Raport 5

Magdalena Potok

2024-01-06

Wstęp

Niech $X_1, \ldots, X_m$ będą niezależnymi zmiennymi losowymi z rozkładem o ciągłej dystrybuancie F. Niech $Y_1, \ldots, Y_n$ będą niezależnymi zmiennymi losowymi z rozkładem o ciągłej dystrybuancie G. Zakładamy, że wszystkie zmienne są niezależne. Rozważamy problem testowania hipotezy

$$H_0 : F = G$$

przeciwko alternatywie

$$H_1 : F \neq G$$

na poziomie istotności $\alpha = 0.05$.

Niech $N = m + n$, a $Z = (Z_1, \ldots, Z_N) = (X_1, \ldots, X_m; Y_1, \ldots, Y_n)$ będzie wektorem połączonych prób. Niech R_i będzie rangą Z_i w próbie Z, $i = 1, \ldots, N$. Klasyczna liniowa statystyka rangowa związana z funkcją wynikową $\phi \in L_2(0, 1)$ ma postać

$$T_\phi = \sqrt{\frac{mn}{N}} \left(\frac{1}{m} \sum_{i=1}^{m} \phi \left(\frac{R_i - 0.5}{N} \right) - \frac{1}{n} \sum_{i=m+1}^{N} \phi \left(\frac{R_i - 0.5}{N} \right) \right),$$

gdzie wybór funkcji ϕ determinuje czułość testu opartego na statystyce T_ϕ. Jeśli $\phi(u) = \phi_1(u) = \sqrt{3}(2u - 1)$, otrzymujemy statystykę Wilcoxona. Wybór $\phi(u) = \phi_2(u) = \sqrt{48}(0.25 - |u - 0.5|)$ prowadzi do statystyki Ansari-Bradley'a. Jeśli $\int_0^1 \phi(u)\, du = 0$ oraz $\int_0^1 \phi^2(u)\, du = 1$, to przy prawdziwości hipotezy zerowej statystyka T_ϕ ma asymptotyczny rozkład standardowy normalny. Ponadto, H_0 odrzucamy na korzyść H_1 dla dużych wartości T_ϕ^2.

W problemie testowania (H_0, H_1) innym klasycznym rozwiązaniem jest na przykład test Kołmogorowa-Smirnowa odrzucający H_0 dla dużych wartości statystyki

$$KS = \sqrt{\frac{mn}{N}} \sup_{x \in \mathbb{R}} |F_m(x) - G_n(x)|,$$

gdzie F_m oraz G_n są dystrybuantami empirycznymi w próbie X-ów i Y-ów, odpowiednio.

Celem raportu jest badanie zachowania funkcji mocy wybranych rozwiązań problemu testowania hipotezy. Dokładniej, będziemy analizować

1. test Wilcoxona oparty na statystyce $W = T_{\phi_1}^2$,
2. test Ansari-Bradley'a oparty na statystyce $AB = T_{\phi_2}^2$,
3. test Lepage'a oparty na statystyce $L = W + AB$,
4. test Kołmogorowa-Smirnowa oparty na statystyce KS.

Zadanie 1

Wygeneruję $m = n = 20$ obserwacji z rozkładu $N(0,1)$. Na ich podstawie obliczę wartości statystyk $W.AB, L$ i KS (odpowiednio dla testu Wilcoxona, Ansari-Bradley'a, Lapage'a oraz Kołmogorowa-Smirnowa). Doświadczenie powtórzę 10000 razy. Wyznaczę wartości krytyczne odpowiadającym im testom prawostronnym.

m,n		W	AB	L	KS
20	Wartość krytyczna	3.997	3.888	6.070	1.265
50	Wartość krytyczna	3.763	3.926	5.927	1.300

Wyniki testów porównawczych wykazały zbliżone wartości krytyczne dla testów Wilcoxona i Ansari-Bradley'a, niezależnie od rozmiaru próby. Te testy, oparte na rozkładzie χ^2 z jednym stopniem swobody, prezentują stabilność wartości krytycznych, co potwierdza zgodność z oczekiwaniami. Co istotne, sposób generowania tych wartości uwzględnia jedynie rangi obserwacji, niezależnie od rozkładu obserwacji w warunkach hipotezy.

Jednakże, wartości krytyczne dla testów Lepage'a i Kołmogorova-Smirnova wykazują pewne różnice w zależności od rozmiaru próby. Dla tych testów wartości krytyczne wydają się być zmienną w stosunku do wielkości próby, co sugeruje inną dynamikę w porównaniu z testami Wilcoxona i Ansari-Bradley'a.

