

Wybrane testy normalności

I'm not insane, my mother had me tested, Sheldon Cooper

Założenie o rozkładzie normalnym zmiennych jest podstawą wielu klasycznych technik modelowania (przykładowo analiza wariancji, regresja gaussowska, modele mieszane). Stąd też popularność technik testowania tego założenia.

W tym dokumencie przedstawiamy czternaście przykładowych testów normalności podzielonych na pięć grup, różniących się sposobem konstrukcji statystyki testowej. Poniżej przedstawiono testy, które mnie osobiście z różnych powodów interesowały, ale to nie są wszystkie testy do badania normalności. Moim celem było pokazanie mnogości podejść do tematu badania zgodności z rodziną rozkładów normalnych a nie wykonanie pełnego przeglądu testów normalności.

Ten dokument był częścią materiałów przygotowanych w ramach szkolenia o testach dla KNFu, a po uzupełnieniu i za zgodą został przekazany fundacji SmarterPoland.pl. Za pomoc w uzupełnieniu i rozbudowaniu pierwszej wersji dziękuję Krzysztofowi Trajkowskiemu, który jak zwykle wskazał wiele punktów, które warto poprawić i testów o których warto napisać.

Mam nadzieję, że dokument okaże się przydatnym. Wszelkie komentarze, sugestie propozycje do tego dokumentu proszę zgłaszać na adres Przemyslaw.Biecek@gmail.com. Z dokumentu można korzystać na zasadzie uznania autorstwa, cytując:

Przemysław Biecek, Wybrane testy normalności, 2013. Materiały Fundacji SmarterPoland.pl.

Wprowadzenie

Interesuje nas testowanie zgodności z rodziną rozkładów normalnych. Obserwujemy próbę prostą n obserwacji x_i , $i = 1, \dots, n$, niezależnych realizacji zmiennej losowej o rozkładzie F . Interesuje nas weryfikacja hipotezy

$$H_0 : \exists_{\mu, \sigma^2} F = \mathcal{N}(\mu, \sigma^2),$$

przeciwko hipotezie

$$H_1 : \neg \exists_{\mu, \sigma^2} F = \mathcal{N}(\mu, \sigma^2).$$

Ponieważ parametrów μ i σ^2 najczęściej nie znamy, więc testowana hipoteza zerowa jest hipotezą złożoną, co nastręcza pewnych trudności. W poniżej opisanych testach najczęściej te parametry są estymowane z próby, z użyciem być może niestandardowych estymatorów, a statystyka testowa odpowiada badaniu zgodności z hipotezą prostą

