

Zanim przejdziemy do prezentacji zadań z raportu 5., przydatne będzie wprowadzenie poniższych oznaczeń.

Niech  $X_1, \dots, X_n$  będą niezależnymi zmiennymi losowymi z rozkładu o ciągłej dystrybucji  $F$ , a  $Y_1, \dots, Y_n$  będą niezależnymi zmiennymi losowymi z rozkładu o ciągłej dystrybucji  $G$ . Rozważać będziemy problem testowania hipotezy

$$H_0 : F = G \text{ przeciwko alternatywie } H_1 : F \neq G,$$

na poziomie istotności  $\alpha = 0.05$ . W tym celu korzystać będziemy z **testów Wilcoxona, Ansari-Bradley'a, Lepage'a** oraz **Kołmogorowa-Smirnowa**.

Niech  $N = m + n$ , a  $Z = (Z_1, \dots, Z_N)$  będzie wektorem połączonych w. w. prób. Niech  $R_i$  będzie rangą  $Z_i$  w próbce  $Z$ ,  $i = 1, \dots, N$ . Definiujemy

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

Jeżeli  $\varphi(u) = \sqrt{3}(2u - 1)$ , to  $T_\varphi^2 = W$  jest statystyką, na której opieramy test Wilcoxona. Jeżeli  $\varphi(u) = \sqrt{48}(0.25 - |u - 0.5|)$ , to  $T_\varphi^2 = AB$  jest statystyką, na której opieramy test Ansari-Bradley'a. Test Lepage'a opieramy na statystyce  $L = W + AB$ , zaś test Kołmogorowa-Smirnowa opieramy na statystyce

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

gdzie  $F_m, G_n$  są dystrybucjami empirycznymi w próbce  $X$ -ów i  $Y$ -ów odpowiednio. Do obliczania w.w. statystyk dla rozważanych prób 20-elementowych wykorzystywać będziemy następujące funkcje:

```
varphi1 <- function(x) {
  return(sqrt(3) * (2*x - 1))
}

varphi2 <- function(x) {
  return(sqrt(48) * (.25 - abs(x - .5)))
}

wilcox.stat.sqr <- function(X1, X2) {
  Z <- c(X1, X2)
  ranks.mod <- (rank(Z) - .5)/40
  return((sqrt(10) * (20^-1 * sum(sapply(ranks.mod[1:20], varphi1)) -
    20^-1 * sum(sapply(ranks.mod[21:40], varphi1))))^2)
}

wilcox.stat <- function(X1, X2) {
  Z <- c(X1, X2)
  ranks.mod <- (rank(Z) - .5)/40
  return(sqrt(10) * (20^-1 * sum(sapply(ranks.mod[1:20], varphi1)) -
```

```

                20^-1 * sum(sapply(ranks.mod[21:40], varphi1))))
}

ansbard.stat.sqr <- function(X1, X2) {
  Z <- c(X1, X2)
  ranks.mod <- (rank(Z) - .5)/40
  return((sqrt(10) * (20^-1 * sum(sapply(ranks.mod[1:20], varphi2)) -
    20^-1 * sum(sapply(ranks.mod[21:40], varphi2))))^2)
}

ansbard.stat <- function(X1, X2) {
  Z <- c(X1, X2)
  ranks.mod <- (rank(Z) - .5)/40
  return(sqrt(10) * (20^-1 * sum(sapply(ranks.mod[1:20], varphi2)) -
    20^-1 * sum(sapply(ranks.mod[21:40], varphi2))))
}

lepage.stat.sqr <- function(X1, X2) return(wilcox.stat.sqr(X1, X2) +
  ansbard.stat.sqr(X1, X2))

kolomsmir.stat <- function(X1, X2) return(sqrt(length(X1)*length(X2)
  /(length(X1) + length(X2))) *
  ks.test(X1, X2)$stat)

```

W przypadku testów **Wilcoxona** i **Ansari-Bradley’a**  $H_0$  odrzucać będziemy gdy  $|T_\varphi|$  będą większa od kwantyla rozkładu normalnego  $N(0, 1)$  rzędu  $1 - \alpha$ . Dla testu **Lepage’a**  $H_0$  odrzucać będziemy gdy  $|L|$  będzie większa od kwantyla rozkładu  $\chi^2_2$  rzędu  $1 - \alpha$ . W teście **Kolmogorowa-Smirnowa**  $H_0$  odrzucać będziemy gdy  $|KS|$  będzie większa od kwantyla rozkładu Kolmogorowa rzędu  $1 - \alpha$ .

**Zadanie 1.** Generujemy  $m = n = 20$  obserwacji z rozkładu  $N(0, 1)$ . Na ich podstawie obliczamy statystyki  $W, AB, L$  i  $KS$ . Otrzymujemy następujące wartości:

W	AB	L	KS
0.919	0.3	1.219	0.791

Doświadczenie powtarzamy 10000 razy. Na tej podstawie wyznaczamy wartości krytyczne opowiadających im testów **prawostronnych**.

```

w <- c()
ab <- c()
l <- c()
ks <- c()

for(i in c(1:10000)) {
  X1 <- rnorm(20)
  X2 <- rnorm(20)
  w[i] <- wilcox.stat(X1, X2)

```

```

ab[i] <- ansbard.stat(X1, X2)
l[i] <- lepage.stat.sqr(X1, X2)
ks[i] <- kolomsmir.stat(X1, X2)
}

W <- sort(w)[9500]
AB <- sort(ab)[9500]
L <- sort(l)[9500]
KS <- sort(ks)[9500]

```

Otrzymujemy następujące wartości:

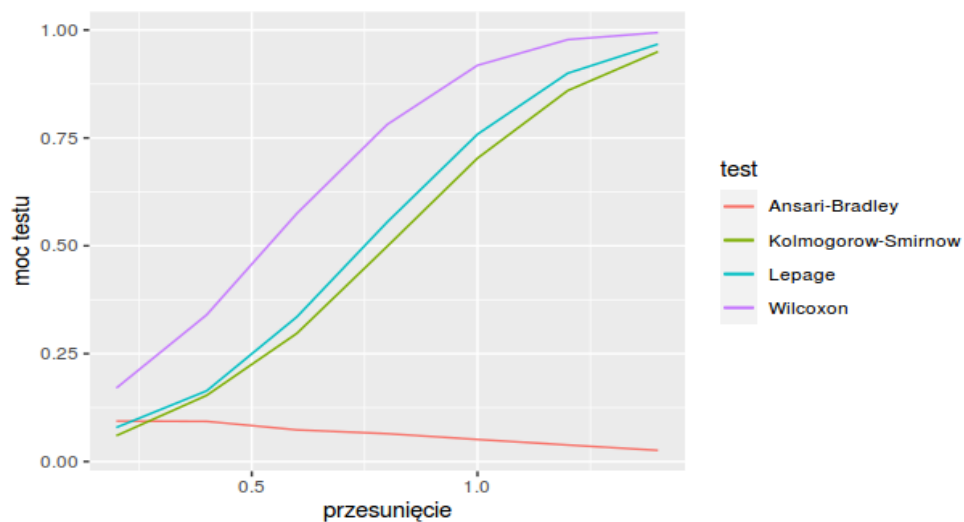
	$W$	$AB$	$L$	$KS$
estymator	1.638	1.698	5.968	1.265
kwantyl	1.645	1.645	5.991	1.36

Zauważmy, że są one zbliżone do wartości kwantyli rzędu  $1 - \alpha$  z odpowiednich rozkładów, co pozwala stwierdzić, że wartości krytyczne generowane w ten sposób są odpowiednie.