Wartości te są istotne, gdyż pomagają w interpretacji istotności statystycznej wyników testów porównawczych, podkreślając różnice między nimi oraz stabilność niektórych metod w różnych warunkach próbnych.

Zadanie 2

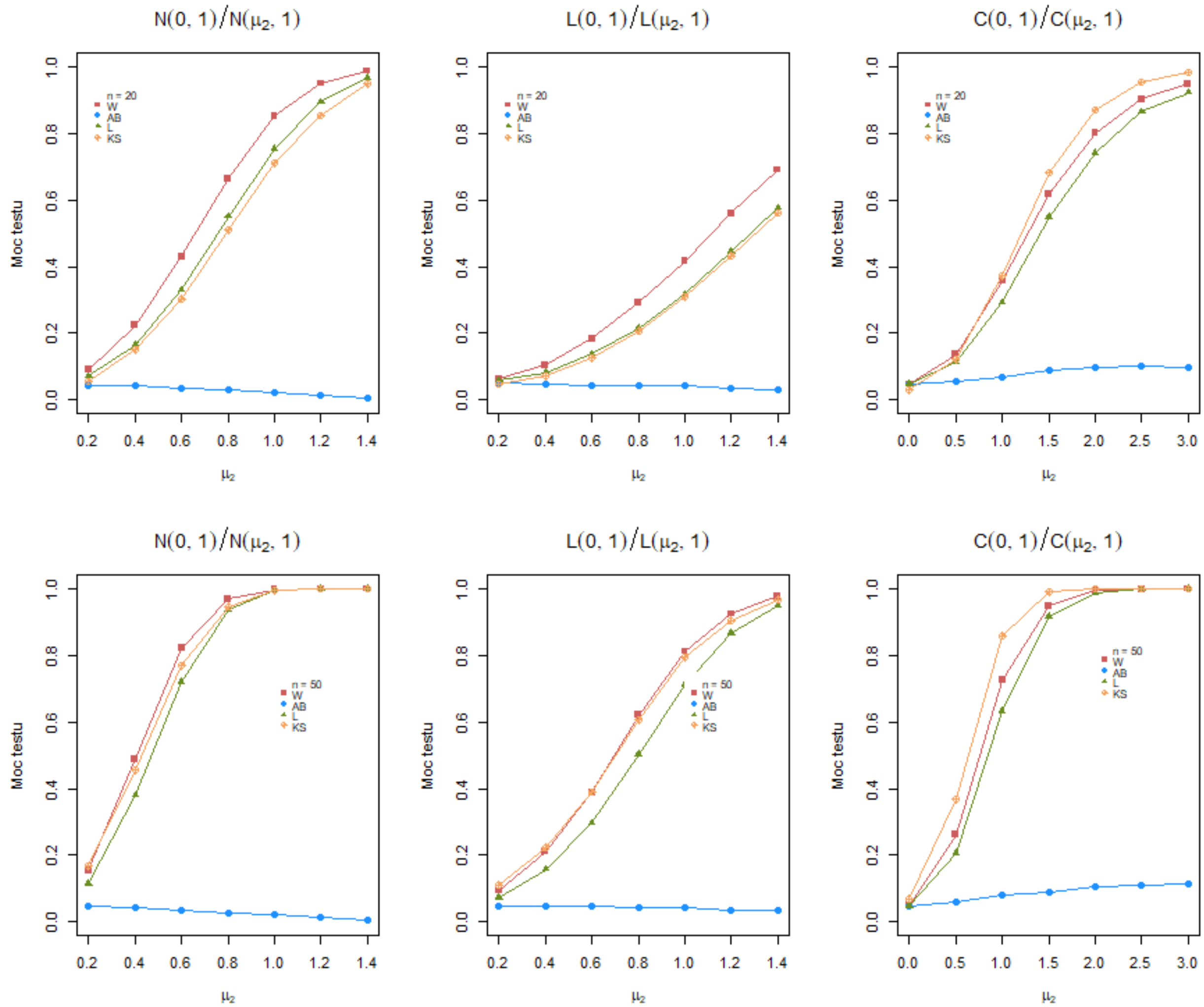
Wygeneruję $m = n = 20$ oraz $m = n = 50$ obserwacji z rozkładów
(a) $N(\mu_1, \sigma_1)$ i $N(\mu_2, \sigma_2)$,
(b) $L(\mu_1, \sigma_1)$ i $L(\mu_2, \sigma_2)$,
(c) $C(\mu_1, \sigma_1)$ i $C(\mu_2', \sigma_2)$,
gdzie $\mu_1 = 0,\ \sigma_1 = 1,\ \sigma_2 = 1,\ \mu_2 \in \{0.2, 0.4, 0.6, 0.8, 1, 1.2, 1.4\}$ oraz $\mu_2' \in \{0, 0.5, 1, 1.5, 2, 2.5, 3\}$. Następnie wyznaczę wartość statystyk $W.AB, L$ i KS. Doświadczenie powtórzę 10000 razy i na tej podstawie oszacuję wartość funkcji mocy analizowanych testów. Narysuję wyestymowane funkcje mocy w zależności od parametrów μ_2 i μ_2'.