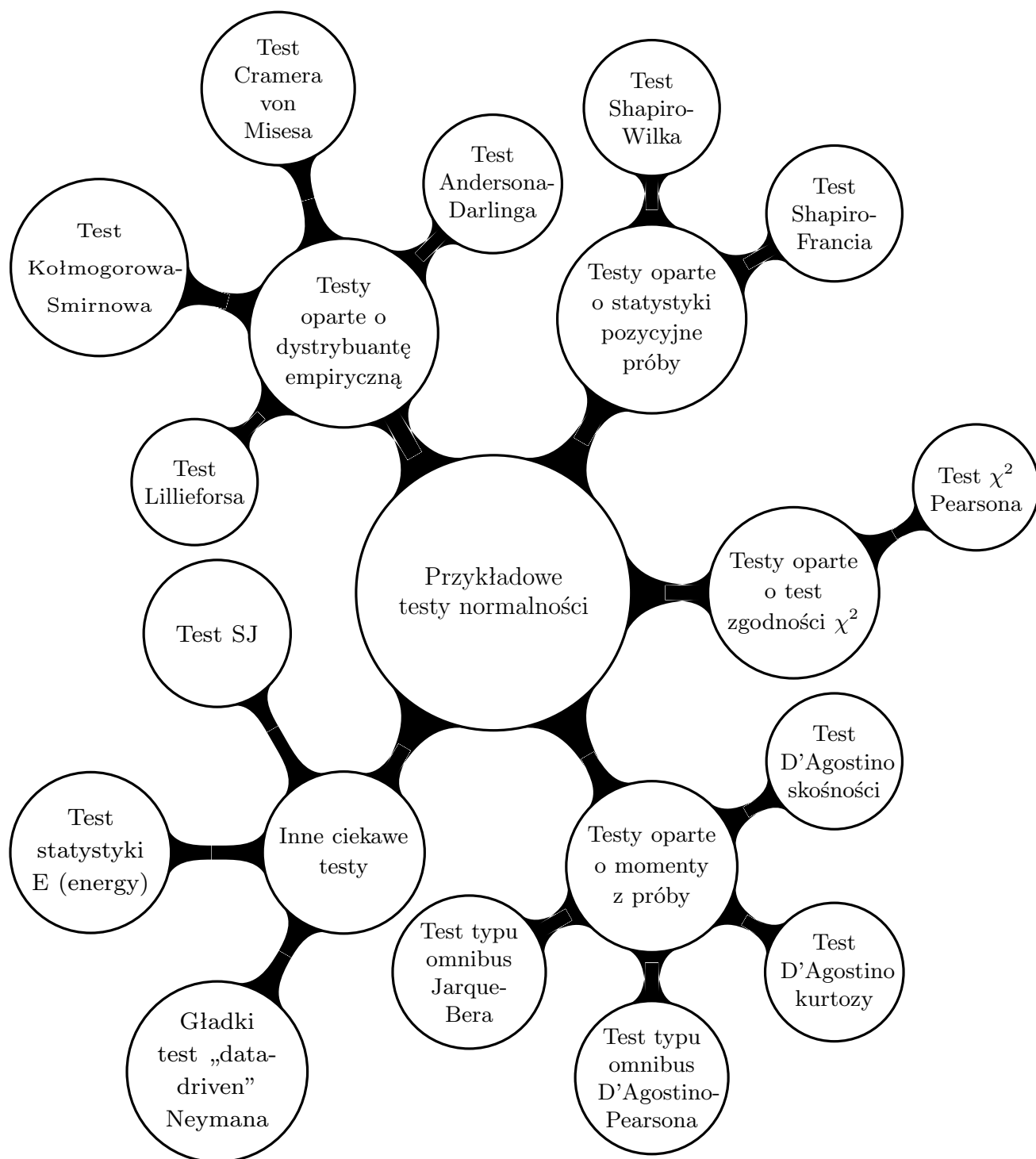
$$H_0 : F = \mathcal{N}(\hat{\mu}, \hat{\sigma}^2).$$

Taka zamiana hipotezy złożonej na prostą ma wpływ na rozkłady estymatorów statystyk testowych.

Poniżej przedstawiamy różne testy do weryfikacji hipotezy o zgodności z rodziną rozkładów normalnych, ale wiele z prezentowanych metod może być zaadoptowanych do zagadnienia testowania zgodności z innymi rodzinami rozkładów.

Dla większości testów przedstawiamy też przykładowe alternatywy, dla których dany test ma największą moc. Wybrane scenariusze alternatyw to jednak kropla w zbiorze możliwych odstępstw od normalności.

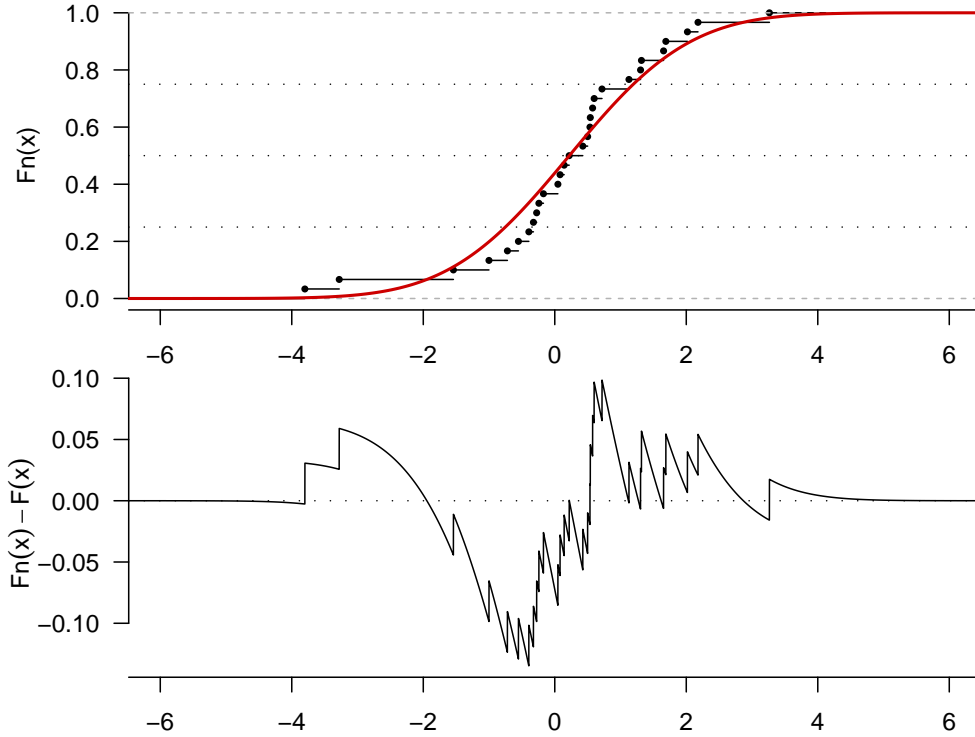
Naturalnym pytaniem jest „którego testu używać”, poniższe przykłady wykazują, że nie ma testu dobrego w każdym przypadku, ale o dwóch warto wspomnieć. Testem często stosowanym jest test Shapiro Wilka, popularny i zaimplementowany w większości pakietów statystycznych. Testem o szerokim spektrum wykrywania możliwych odstępstw jest test gładki test Neymana zaimplementowany w pakiecie R.