**Zadanie 2.** Celem tego zadania będzie zbadanie, który z wymienionych wyżej testów sprawdza się najlepiej w przypadku rozkładów **różniących się** wartością parametru **przesunięcia**. Zaczniemy od wygenerowania  $m = n = 20$  obserwacji z rozkładu **normalnego** z parametrem przesunięcia  $\mu_1 = 0, \mu_2$  oraz skali  $\sigma_1 = 1, \sigma_2 = 1$ . Otrzymujemy następujące statystyki:

	$W$	$AB$	$KS$	$L$
$\mu_2 = 0.2$	1.452	0.588	0.791	2.040
$\mu_2 = 0.4$	1.875	0.108	0.791	1.983
$\mu_2 = 0.6$	4.219	0.243	1.265	4.462
$\mu_2 = 0.8$	0.919	0.432	0.791	1.351
$\mu_2 = 1.0$	13.669	0.075	1.739	13.744
$\mu_2 = 1.2$	6.769	1.587	1.581	8.356
$\mu_2 = 1.4$	11.907	0.003	1.739	11.910

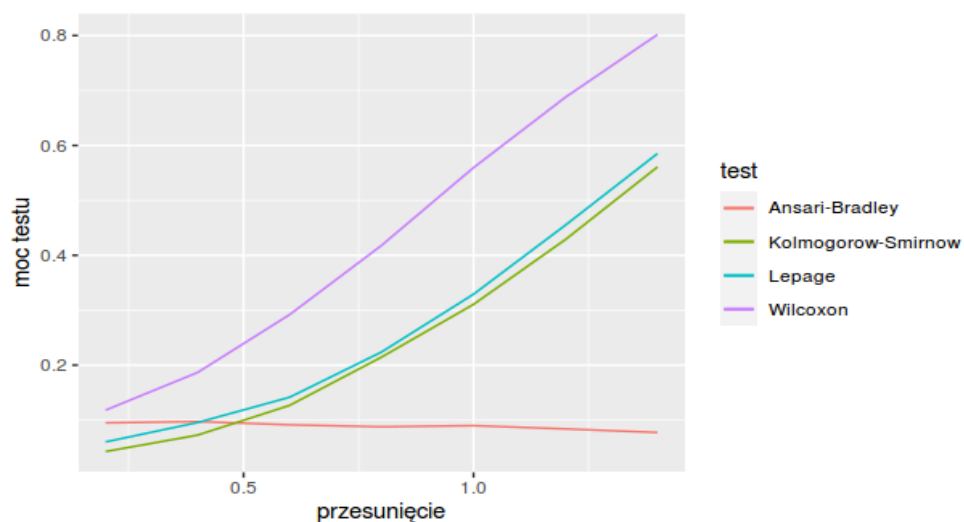
Powtarzamy doświadczenie 10000 razy. Na tej podstawie otrzymujemy następujące oszacowanie funkcji mocy analizowanych testów.



Powtarzamy powyższe doświadczenie dla rozkładu **logistycznego** z parametrem przesunięcia  $\mu_1 = 0, \mu_2$  oraz skali  $\sigma_1 = 1, \sigma_2 = 1$ . Dla próby  $m = n = 20$ -elementowej otrzymujemy następujące statystyki:

	$W$	$AB$	$KS$	$L$
$\mu_2 = 0.2$	0.972	1.875	0.791	2.847
$\mu_2 = 0.4$	0.169	2.700	0.791	2.869
$\mu_2 = 0.6$	1.587	6.627	1.265	8.214
$\mu_2 = 0.8$	0.432	3.072	0.791	3.504
$\mu_2 = 1.0$	1.387	2.187	0.949	3.574
$\mu_2 = 1.2$	8.748	0.588	1.581	9.336
$\mu_2 = 1.4$	0.469	2.352	0.791	2.821

Zaś oszacowana funkcja mocy testów wygląda następująco:

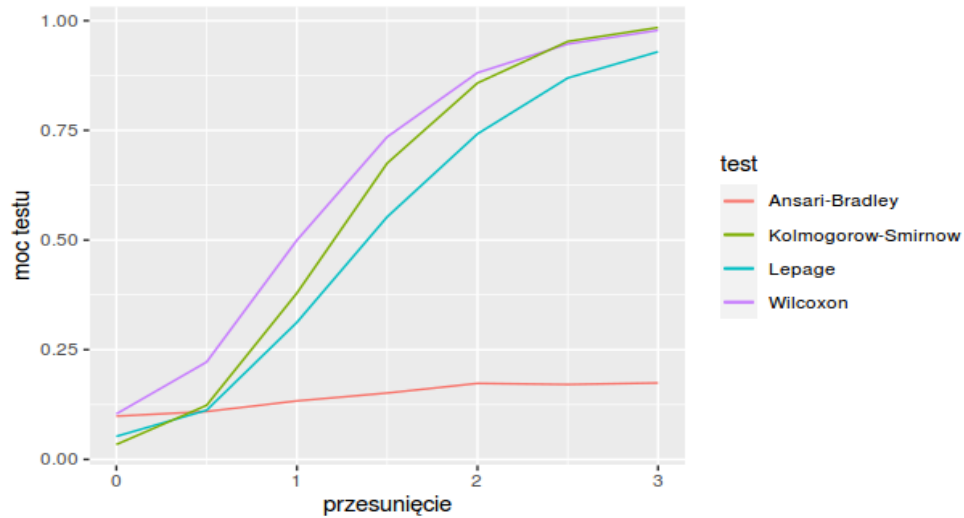


Na koniec powtarzamy powyższy eksperyment dla rozkładu **Cauchy'ego** z parametrem

przesunięcia  $\mu_1 = 0, \mu_2$  oraz skali  $\sigma_1 = 1, \sigma_2 = 1$ . Dla próby  $m = n = 20$ -elementowej otrzymujemy następujące statystyki:

	$W$	$AB$	$KS$	$L$
$\mu_2 = 0.0$	1.261	0.867	0.791	2.128
$\mu_2 = 0.5$	0.091	0.507	0.474	0.598
$\mu_2 = 1.0$	2.028	0.300	0.949	2.328
$\mu_2 = 1.5$	6.075	2.523	1.581	8.598
$\mu_2 = 2.0$	7.803	0.012	1.423	7.815
$\mu_2 = 2.5$	7.957	0.027	1.897	7.984
$\mu_2 = 3.0$	5.941	0.108	1.739	6.049

Zaś oszacowana funkcja mocy testów wygląda następująco:

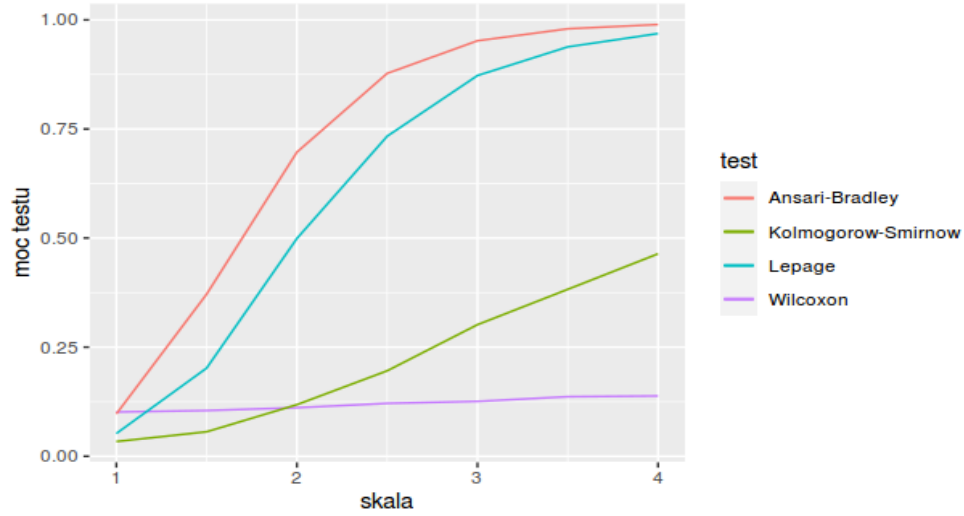


Możemy zauważyć, że w przypadku każdego z rozkładów najlepiej sprawdza się test Wilcoxona. Niewiele gorzej wypadają testy Kolmogorowa-Smirnowa i Lepage'a. Ponadto w przypadku każdego z tych testów, moc testu rośnie wraz ze wzrostem różnicy między parametrami przesunięcia. Test Ansari-Bradley'a natomiast nie sprawdza się dobrze. W szczególności dla rozkładów normalnego i logistycznego jego moc maleje wraz ze wzrostem różnicy między parametrami przesunięcia, co nie jest zjawiskiem pożądanym.

**Zadanie 3.** W tym zadaniu badać będziemy zachowanie wspomnianych wcześniej testów w przypadku rozkładów **różniących się** wartością **skali**. Zaczniemy od wygenerowania  $m = n = 20$  obserwacji z rozkładu **normalnego** z parametrem przesunięcia  $\mu_1 = 0, \mu_2 = 0$  oraz skali  $\sigma_1 = 1, \sigma_2$ . Otrzymujemy następujące statystyki:

	$W$	$AB$	$KS$	$L$
$\sigma_2 = 1.0$	4.332	1.323	1.581	5.655
$\sigma_2 = 1.5$	0.675	2.028	0.791	2.703
$\sigma_2 = 2.0$	1.452	7.803	1.265	9.255
$\sigma_2 = 2.5$	0.027	7.500	1.107	7.527
$\sigma_2 = 3.0$	2.352	10.800	1.423	13.152
$\sigma_2 = 3.5$	0.037	7.500	0.949	7.375
$\sigma_2 = 4.0$	1.728	13.872	1.581	15.600

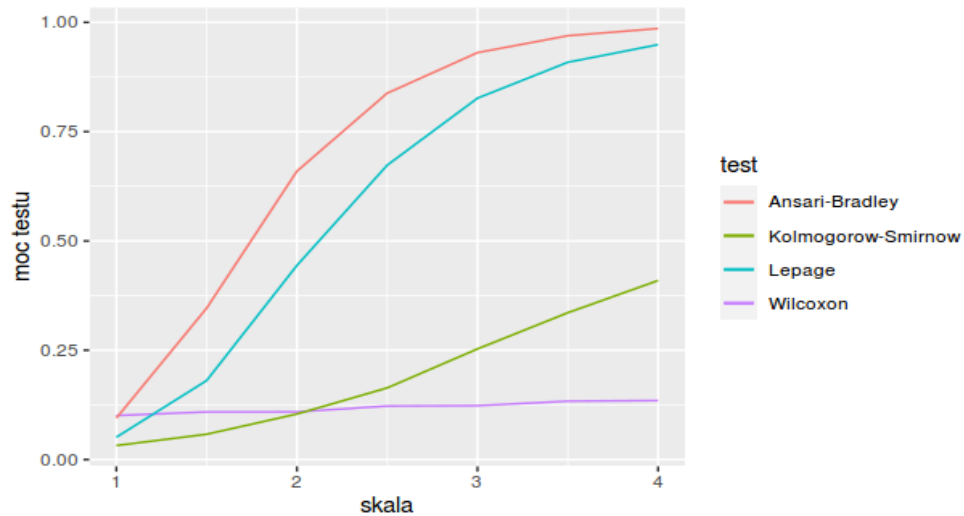
Powtarzamy doświadczenie 10000 razy. Na tej podstawie otrzymujemy następujące oszacowanie funkcji mocy analizowanych testów.