Wykresy zostały przedstawione na następnej stronie.

Dla przypadku, gdy m = n = 20, analiza skuteczności testów w zależności od rozkładu obserwacji wykazuje interesujące trendy. Gdy obserwacje pochodzą zarówno z rozkładu normalnego, jak i logistycznego, test Wilcoxona wypada najlepiej. Ten test wykazuje się wysoką skutecznością w wykrywaniu różnic między dystrybucjami, co czyni go preferowanym wyborem. Bezpośrednio po nim znajduje się test Lepage'a, który również prezentuje dobrą skuteczność. Dodatkowo, test Kołmogorowa-Smirnowa plasuje się wysoko w hierarchii skuteczności dla tych dwóch rozkładów. Natomiast test Ansari-Bradley'a wykazuje się nieskutecznością w tym scenariuszu, co oznacza, że jest mniej przydatny w wykrywaniu różnic między rozkładami dla tego konkretnego zestawu próbek.

Gdy obserwacje pochodzą z rozkładu Cauchy'ego, test Kołmogorowa-Smirnowa okazuje się najbardziej skuteczny w wykrywaniu różnic w rozkładach. Jednakże testy Wilcoxona i Lepage'a wciąż prezentują się dobrze, choć ich skuteczność jest niższa w porównaniu do testu Kołmogorowa-Smirnowa. Test Ansari-Bradley'a nadal wykazuje nieskuteczność, co sugeruje, że dla tego konkretnego rodzaju danych jest mniej przydatny w wykrywaniu różnic w rozkładach.