1 Testy zgodności oparte o dystrybuantę empiryczną

Testy wymienione w tym rozdziale badają zgodność z rozkładem normalnym przez ocenę odległości pomiędzy dystrybuantą empiryczną a dystrybuantą rozkładu normalnego.

Opisane w tym rozdziale testy określa się terminem „distribution free”, co oznacza, że rozkład statystyki testowej przy prawdziwej hipotezie zerowej nie zależy od rozkładu, z którym badana jest zgodność.



Rysunek 1: Ilustracja dystrybuanty empirycznej standaryzowanej próby (czarna krzywa) i dystrybuanty rozkładu $\mathcal{N}(0, 1)$ oraz różnica tych dwóch dystrybuant.

1.1 Test Cramera von Misesa

Ten test jest oparty o odległość Cramera von Misesa pomiędzy dystrybuantami empiryczną F_n i teoretyczną F

$$d(F_n, F) = n \int_{-\infty}^{\infty} (F_n(x) - F(x))^2 dF(x). \quad (1)$$

Statystykę testową dla próby x_i zbudowaną na powyższej odległości można zapisać jako

$$T = \frac{1}{12n} + \sum_{i=1}^n \left[\frac{2i-1}{2n} - F(x_i) \right]^2. \quad (2)$$

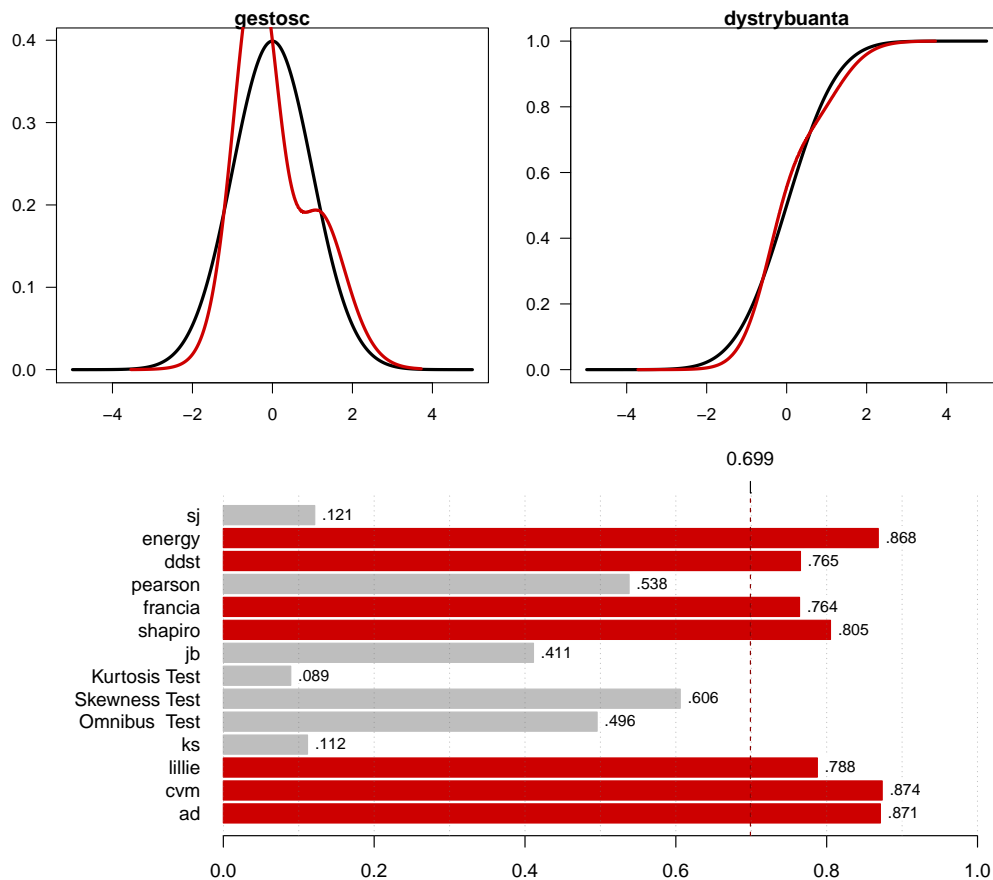
W przypadku badania zgodności z rozkładem normalnym dystrybuanta F jest zastępowana przez $\Phi(\hat{\mu}, \hat{\sigma}^2)$ dla oszacowanych parametrów średniej i wariancji. Rozkład statystyki przybliża się po uprzedniej modyfikacji [Stephens1986]