Powtarzamy powyższe doświadczenie dla rozkładu **logistycznego** z parametrem przesunięcia  $\mu_1 = 0, \mu_2 = 0$  oraz skali  $\sigma_1 = 1, \sigma_2$ . Dla próby  $m = n = 20$ -elementowej otrzymujemy następujące statystyki:

	$W$	$AB$	$KS$	$L$
$\sigma_2 = 1.0$	0.432	0.012	0.632	0.444
$\sigma_2 = 1.5$	1.951	2.352	1.107	4.303
$\sigma_2 = 2.0$	0.331	7.500	1.107	7.831
$\sigma_2 = 2.5$	0.271	11.163	1.107	11.434
$\sigma_2 = 3.0$	0.007	16.875	1.265	16.882
$\sigma_2 = 3.5$	0.001	11.907	1.265	11.908
$\sigma_2 = 4.0$	0.817	7.203	1.107	8.020

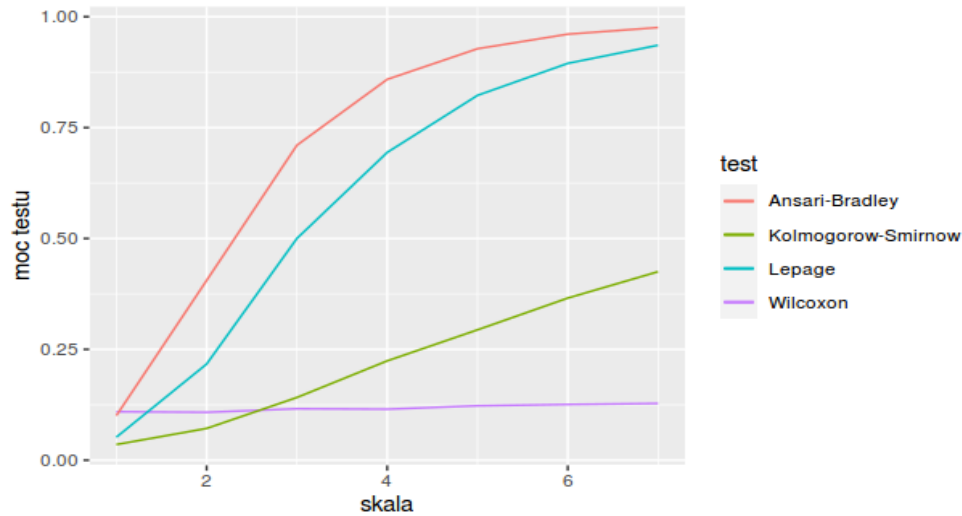
Zaś oszacowana funkcja mocy testów wygląda następująco:



Na koniec powtarzamy powyższy eksperyment dla rozkładu **Cauchy’ego** z parametrem przesunięcia  $\mu_1 = 0, \mu_2 = 0$  oraz skali  $\sigma_1 = 1, \sigma_2$ . Dla próby  $m = n = 20$ -elementowej otrzymujemy następujące statystyki:

	$W$	$AB$	$KS$	$L$
$\sigma_2 = 1.0$	2.352	0.075	1.107	2.427
$\sigma_2 = 2.0$	2.611	5.808	1.265	8.419
$\sigma_2 = 3.0$	0.075	0.768	0.791	0.843
$\sigma_2 = 4.0$	0.919	0.768	1.107	1.687
$\sigma_2 = 5.0$	1.519	1.587	1.107	3.106
$\sigma_2 = 6.0$	0.048	7.803	0.949	7.851
$\sigma_2 = 7.0$	0.972	9.747	1.265	10.719

Zaś oszacowana funkcja mocy testów wygląda następująco:

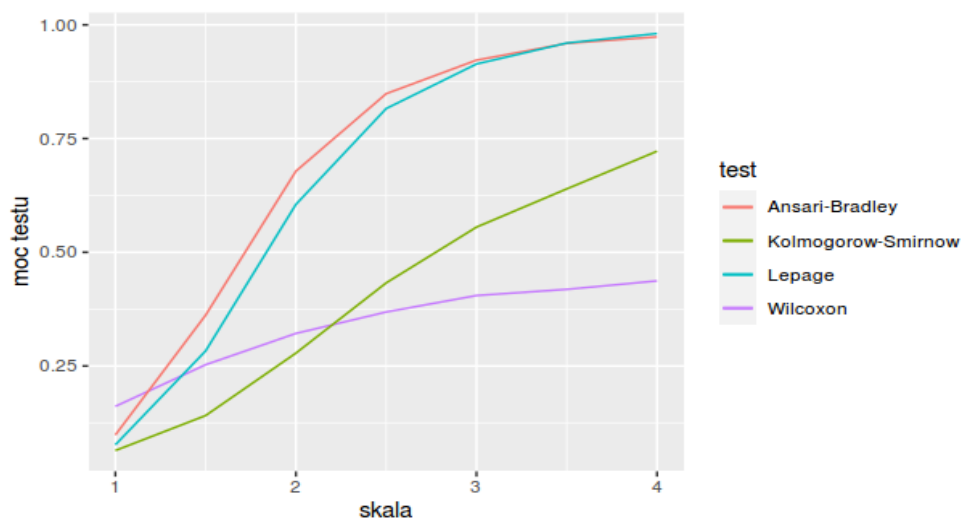


Możemy zauważyć, że w przypadku każdego z rozkładów najlepiej sprawdza się test Ansari-Bradley’a. Niewiele gorzej wypada test Lepage’a. Znacznie słabiej wypada test Kołmogorowa-Smirnowa, jednak również ma poprawne własności tj. jego funkcja mocy wzrasta wraz ze wzrostem różnicy skal. Test Wilcoxona natomiast nie sprawdza się dobrze. W przypadku każdego z rozkładów jego funkcja mocy jest w przybliżeniu funkcją stałą oraz nie przekracza wartości 0.25.

**Zadanie 4.** W tym zadaniu badać będziemy zachowanie wspomnianych wcześniej testów w przypadku rozkładów **różniących się** zarówno wartością **skali**, jak i wartością parametru **przesunięcia**. Zaczniemy od wygenerowania  $m = n = 20$  obserwacji z rozkładu **normalnego** z parametrem przesunięcia  $\mu_1 = 0, \mu_2$  oraz skali  $\sigma_1 = 1, \sigma_2$ . Otrzymujemy następujące statystyki:

	$W$	$AB$	$KS$	$L$
$\mu_2 = 0.2, \sigma_2 = 1.0$	0.397	2.883	0.791	3.280
$\mu_2 = 0.4, \sigma_2 = 1.5$	5.677	1.452	1.423	7.129
$\mu_2 = 0.6, \sigma_2 = 2.0$	2.791	8.112	1.265	10.903
$\mu_2 = 0.8, \sigma_2 = 2.5$	3.997	7.500	1.581	11.497
$\mu_2 = 1.0, \sigma_2 = 3.0$	1.387	6.348	1.265	7.735
$\mu_2 = 1.2, \sigma_2 = 3.5$	2.977	13.467	1.423	16.444
$\mu_2 = 1.4, \sigma_2 = 4.0$	0.397	12.675	1.265	13.072

Powtarzamy doświadczenie 10000 razy. Na tej podstawie otrzymujemy następujące oszacowanie funkcji mocy analizowanych testów.