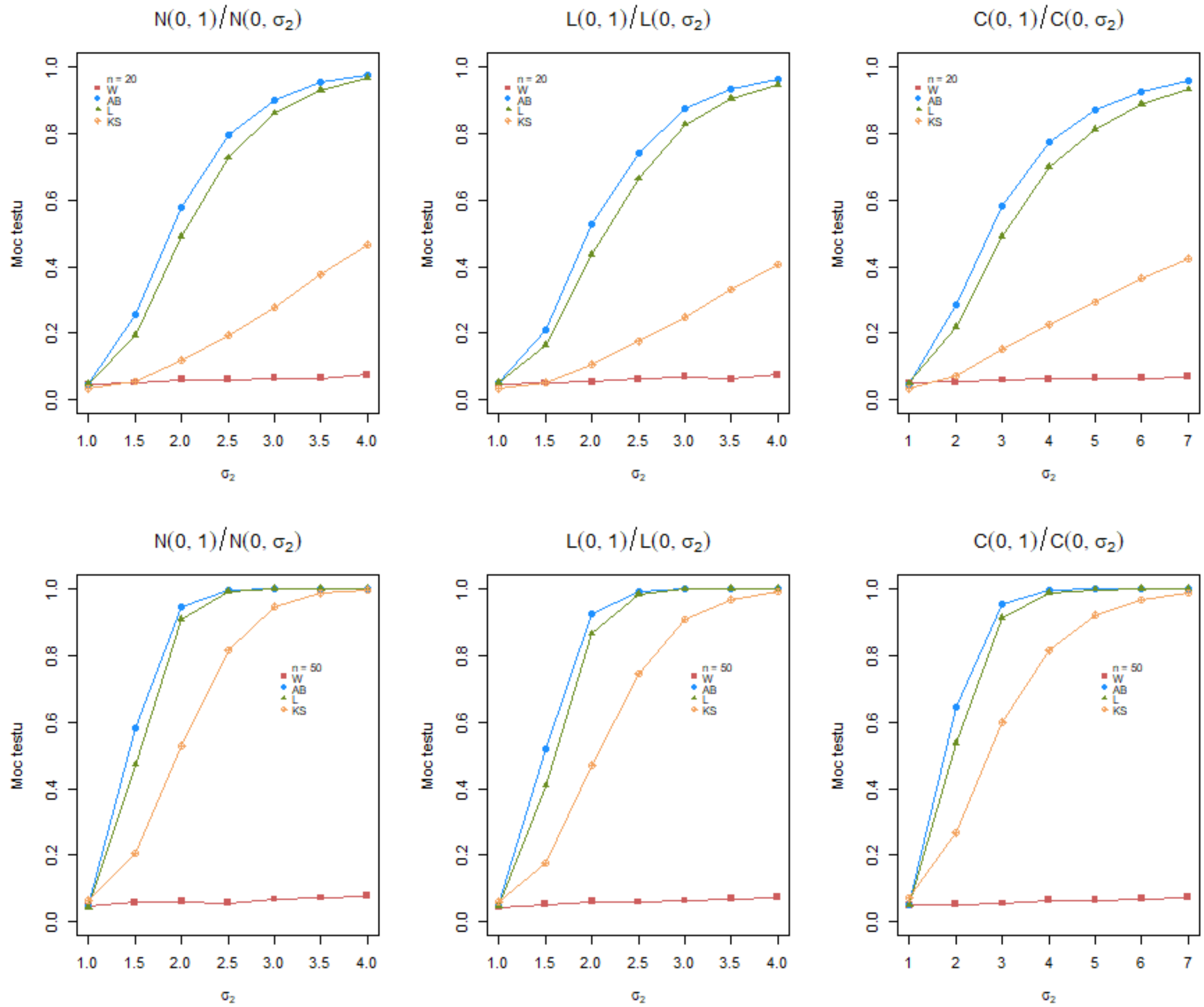
Przy rozważeniu większego rozmiaru próby (m = n = 50), obserwuje się wzrost skuteczności testów dla niższych wartości parametru μ_2 niż w poprzednio analizowanym przypadku (m = n = 20). Test Kołmogorowa-Smirnowa zyskuje na skuteczności wraz ze zwiększeniem próby, szczególnie w przypadku obserwacji z rozkładu normalnego lub logistycznego. Zauważalne jest również, że większy rozmiar próby nie wpływa na zwiększenie skuteczności testu Ansari-Bradley'a, który nadal pozostaje nieskuteczny w wykrywaniu różnic między rozkładami.



Podsumowując, zwiększenie rozmiaru próby przynosi korzyści, szczególnie dla testu Kołmogorowa-Smirnowa, podczas gdy test Ansari-Bradley'a pozostaje nieskuteczny niezależnie od rozmiaru próby. Test Wilcoxona i Lepage'a pozostają skuteczne, choć ich hierarchia skuteczności może się różnić w zależności od rozkładu obserwacji i rozmiaru próby.

Zadanie 3

Wygeneruję $m = n = 20$ oraz $m = n = 50$ obserwacji z rozkładów
(a) $N(\mu_1, \sigma_1)$ i $N(\mu_2, \sigma_2)$,
(b) $L(\mu_1, \sigma_1)$ i $L(\mu_2, \sigma_2)$,
(c) $C(\mu_1, \sigma_1)$ i $C(\mu_2, \sigma_2')$,
gdzie $\mu_1 = 0,\ \sigma_1 = 1,\ \mu_2 = 1,\ \sigma_2 \in \{1, 1.5, 2, 2.5, 3, 3.5, 4\}$ oraz $\sigma_2' \in \{1, 2, 3, 4, 5, 6, 7\}$. Następnie wyznaczę wartość statystyk $W.AB, L$ i KS. Doświadczenie powtórzę 10000 razy i na tej podstawie oszacuję wartość funkcji mocy analizowanych testów. Narysuję wyestymowane funkcje mocy w zależności od parametrów σ_2 i σ_2'.



Dla przypadku, gdy m = n = 20, analiza wykazała interesujące różnice w skuteczności testów w zależności od badanych rozkładów. Test Ansari-Bradley'a wykazał się jako najbardziej skuteczny dla wszystkich trzech badanych rozkładów, co świadczy o jego czułości na zmiany parametru skali w porównaniu do pozostałych testów. Jest to istotna obserwacja, szczególnie w kontekście analizy zmienności parametrów rozkładu.