$$T^* = T(1.0 + 0.5/n). \quad (3)$$

Przyglądając się odległości $d(F_n, F)$ widzimy, że „zauważane” są różnice w centrum rozkładu, tam gdzie $dF(x)$ jest duża. Spodziewamy się „czułości” testu dla alternatyw zniekształconych w środku rozkładu.

W programie R ten test jest zaimplementowany w funkcji `cvm.test{nortest}`.

```
set.seed(1313); x <- rnorm(100)
library(nortest)
cvm.test(x)
#
#      Cramer-von Mises normality test
#
#data:  x
#W = 0.0607, p-value = 0.3665
#
cvm.test(x)$p.value
# 0.3664784
```



Rysunek 2: Moc vs. alternatywa Legendre’a rzędu 4 (opis tej klasy alternatyw jest w rozdziale 5), $n=100$, $\alpha = 0.05$. Na górnym panelu gęstość i dystrybuantę rozkładu dla którego badano moc (na czerwono) porównano z gęstością i dystrybuantą rozkładu normalnego (na czarno). Na dolnym panelu przedstawiono moc wybranych testów. Na czerwono zaznaczono wyniki nie gorsze niż 80% mocy najlepszego z testów (patrz pionowa czerwona linia). Mniej klasyczna alternatywa, ale też przykład w którym najlepiej wypadł test Cramera von Misesa. Moc oszacowana na bazie 10 000 powtórzeń.

1.2 Test Andersona-Darlinga

Ten test jest oparty o ważoną odległość Cramera von Misesa pomiędzy dystrybuantami empiryczną F_n i teoretyczną F z wagami odpowiadającymi odwrotności wariacji dystrybuanty empirycznej [Anderson1954] (pamiętamy, że $F_n(x)$ ma rozkład dwumianowy).

$$d(F_n, F) = n \int_{-\infty}^{\infty} \frac{(F_n(x) - F(x))^2}{[F(x)(1 - F(x))]} dF(x). \quad (4)$$

Statystykę testową opartą o powyższą odległość dla próby prostej x_i można przedstawić jako

$$\begin{aligned} S &= \sum_{k=1}^n \frac{2k-1}{n} [\ln(F(x_k)) + \ln(1 - F(x_{n+1-k}))], \\ A^2 &= -n - S. \end{aligned} \quad (5)$$

Podobnie jak w teście Cramera von Misesa, testując zgodność z rozkładem normalnym przyjmuje się $F = \Phi(\hat{\mu}, \hat{\sigma}^2)$, a rozkład statystyki testowej przybliża się po uprzedniej korekcie [Stephens1986] (w literaturze spotyka się też inne korekty, poniżej prezentowana jest używana w programie R)

$$A^{*,2} = A^2(1.0 + 0.75/n + 2.25/n^2). \quad (6)$$

W programie R ten test jest zaimplementowany w funkcji `ad.test{nortest}`.

```
set.seed(1313); x <- rnorm(100)
library(nortest)
ad.test(x)
#
#      Anderson-Darling normality test
#
#data:  x
#A = 0.4141, p-value = 0.3301
#
ad.test(x)$p.value
# [1] 0.3300575
```

