Sprawozdanie 2

ZUZANNA NASIŁOWSKA, MARIA NOWCKA

1. Wprowadzenie

Niniejszy raport skupia się na analizie mocy testów statystycznych w sytuacjach, gdy założenia dotyczące danych są stopniowo osłabiane. Testy statystyczne stanowią kluczowy element weryfikacji hipotez i oceny, czy badane próbki spełniają określone warunki rozkładu. W statystyce, jednym z głównych wyzwań jest dobór odpowiednich testów, które nie tylko wykrywają odchylenia od normalności, ale również są odporne na różne formy zaburzeń danych. Celem raportu jest więc nie tylko zbadanie skuteczności testów w optymalnych warunkach, ale także analiza ich działania w sytuacjach, gdy założenia o normalności rozkładu danych nie są w pełni spełnione.

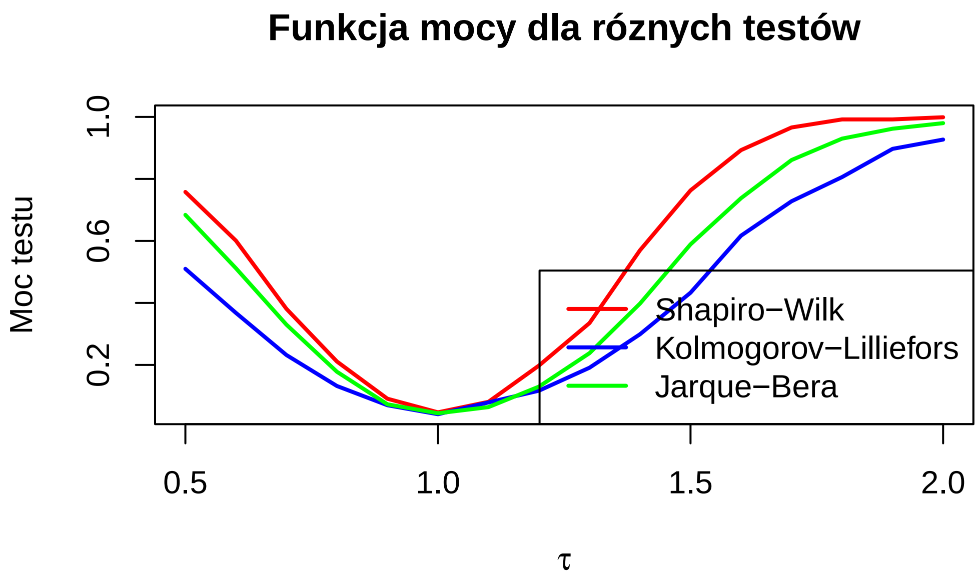
W naszych badanich wykorzystałyśmy trzy popularne testy normalności:

- Test Shapiro-Wilka: to test parametryczny, który weryfikuje normalność danych poprzez analizę stosunku wariancji próbki do wariancji oczekiwanej w przypadku normalności,
- Test Kołmogorowa-Lillieforsa: to test nieparametryczny, który mierzy maksymalną różnicę między dystrybuantą empiryczną a teoretyczną funkcją rozkładu normalnego,
- Test Jarque-Bera: to test statystyczny oparty na skośności i kurtozie, który ocenia odchylenia od normalności na podstawie momentów wyższych rzędów.

Każdy z powyżej wymienionych testów ma swoje zalety i ograniczenia, a ich skuteczność zależy od charakterystyki danych. Celem analizy jest porównanie ich mocy w różnych warunkach.

Zadanie 1

Rozważmy próbę $(X_1, ...X_{100})$ z rozkładu normalnego $\mathcal{N}(2,\ 1)$ przekształconego przez transformatę Sinh-arcsinh z $\nu = 0$. Korzystając z symulacji Monte Carlo wykonaj wykres funkcji mocy w zależności od τ na przedziale $(0.5,\ 2)$ dla wszystkich trzech testów. Czy istnieje test jednostajnie najmocniejszy spośród nich?



UMPT to test to test statystyczny, który dla każdego poziomu istotności α ma największą moc spośród wszystkich testów w danym problemie.

Moc testów: Moc testu statystycznego to prawdopodobieństwo odrzucenia hipotezy zerowej, gdy jest ona fałszywa.