Niezwykle istotne jest, że test Wilcoxona, który sprawdzał się dobrze przy zmianach parametru przesunięcia, w przypadku zmiany parametru skali okazał się mniej skuteczny. Z kolei test Lepage'a, podobnie jak dla zmian parametru przesunięcia, również prezentuje się skutecznie w wykrywaniu różnic między rozkładami przy zmianach parametru skali.

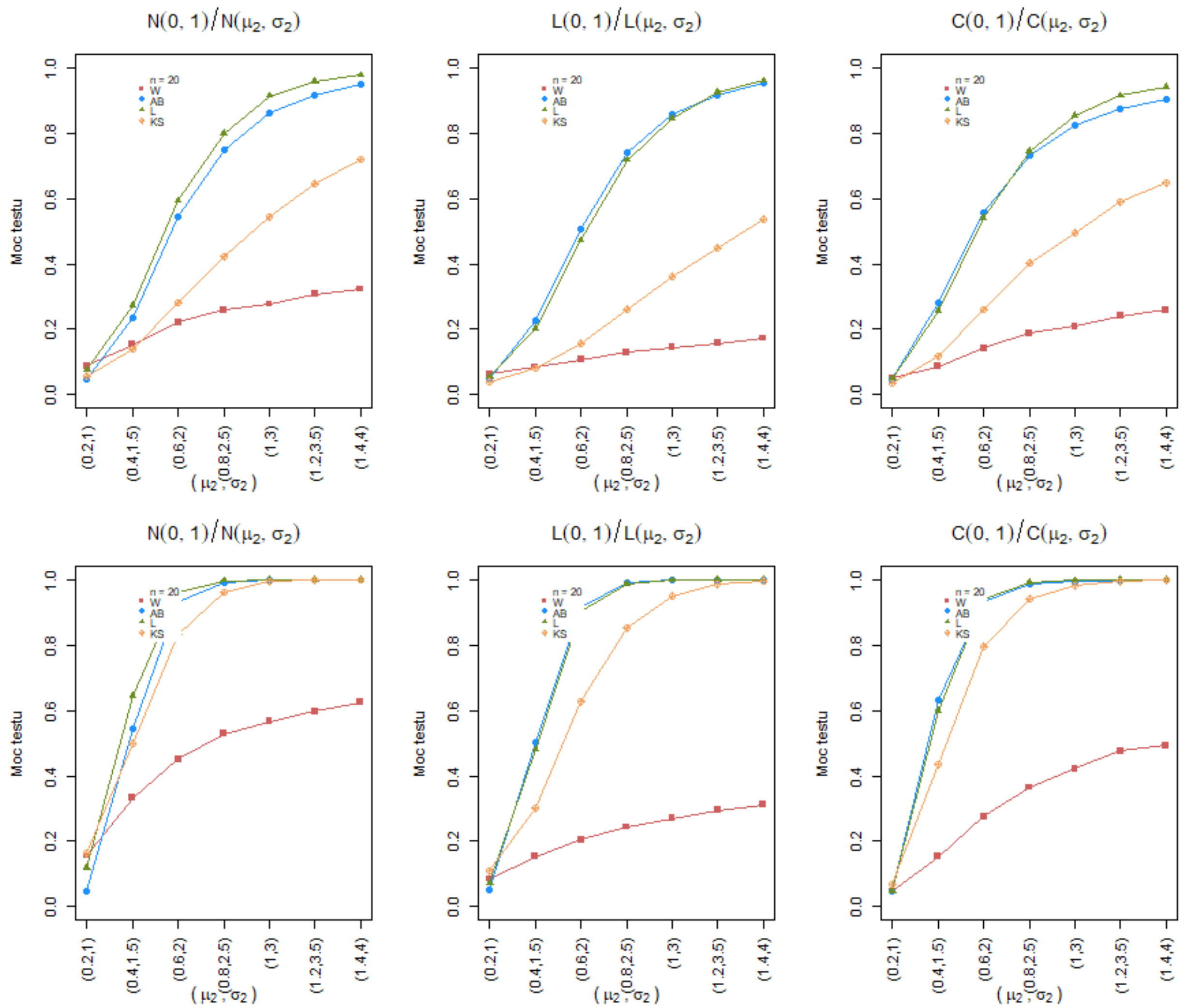
Natomiast test Kołmogorowa-Smirnowa wypadł gorzej, gdy zmieniał się parametr skali, co jest interesującym odwróceniem sytuacji w porównaniu do zmiany parametru przesunięcia. To sugeruje, że test ten może być bardziej wrażliwy na pewne typy zmian w rozkładzie, ale niekoniecznie na wszystkie rodzaje zmian.