1.3 Test Kołmogorowa-Smirnowa

Test oparty o odległość supremum pomiędzy dystrybucjami empiryczną F_n i teoretyczną F

$$d(F_n, F) = \sup_x |F_n(x) - F(x)|. \quad (7)$$

Statystyka testowa oparta o powyższą odległość sprowadza się do liczenia maksimum modułu różnicy dystrybucji w punktach skoku dystrybucji empirycznej

$$D = \max_{x_i} |F_n(x_i) - F(x_i)|. \quad (8)$$

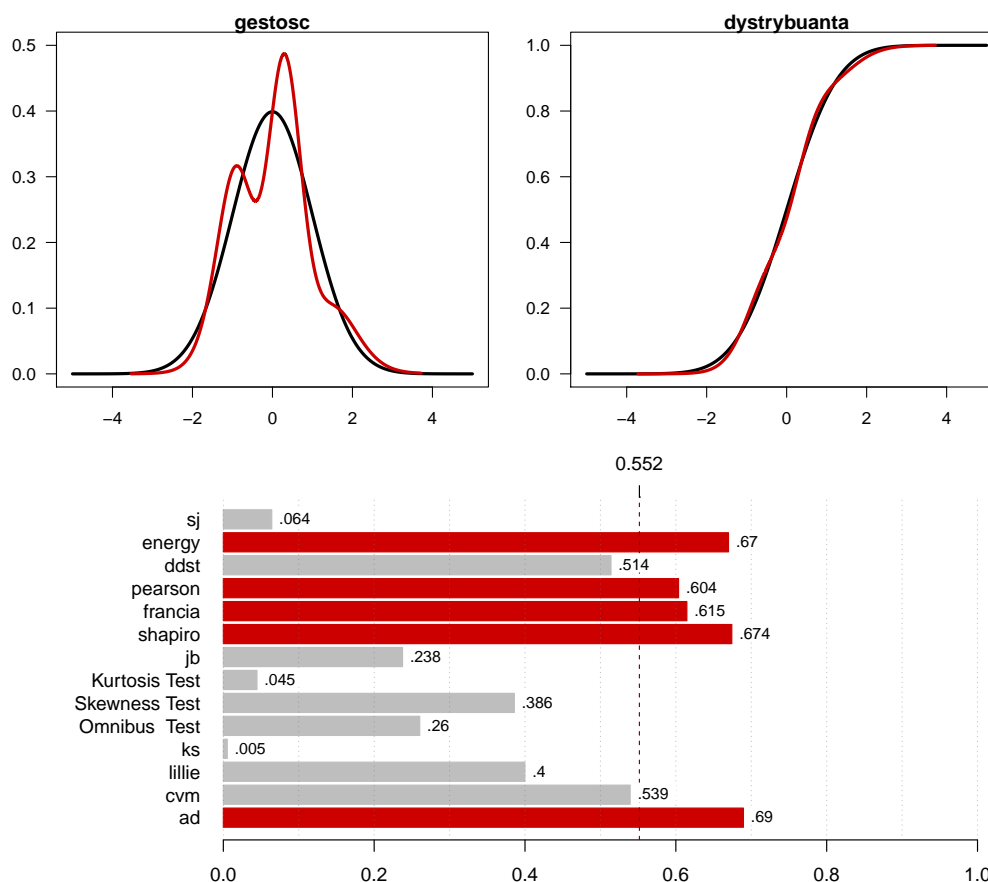
Rozważa się też ważoną wersję testu Kołmogorowa Smirnowa, bardziej czułą na różnice w ogonach

$$D' = \max_{x_i} \frac{|F_n(x_i) - F(x_i)|}{\sqrt{F(x_i)(1 - F(x_i))}}.$$

Rozkład statystyki testowej można wyznaczyć w sposób dokładny (rekurencyjnie) dla prostej hipotezy zerowej, a więc dla porównania z jednym określonym rozkładem. Asymptotycznie, ta statystyka przemnożona przez \sqrt{n} ma rozkład Kołmogorowa. Test ten pomimo łatwego opisu probabilistycznego nie jest stosowany z uwagi na moc niższą niż konkurencja.