Wyniki przedstawione na powyższym wykresie pokazują, jak moc testów zmienia się w zależności od parematry τ:

- Wartości τ **bliskie 1**: Moc testów jest najmniejsza, co oznacza, że trudniej im wykryć odchylenia od normalności,
- Wartości dalekie od 1: Moc testów wzrasta, co wskazuje, że testy lepiej indentyfikują niezgodność z rozkładem normalnym.

Dyskusja na temat mocy testów:

- Test Shapiro-Wilka wykazuje największą moc w niemal całym zakresie τ, zwłaszcz dla skrajnych wartości,
- Test Kołmogorowa-Lillieforsa ma niższą moc niż Shapiro-Wilk i Jarque-Bera, szczególnie dla mniejszych odchyleń. Jest mniej efektywny w wykrywaniu subtelnych różnic,
- Test Jarque-Bera osiąga wartości mocy podobne do Shapiro-Wilka, ale jest mniej czuły na niektóre odchylenia.

Porównanie oraz wnioski

- **Shapiro-Wilk** ma najwyższą moc w całym zakresie τ, co oznacza, że jest najlepszy w wykrywaniu nienormalności, zwłaszcza w małych próbach,
- **Kołmogorov-Lillefor** ma najniższą moc, co sugeruje, że nie jest najlepszym wyborem w tym wypadku,
- Patrząc na wykres, test Shapiro-Wilka jest najczęściej najmocniejszy, ale w niektórych zakresach τ test Jarque-Bera może osiągać podobne wartości. Jednak żaden test nie dominuje całkowicie we wszystkich przypadkach, więc nie można stwierdzić, że istnieje jednozncznie najmocniejszy test. Shapiro-Wilk jest jednak najlepszym wyborem sposród trzech testów.

Zadanie 2

Mamy próbę $(X_1, ...X_{100})$ z rozkładu normalnego $\mathcal{N}(2, 1)$ przekształconego przez transformatę Sinh-arcsinh z $\tau = 1$. Korzystając z symulacji Monte Carlo wykonaj wykres funkcji mocy w zależności od   na przedziale $(-2, 2)$ dla wszystkich trzech testów. Czy istnieje test jednostajnie najmocniejszy spośród nich?

Parametr v zmienia się w przedziale (-2,2), co osłabia założenia normalności, transformacja Sinh-arcsinh zmienia kształt rozkładu, co powduje, że dane nie są już dokładnie normalne:

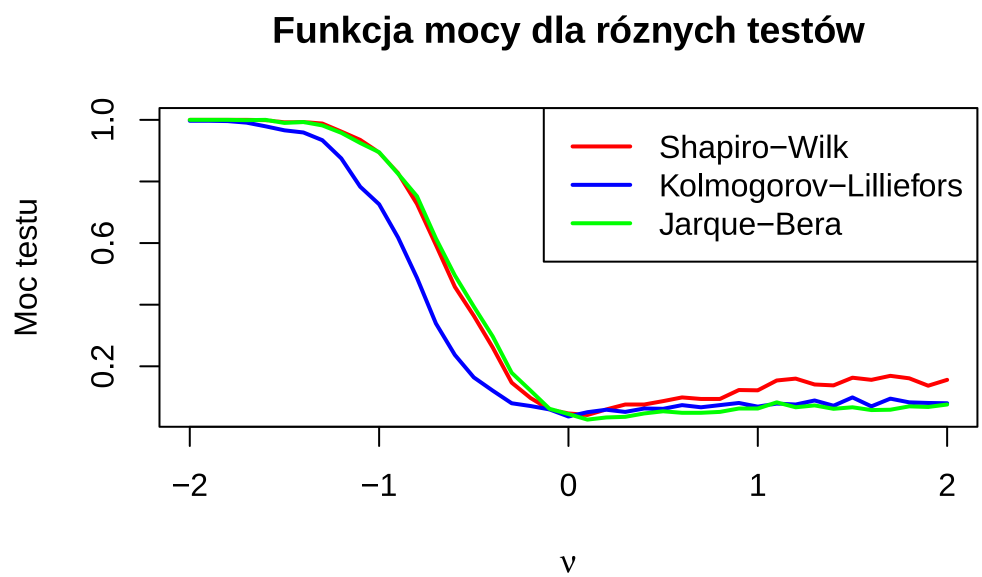
Z wykresu można zauważyć kilka istotnych rzeczy:

- **Dla wartości** v bliskich 0 (czyli gdy dane są bliskie normalności), moc testów jest bardzo niska: Oznacza to, że testy nie wykrywają odchyleń od normalności, co jest spodziewane, bo dane są wtedy w przybliżeniu normalne,
- **Dla dużych wartości** $|v|$, szczególnie ujemnych, most testu wzrasta: Oznacza to, że im większa deformacja rozkładu, tym łatwiej testom wykryć, że dane nie pochodzą z normalności.