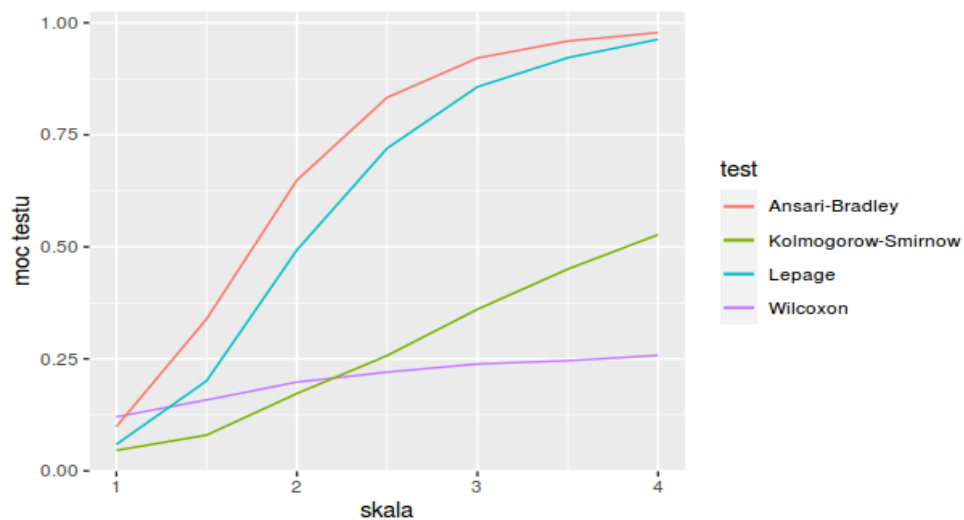


Powtarzamy powyższe doświadczenie dla rozkładu **logistycznego** z parametrem przesunięcia  $\mu_1 = 0, \mu_2$  oraz skali  $\sigma_1 = 1, \sigma_2$ . Dla próby  $m = n = 20$ -elementowej otrzymujemy następujące statystyki:

	$W$	$AB$	$KS$	$L$
$\mu_2 = 0.2, \sigma_2 = 1.0$	2.977	0.075	0.949	3.052
$\mu_2 = 0.4, \sigma_2 = 1.5$	0.127	0.507	0.632	0.634
$\mu_2 = 0.6, \sigma_2 = 2.0$	1.141	0.588	0.949	1.729
$\mu_2 = 0.8, \sigma_2 = 2.5$	11.347	5.043	2.214	16.390
$\mu_2 = 1.0, \sigma_2 = 3.0$	0.037	3.468	0.949	3.505
$\mu_2 = 1.2, \sigma_2 = 3.5$	5.167	5.043	1.739	10.210
$\mu_2 = 1.4, \sigma_2 = 4.0$	0.397	4.800	0.949	5.197

Zaś oszacowana funkcja mocy testów wygląda następująco:

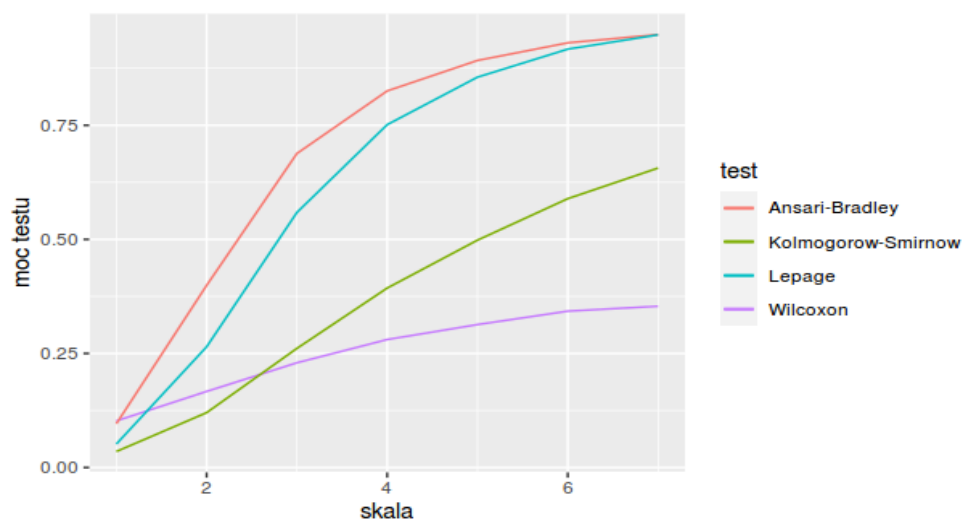




Na koniec powtarzamy powyższy eksperyment dla rozkładu **Cauchy'ego** z parametrem przesunięcia  $\mu_1 = 0, \mu_2$  oraz skali  $\sigma_1 = 1, \sigma_2$ . Dla próby  $m = n = 20$ -elementowej otrzymujemy następujące statystyki:

	$W$	$AB$	$KS$	$L$
$\mu_2 = 0.0, \sigma_2 = 1.0$	0.972	1.728	0.791	2.700
$\mu_2 = 0.5, \sigma_2 = 2.0$	1.801	5.292	1.107	7.093
$\mu_2 = 1.0, \sigma_2 = 4.0$	0.007	13.068	1.107	13.075
$\mu_2 = 1.5, \sigma_2 = 4.0$	0.147	15.987	1.107	16.134
$\mu_2 = 2.0, \sigma_2 = 5.0$	5.043	10.800	1.581	15.843
$\mu_2 = 2.5, \sigma_2 = 6.0$	3.267	13.068	1.739	16.335
$\mu_2 = 3.0, \sigma_2 = 7.0$	0.127	17.328	1.423	17.455

Zaś oszacowana funkcja mocy testów wygląda następująco:



Możemy zauważyć, że w przypadku każdego z rozkładów, podobnie jak w poprzednim zadaniu,

niu, najlepiej sprawdza się test Ansari-Bradley’a oraz nieznacznie gorzej wypada test Lepage’a. Test Kołmogorowa-Smirnowa znów wypada dosyć słabiej, ale znacznie lepiej niż w zadaniu 3. Jego funkcja mocy nadal wzrasta wraz ze wzrostem różnicy skal i przesunięć. Test Wilcoxona znów sprawdza się najslabiej, jednak podobnie jego funkcja mocy jest wyższa niż w zadaniu 3. Dodatkowo w tym przypadku w szczególności jest ona funkcją rosnącą. Możemy podejrzewać, na podstawie obserwacji z zadania 1., że na poprawę zachowania się testów Kołmogorowa-Smirnowa oraz Wilcoxona ma wpływ zastosowanie rozkładów o różnych parametrach przesunięcia.