W programie R ten test jest zaimplementowany w funkcji `ks.test{stats}`.

```
set.seed(1313)
x <- rnorm(100)
ks.test(scale(x), "pnorm")
#
#      One-sample Kolmogorov-Smirnov test
#
```



Rysunek 3: Moc vs. alternatywa Legendre'a rzędu 6 (opis tej klasy alternatyw jest w rozdziale 5), $n=250$, $\alpha = 0.05$. Na górnym panelu gęstość i dystrybuantę rozkładu dla którego badano moc (na czerwono) porównano z gęstością i dystrybuantą rozkładu normalnego (na czarno). Na dolnym panelu przedstawiono moc wybranych testów. Na czerwono zaznaczono wyniki nie gorsze niż 80% mocy najlepszego z testów (patrz pionowa czerwona linia). Z wykryciem tej alternatywy najlepiej poradził sobie test Andersona-Darlinga, zapewne przez cięższy prawy i lżejszy lewy ogon. Moc oszacowana na bazie 10 000 powtórzeń.

```
#data: scale(x)
#D = 0.0543, p-value = 0.9297
#alternative hypothesis: two-sided
#
ks.test(scale(x), "pnorm")$p.value
# [1] 0.9297479
```

1.4 Test Lillieforsa

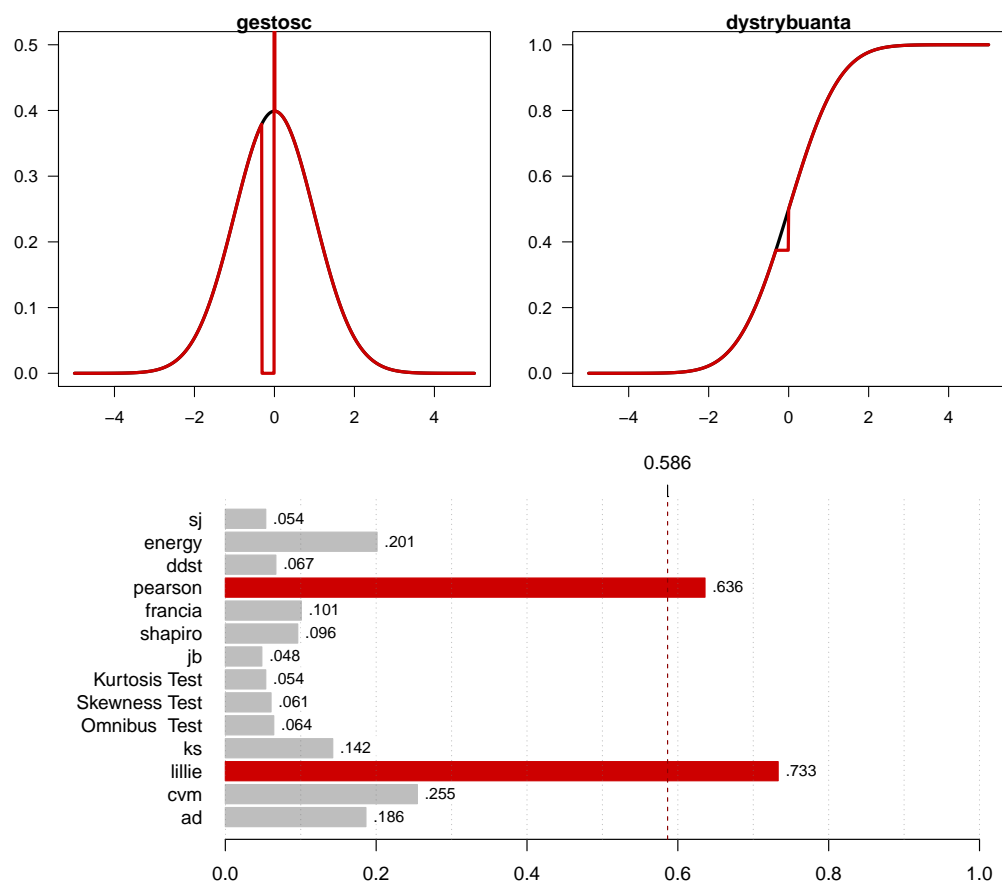
Ten test to modyfikacja testu Kołmogorowa-Smirnowa zaproponowana przez Huberta Lillieforsa [Lilliefors1967], pozwalająca na testowanie zgodności z całą rodziną rozkładów normalnych, bez znajomości parametrów średniej i odchylenia standardowego (test Kołmogorowa-Smirnowa pozwala na zbadanie zgodności z jednym określonym rozkładem).

Statystyka testowa w przypadku testu Lillieforsa wygląda tak samo jak w przypadku testu Kołmogorowa Smirnowa. Różnica polega na zastosowaniu innego rozkładu dla statystyki testowej (przybliżenie rozkładu dokładnego), uwzględniającego to, że hipoteza zerowa jest hipotezą złożoną.