Przy zwiększeniu rozmiaru próby (m = n = 50), obserwujemy podobne tendencje co dla m = n = 20. Test Ansari-Bradley'a wciąż wykazuje się jako najskuteczniejszy, a test Lepage'a plasuje się tuż za nim. Interesująco jest zauważyć, że test Kołmogorowa-Smirnowa uzyskuje wyższą skuteczność wraz ze wzrostem rozmiaru próby, co może sugerować jego większą moc w sytuacjach z większymi próbkami.

Jednakże test Wilcoxona nadal wykazuje się niższą skutecznością, szczególnie w kontekście zmiany parametru skali, co podkreśla jego ograniczenia w wykrywaniu pewnych rodzajów zmian w badanych rozkładach pomimo zwiększenia rozmiaru próby. Ta obserwacja wskazuje na istotność wyboru odpowiedniego testu statystycznego w zależności od konkretnych parametrów, które podlegają analizie.

Zadanie 4

Wygeneruję $m = n = 20$ oraz $m = n = 50$ obserwacji z rozkładów
(a) $N(\mu_1, \sigma_1)$ i $N(\mu_2, \sigma_2)$,
(b) $L(\mu_1, \sigma_1)$ i $L(\mu_2, \sigma_2)$,
(c) $C(\mu_1, \sigma_1)$ i $C(\mu_2', \sigma_2')$,
gdzie $\mu_1 = 0$, $\sigma_1 = 1$, $\sigma_2 \in \{1, 1.5, 2, 2.5, 3, 3.5, 4\}$, $\mu_2 \in \{0.2, 0.4, 0.6, 0.8, 1, 1.2, 1.4\}$ oraz $\sigma_2' \in \{1, 2, 3, 4, 5, 6, 7\}$, $\mu_2' \in \{0, 0.5, 1, 1.5, 2, 2.5, 3\}$. Następnie wyznaczę wartość statystyk $W.AB, L$ i KS. Doświadczenie powtórzę 10000 razy i na tej podstawie oszacuję wartość funkcji mocy analizowanych testów. Narysuję wyestymowane funkcje mocy w zależności od wektora parametrów (μ_2, σ_2) i (μ_2', σ_2').



Dla przypadku, gdy m = n = 20, zauważono spójne tendencje niezależnie od badanych rozkładów. Test Lepage'a oraz test Ansari-Bradley'a prezentują się jako najskuteczniejsze narzędzia. Ich wysoka skuteczność w wykrywaniu zmian w rozkładach wynika z ich wrażliwości na różne parametry rozkładu. Natomiast test Wilcoxona wypada najgorzej, co wynika z jego ograniczonej zdolności do reagowania na zmianę parametru skali, co było widoczne we wszystkich trzech rozważanych rozkładach. Test Kołmogorowa-Smirnowa zajmuje pozycję pośrednią, prezentując się lepiej niż test Wilcoxona, ale gorzej niż testy Lepage'a i Ansari-Bradley'a.

Gdy rozważamy większy rozmiar próby (m = n = 50), obserwujemy podobne trendy jak wcześniej. Moc testu Kołmogorowa-Smirnowa wzrasta wraz ze wzrostem rozmiaru próby, co potwierdza jego wyższą skuteczność w sytuacjach, gdy analizujemy większe ilości danych. Moc testu Wilcoxona również wzrosła w porównaniu do przypadku m = n = 20, ale nadal jest on najmniej skuteczny spośród analizowanych testów. Pomimo tego wzrostu skuteczności, test Wilcoxona nadal wykazuje swoje ograniczenia, szczególnie w kontekście wrażliwości na zmiany parametru skali.

Podsumowując, analiza dla obu rozmiarów prób (m = n = 20 oraz m = n = 50) potwierdza, że testy Lepage'a i Ansari-Bradley'a wypadają najlepiej niezależnie od rozpatrywanego rozkładu, ze względu na ich zdolność do wykrywania zmian w różnych parametrach rozkładu. Test Kołmogorowa-Smirnowa wykazuje wzrost skuteczności wraz ze zwiększeniem próby, co jest korzystne przy analizie większych zbiorów danych. Natomiast test Wilcoxona, mimo pewnego wzrostu skuteczności przy większym rozmiarze próby, nadal pozostaje najmniej skutecznym testem spośród analizowanych, szczególnie w kontekście zmiany parametru skali.