej. To oznacza, że studenci doświadczają inflacji w sposób istotnie różny od oficjalnych wskaźników.

Wykresy pudełkowe jasno obrazują różnice między oboma typami inflacji. Inflacja studencka wykazuje większą zmienność, co sugeruje, że doświadczenia studenckie w zakresie inflacji mogą być bardziej zróżnicowane niż oficjalne wskaźniki sugerują. Wykres pudełkowy różnic dodatkowo potwierdza te różnice, nie tylko pod kątem statystycznym, ale również w kontekście praktycznym.

Wybór testu Wilcoxona był uzasadniony, biorąc pod uwagę charakter danych oraz ich brak normalności. Test ten jest bardziej odporny na odstępstwa od założenia normalności i obecność wartości odstających, co czyni go odpowiednim narzędziem dla naszych danych. Ostatecznie, wyniki testu Wilcoxona potwierdzają istnienie statystycznie istotnej różnicy między inflacją studencką a oficjalną i potwierdzają, że nasze założenia były trafne.

3 Problem 3 – estymacja

Wiadomo, że estymatory największej wiarygodności mają asymptotycznie rozkład normalny. Proszę zilustrować tę cechę dowolnym przykładem (próby można tworzyć generatorem liczb

losowych).

3.1 Metoda rozwiązania problemu

Asymptotyczna zbieżność estymatorów największej wiarygodności do rozkładu normalnego polega na tym, że wraz ze wzrostem liczności próby rozkład tych estymatorów zbiega do rozkładu normalnego. Ta własność jest prawdziwa dla wszystkich rozkładów tj. próbki można generować z dowolnie przyjętego rozkładu. W szczególności:

$$\sqrt{nI(\theta)}(\hat{\theta}_n - \theta) \xrightarrow{d} \mathcal{N}(0, 1) \quad \text{gdy } n \rightarrow \infty \quad (1)$$

gdzie n to liczność próby, a $I(\theta)$ to informacja Fishera.

W celu zilustrowania tego faktu posłużono się rozkładem Poissona z parametrem $\lambda=0.5$, na którym przeprowadzono symulacje komputerowe. Polegały one na tym, że dla różnych licznosci prób ($n = 10, 100, 1000, 10000, 100000, 1000000$) wykonano 5000 symulacji, generując próbki z wybranego rozkładu Poissona.

Następnie, korzystając z faktu, że estymator największej wiarygodności parametru λ w rozkładzie Poissona jest średnią arytmetyczną próby, dla każdej próbki obliczono wartość estymatora $\hat{\lambda}$. W ten sposób dla każdej licznosci n uzyskano 5000 przykładowych wartości estymatora $\hat{\lambda}$, które posłużyły do wyrysowania histogramów.

Jednakże przed stworzeniem wykresów dokonano normalizacji otrzymanego estymatora. Korzystając z 1 oraz faktu, że informacja Fishera dla rozkładu Poissona przyjmuje postać $\frac{1}{\lambda}$, otrzymujemy znormalizowany estymator postaci:

$$\hat{\lambda}_{\text{norm}} = (\hat{\lambda} - \lambda) \sqrt{\frac{n}{\lambda}}$$

Mając znormalizowane estymatory, stworzono ich histogramy, na które naniesiono wykres funkcji gęstości rozkładu $\mathcal{N}(0, 1)$ w celu porównania.

3.2 Prezentacja wyników

Wyniki testu Shapiro-Wilka dla różnych licznosci prób przedstawiono w tabeli:

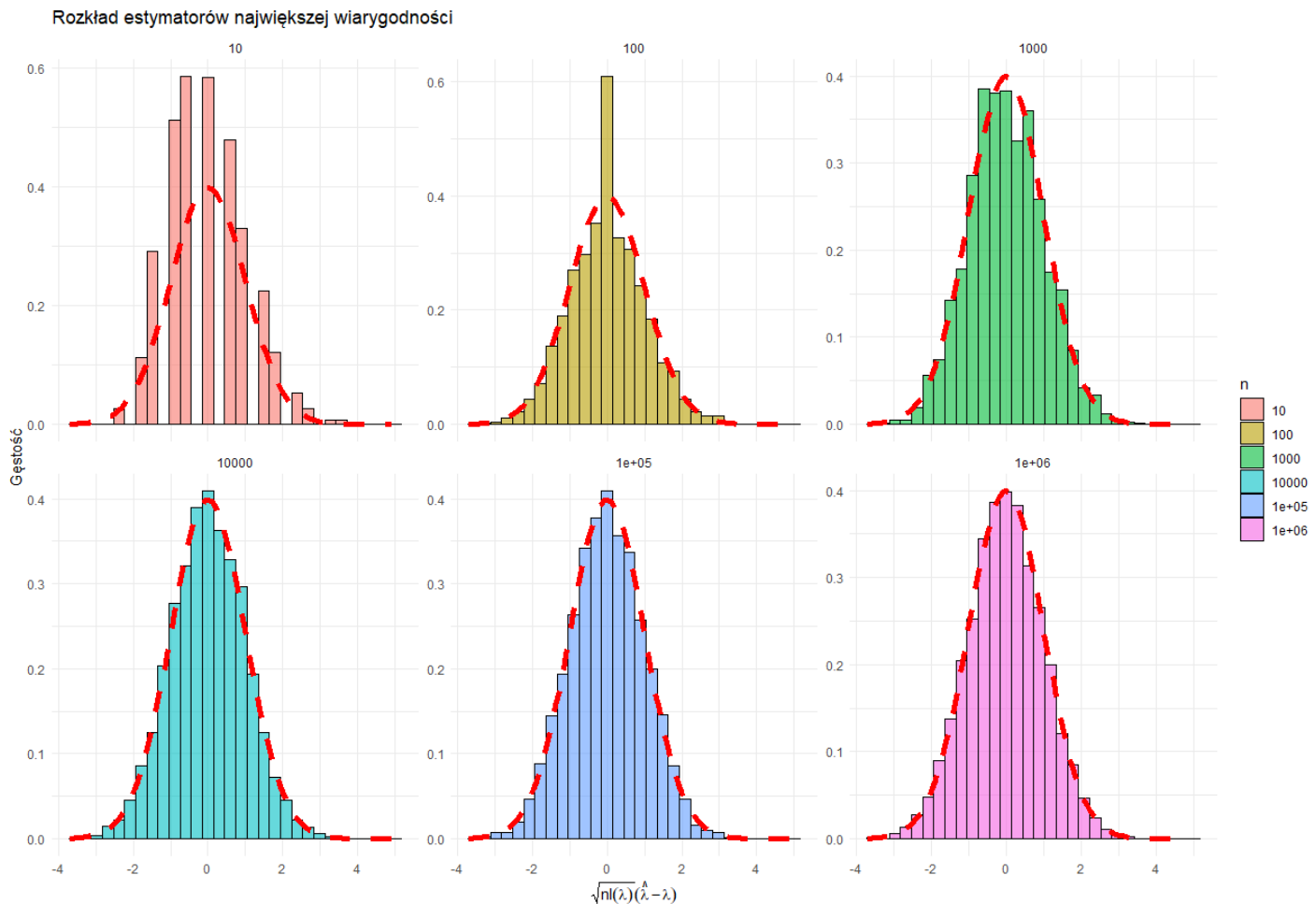
3.2.1 Testy normalności

n	Wartość W	p-wartość
10	0.96976	< 2.2e-16
100	0.99739	1.446e-07
1000	0.99889	0.00195
10000	0.99977	0.8908
100000	0.99958	0.3623
1000000	0.99989	0.9991

Rysunek 9: Wynik testu statystycznego

Na wykresach poniżej przedstawiono histogramy znormalizowanych estymatorów dla różnych licznosci prób wraz z teoretycznym rozkładem normalnym (czerwona przerywana linia):

3.2.2 Histogramy



Rysunek 10: Histogramy

3.3 Interpretacja wyników

Testy Shapiro-Wilka pokazują, że dla większych licznosci prób ($n \geq 10000$), rozkład estymatorów największej wiarygodności jest zgodny z rozkładem normalnym (wysokie p-wartości). Dla mniejszych prób ($n < 10000$), estymatory nie spełniają założeń normalności (niskie p-wartości). Histogramy z nałożonym teoretycznym rozkładem normalnym potwierdzają te wyniki wizualnie, pokazując, że znormalizowane estymatory coraz lepiej dopasowują się do rozkładu normalnego wraz ze wzrostem licznosci prób.