W programie R ten test jest zaimplementowany w funkcji `lillie.test{nortest}` i `lillieTest{fBasics}`.

```
set.seed(1313); x <- rnorm(100)
lillie.test(x)$p.value
```

```
# [1] 0.6643097
lillieTest(x)$test$p.value
# 0.6643097
```



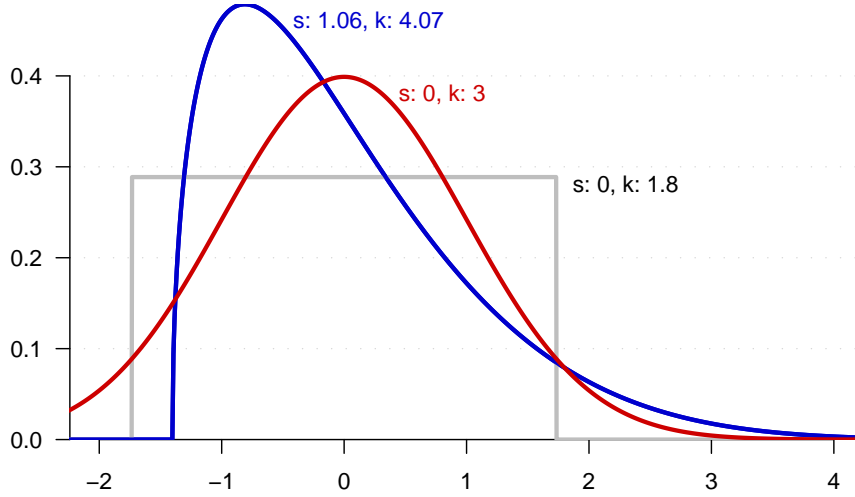
Rysunek 4: Moc vs. alternatywa w której część rozkładu skupiono w punkcie, $n=100$, $\alpha = 0.05$. Na górnym panelu gęstość i dystrybuantę rozkładu dla którego badano moc (na czerwono) porównano z gęstością i dystrybuantą rozkładu normalnego (na czarno). Na dolnym panelu przedstawiono moc wybranych testów. Na czerwono zaznaczono wyniki nie gorsze niż 80% mocy najlepszego z testów (patrz pionowa czerwona linia). Z wykryciem tej alternatywy najlepiej poradził sobie test Lillieforsa. Moc oszacowana na bazie 10 000 powtórzeń.

2 Testy oparte o momenty z próby

Testy oparte o dystrybucję empiryczną można było stosować do badania zgodności z dowolnym rozkładem. Poniższe testy wykorzystują specyficzną właściwość rozkładu normalnego, tzn. zerową skośność i kurtozę równą 3 (licząc zgodnie z definicją m_4/m_2^2) lub kurtozę równą 0 (licząc zgodnie z definicją $m_4/m_2^2 - 3$).

Aby łatwiej nam było opisać poniższe testy, przyjmijmy następujące oznaczenia na m_k to k-ty centralny moment z próby, skośność i kurtozę.

$$\begin{aligned} m_k &= \frac{1}{n} \sum_i (x_i - \bar{x})^k, \\ \sqrt{b_1} &= m_3/m_2^{3/2}, \\ b_2 &= m_4/m_2^2. \end{aligned} \tag{9}$$



Rysunek 5: Gęstości trzech przykładowych rozkładów, normalnego, jednostajnego (bardziej spłaszczony niż normalny) i lognormalnego (skośny)