Porównanie testów i wnioski:

- Shapiro-Wilk jest najbardziej czuły na odchylenia: dla dużych wartości v jego moc rośnie najszybciej,
- Jarque-Bera ma podobną moc do Shapiro-Wilka, ale jest nieco mniej efektywny w wykrywaniu nieznacznych odchyleń,
- Kołmogorov-Lillefors ma najniższą moc w całym zakresie v, co oznacza, że jest najmniej efektywny w wykrywaniu odchyleń od normalności,
- Podobnie jak w zadaniu 1, **test Shapiro-Wilka wykazuje najwyższą moc w większości przypadków**, co sugeruje, że jest najlepszym wyborem. Jednak dalej w niektórych zakresach, test Jarque-Bera ma zbliżoną skuteczność. Nie można jednoznacznie stwierdzić, że istnieje test jednostajnie najmocniejszy.

Wyniki tego zadania pokazują, że gdy założenia normalności są osłabione (tj. dane są przekształcone przez transformację Sinh-arcsinh z różnymi wartościami v), moc testów znacząco się zmienia. W szczególności dla wartości $v \approx 0$, czyli gdy dane są bliskie normalności, moc testów jest bardzo niska. To oznacza, że w tym zakresie testy nie są w stanie odróżnić lekko zmodyfikowanego rozkładu od normalnego. Jeśli dane tylko lekko odbiegają od normalności, testy mogą nie wykryć tego odchylenia i błędnie zaakceptować hipoteze zerową. Jeśli założenia normalności są mocniej naruszone (duże wartości $|v|$), testy zaczynają działać lepiej.



Zadanie 3

Mamy próbę $(X_1, ...X_{100})$ taką, że zmienne losowe $Y_i = \frac{X_i - 2}{1}$ są z rozkładu t-Studenta $\mathcal{T}(\nu)$. Korzystając z symulacji Monte Carlo wykonaj wykres funkcji mocy w zależności od na przedziale $(0.05, 20)$ dla wszystkich trzech testów. Czy istnieje test jednostajnie najmocniejszy spośród nich?

W niniejszym zadaniu dokonaliśmy transformacji, co sprawia, że zmienne Y_i mają rozkład t-tudenta. Parametr v kontroluje liczbę stopni swobody, co bezpośrednio wpływa na kształt rozkładu. Dla dużych v rozkład $T(v)$ staje się coraz bardziej zbliżony do rozkładu normalnego. **Na podstawie wykresu możemy stwierdzić, że:**

- Dla małych wartości v moc testów jest wysoka,
- Gdy v jest małe, rozkład t-Studenta znacznie rożni się od normalnego (ma cięższe ogony). Wszystkie testy dobrze wykrywają to odchylenie i często odrzucają hipoteze zerową,
- Największą moc w tej strefie (małe v) ma Shapiro-Wilk i Jarque-Bera, natomiast Kołmogorov-Lillefors jest najsłabszy,

- Dla dużych wartości v moc testów drastycznie spada. Gdy v rośnie, rozkład t-Studenta coraz bardziej przypomina normalny, więc testy mają trudność z odrzuceniem hipotezy zerowej.

Porównanie:

- **Jarque-Bera**: Najlepszy w wykrywaniu odchyleń od normalności. Dobrze wykrywa odchylenia w skośności i kurtozie, które są kluczowe dla rozkładu t-Studenta,
- **Shapiro-Wilk**: Jest skuteczny dla małych wartości v, ale traci moc szybciej niż Jarque-Bera. Działa gorzej, ponieważ jest wrażliwy na rożnice w kształcie rozkładu,
- **Kołmogorov-Lilleforsa**: Wypada najgorzej, ma najniższą moc prawie w całym zakresie.

Im bardziej osłabione są założenia, tym trudniej testom poprawnie wykryć brak normalności.