**Zadanie 5.** Generujemy  $m = n = 50$  obserwacji z rozkładu  $N(0, 1)$ . Na ich podstawie, przy użyciu zmodyfikowanych funkcji używanych wcześniej dla prób  $m = n = 20$ -elementowych, obliczamy statystyki  $W, AB, L$  i  $KS$ . Otrzymujemy następujące wartości:

W	AB	L	KS
0.786	10.830	11.616	0.700

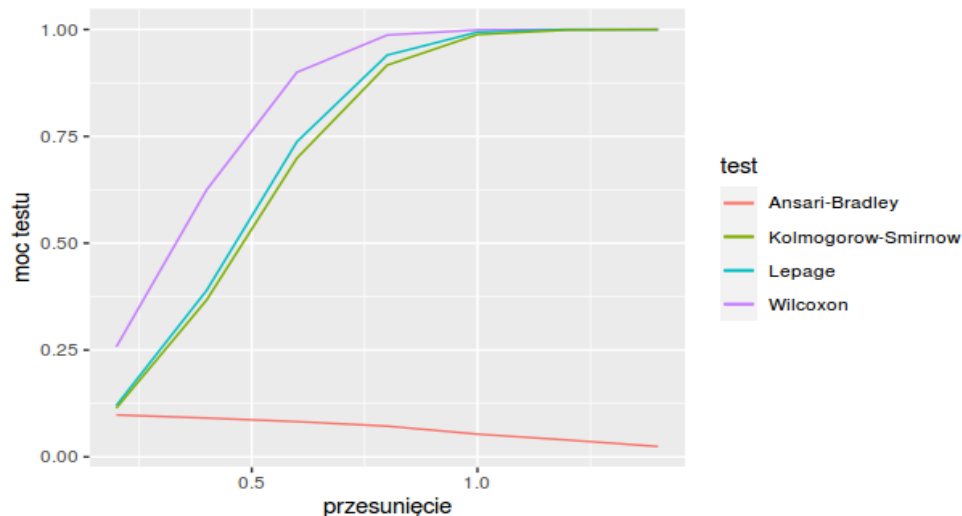
Doświadczenie powtarzamy 10000 razy. Na tej podstawie wyznaczamy wartości krytyczne opowiadających testów **prawostronnych**. Otrzymujemy następujące wartości:

	$W$	$AB$	$L$	$KS$
estymator	1.642	1.649	6.050	1.300
kwantyl	1.645	1.645	5.991	1.36

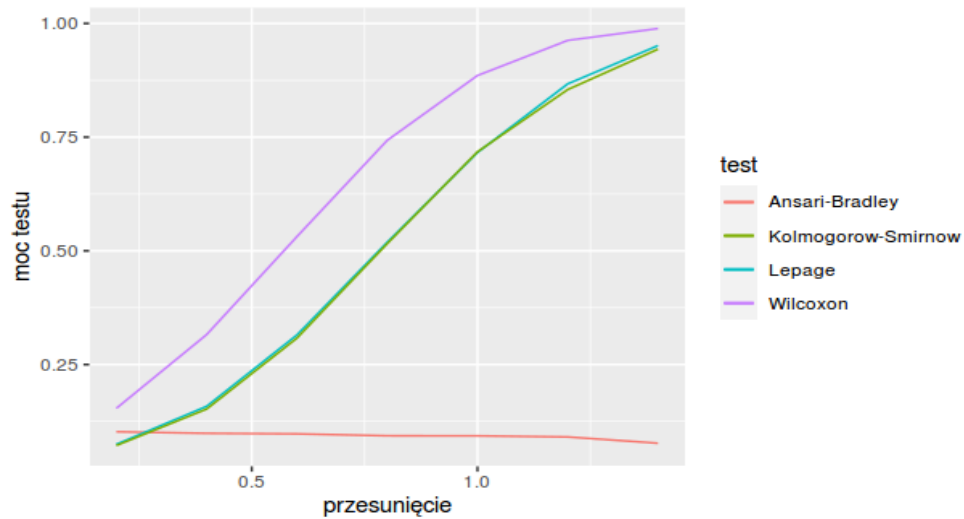
Zauważmy, że znów są one zbliżone do wartości kwantyli rzędu  $1 - \alpha$  z odpowiednich rozkładów. Ponadto są one bliższe teoretycznym kwantylom, niż w przypadku prób  $m = n = 20$ -elementowych.

**Zadanie 6.** W tym zadaniu powtórzmy eksperymenty dotyczące szacowania wartości funkcji mocy omawianych testów dla rozkładów z zadań 2, 3, 4 i prób  $m = n = 50$ -elementowych.

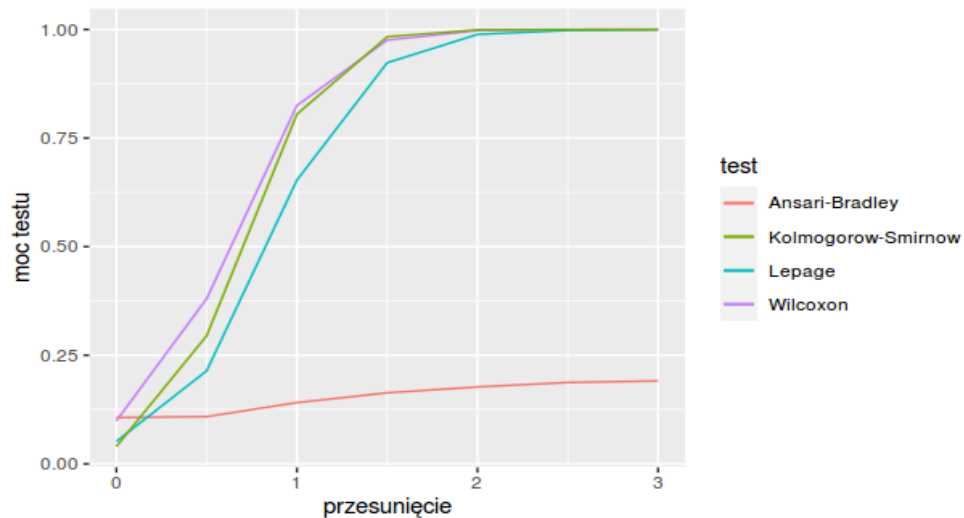
Zacniemy od rozkładów z zadania 2. Dla rozkładu **normalnego** z parametrem przesunięcia  $\mu_1 = 0, \mu_2$  oraz skali  $\sigma_1 = 1, \sigma_2 = 1$  otrzymujemy:



Następnie dla rozkładu **logistycznego** z parametrem przesunięcia  $\mu_1 = 0, \mu_2$  oraz skali  $\sigma_1 = 1, \sigma_2 = 1$  otrzymujemy:

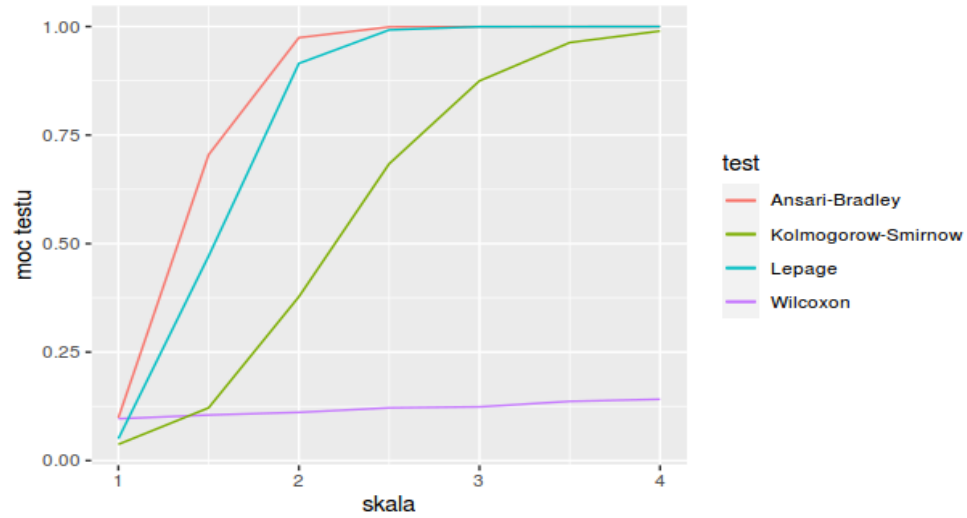


Dla rozkładu **Cauchy'ego** z parametrem przesunięcia  $\mu_1 = 0, \mu_2$  oraz skali  $\sigma_1 = 1, \sigma_2 = 1$  otrzymujemy zaś:

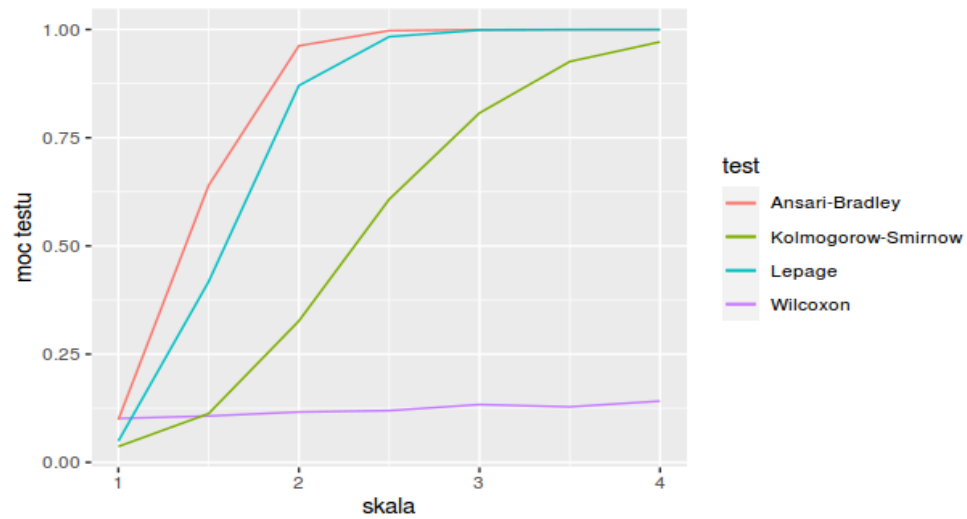


Widzimy, że kształty funkcji są zbliżone o tych uzyskanych dla prób  $m = n = 20$ -elementowych, natomiast dla większej próby ich wzrost dla testów Wilcoxona, Kołmogorowa-Smirnowa oraz Lepage'a jest szybszy. Ponadto w tym przypadku znów największe wartości osiąga funkcja mocy dla testu Wilcoxona.

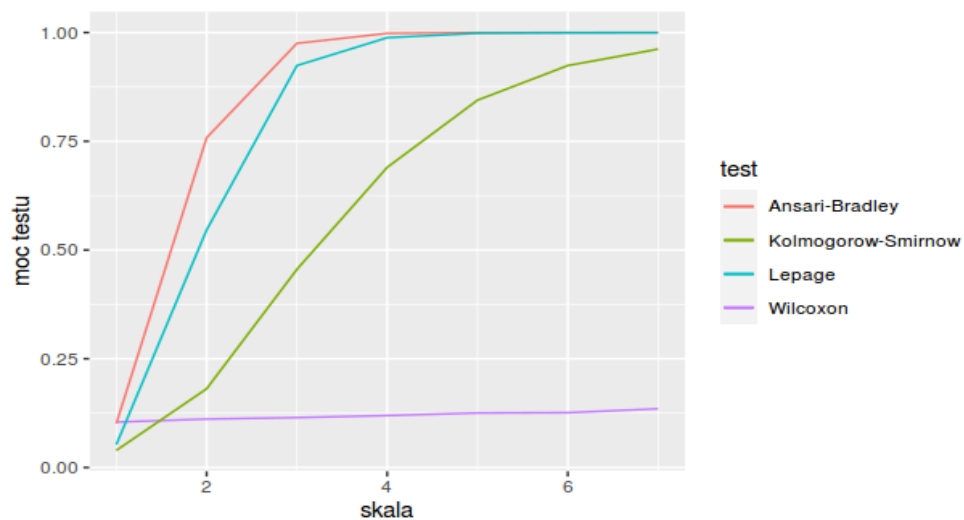
Przejdziemy do rozkładów z zadania 3. Dla rozkładu **normalnego** z parametrem przesunięcia  $\mu_1 = 0, \mu_2 = 0$  oraz skali  $\sigma_1 = 1, \sigma_2$  otrzymujemy:



Następnie dla rozkładu **logistycznego** z parametrem przesunięcia  $\mu_1 = 0, \mu_2 = 0$  oraz skali  $\sigma_1 = 1, \sigma_2$  otrzymujemy:

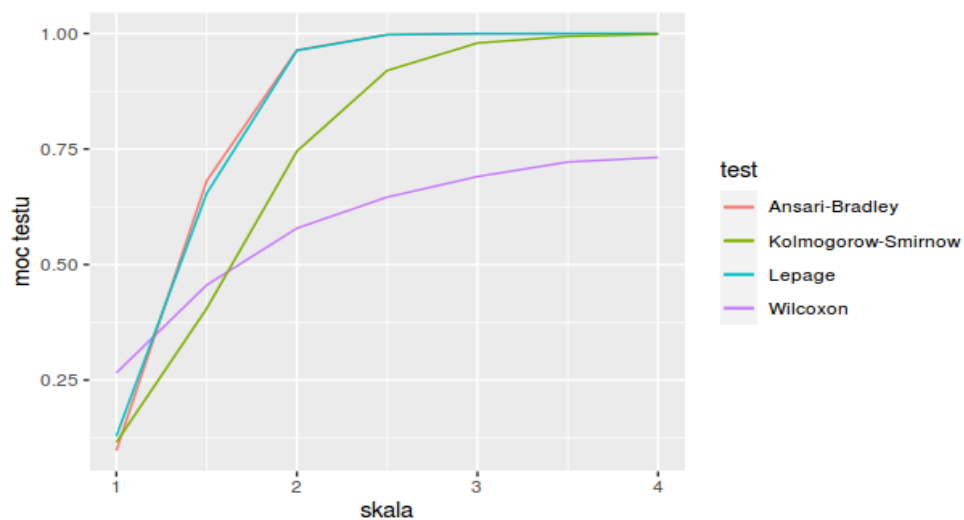


Dla rozkładu **Cauchy'ego** z parametrem przesunięcia  $\mu_1 = 0, \mu_2 = 0$  oraz skali  $\sigma_1 = 1, \sigma_2$  otrzymujemy zaś:

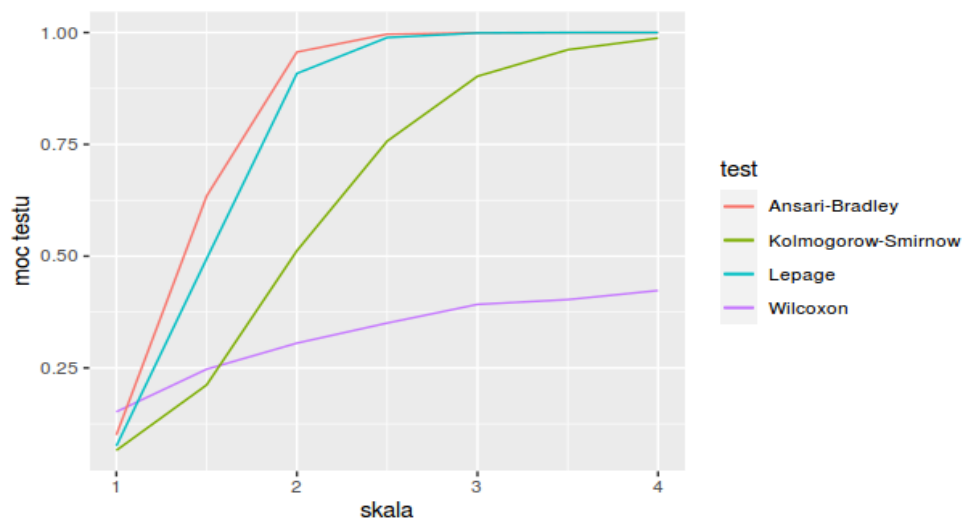


Widzimy, że kształty funkcji znów są zbliżone o tych uzyskanych dla prób  $m = n = 20$ -elementowych, natomiast dla większej próby ich wzrost dla testów Ansari-Bradley'a, Kołmogorowa-Smirnowa oraz Lepage'a jest szybszy. Ponadto w tym przypadku znów największe wartości osiąga funkcja mocy dla testu Ansari-Bradley'a.

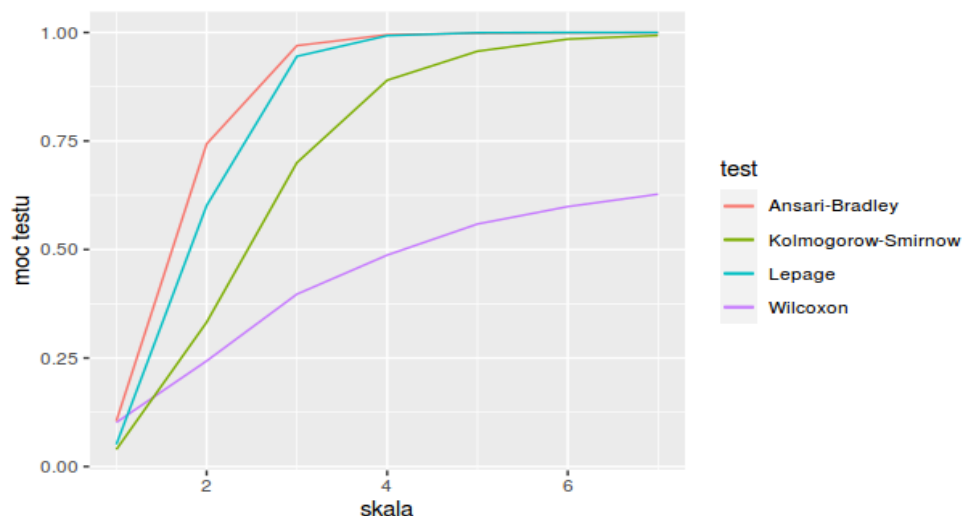
Przejdziemy do rozkładów z zadania 4. Dla rozkładu **normalnego** z parametrem przesunięcia  $\mu_1 = 0, \mu_2$  oraz skali  $\sigma_1 = 1, \sigma_2$  otrzymujemy:



Następnie dla rozkładu **logistycznego** z parametrem przesunięcia  $\mu_1 = 0, \mu_2$  oraz skali  $\sigma_1 = 1, \sigma_2$  otrzymujemy:



Dla rozkładu **Cauchy’ego** z parametrem przesunięcia  $\mu_1 = 0, \mu_2$  oraz skali  $\sigma_1 = 1, \sigma_2$  otrzymujemy zaś:



Widzimy, że tutaj funkcje mocy testów zachowują się analogicznie do tych dla prób  $m = n = 20$ -elementowych przy czym szybciej rosną.

Podsumowując możemy stwierdzić, że wraz ze wzrostem liczebności prób, wzrasta moc opowiednich testów, a zatem w szczególności ich poprawność. Do porównywania rozkładów o **różnych parametrach przesunięcia i równych skalach** najlepiej nadaje się test **Wilcoxona**, zaś nie sprawdza się zupełnie test **Ansari-Bradley’a**. Do porównywania rozkładów o **równych parametrach przesunięcia i różnych skalach** najlepiej nadaje się test **Ansari-Bradley’a**, zaś nie sprawdza się zupełnie test **Wilcoxona**. W przypadku porównywania rozkładów o **różnych parametrach przesunięcia i różnych skalach** najlepiej sprawdzają się test **Ansari-Bradley’a** oraz test **Lepage’a**, zaś niezbyt sprawdza się test **Wilcoxona**.