2.1 Test D’Agostino oparty o skośność

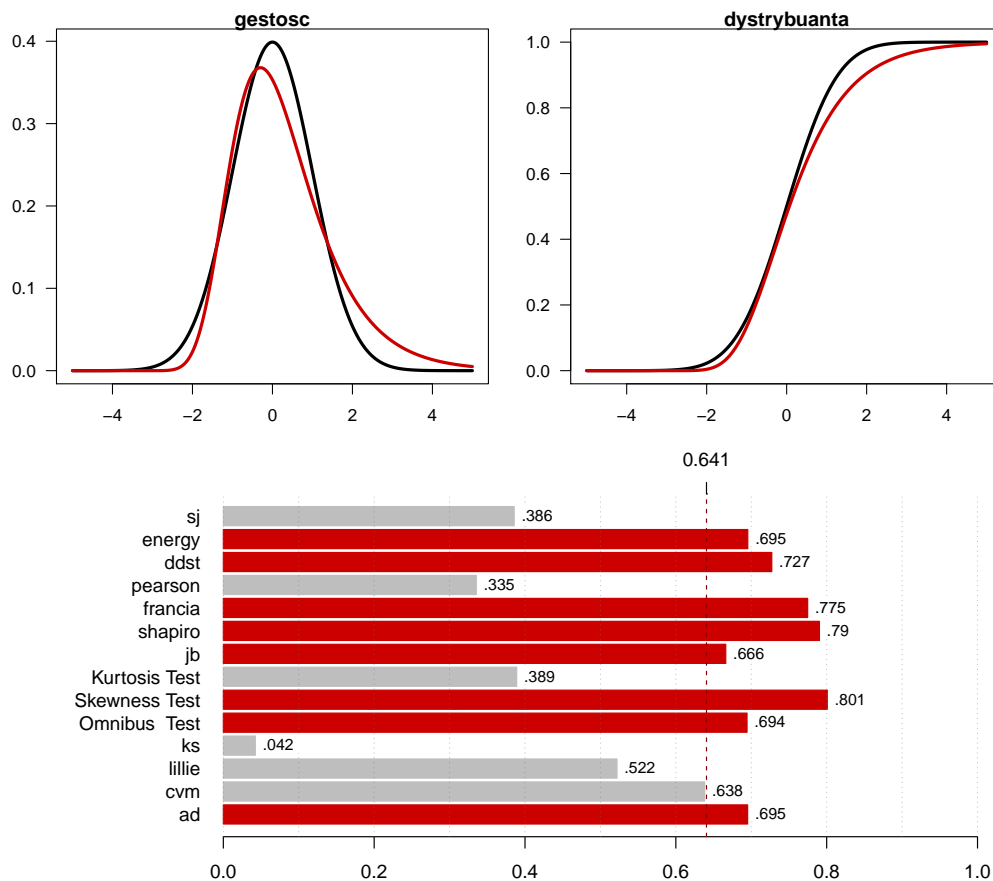
Test D’Agostino (1970) za statystykę testową wykorzystuje przekształconą próbkową skośność. W celu „normalizacji” rozkładu statystyki testowej stosuje się następujące przekształcenie

$$\begin{aligned} Y &= \sqrt{b_1} \left(\frac{(n+1)(n+3)}{6(n+2)} \right)^{1/2}, \\ \beta_2(\sqrt{b_1}) &= \frac{3(n^2+27n-70)(n+1)(n+3)}{(n-2)(n+5)(n+7)(n+9)}, \\ W^2 &= -1 + (2(\beta_2(\sqrt{b_1}) - 1))^{1/2}, \\ \delta &= 1/\sqrt{\ln W}, \\ \alpha &= (2/(W^2 - 1))^{1/2}, \\ Z(\sqrt{b_1}) &= \delta \ln \left(Y/\alpha + ((Y/\alpha)^2 + 1)^{1/2} \right). \end{aligned} \tag{10}$$

Statystyką testową jest $Z(\sqrt{b_1})$. Przy założeniu prawdziwej hipotezy zerowej ta statystyka ma asymptotyczny rozkład normalny [Agostino1970].

W programie R ten test jest zaimplementowany w funkcji `dagoTest{fBasics}`.

```
set.seed(1313); x <- rnorm(100)
dagoTest(x)@test$p.value[2]
# Skewness Test
# 0.9287743
```



Rysunek 6: Moc vs. rozkład Gumbela (wartości ekstremalnej), $n=60$, $\alpha = 0.05$. Na górnym panelu gęstość i dystrybuantę rozkładu dla którego badano moc (na czerwono) porównano z gęstością i dystrybuantą rozkładu normalnego (na czarno). Na dolnym panelu przedstawiono moc wybranych testów. Na czerwono zaznaczono wyniki nie gorsze niż 80% mocy najlepszego z testów (patrz pionowa czerwona linia). Ciężki prawy ogon rozkładu wartości ekstremalnej najlepiej wykrywa test skośności D’Agostino. Moc oszacowana na bazie 10 000 powtórzeń.

2.2 Test D’Agostino oparty o kurtozę

Test D’Agostino za statystykę testową wykorzystuje przekształconą próbkową kurtozę.

Bazując na wynikach [Anscombe1983] w celu „normalizacji” rozkładu oceny kurtozy stosuje się następujące przekształcenie

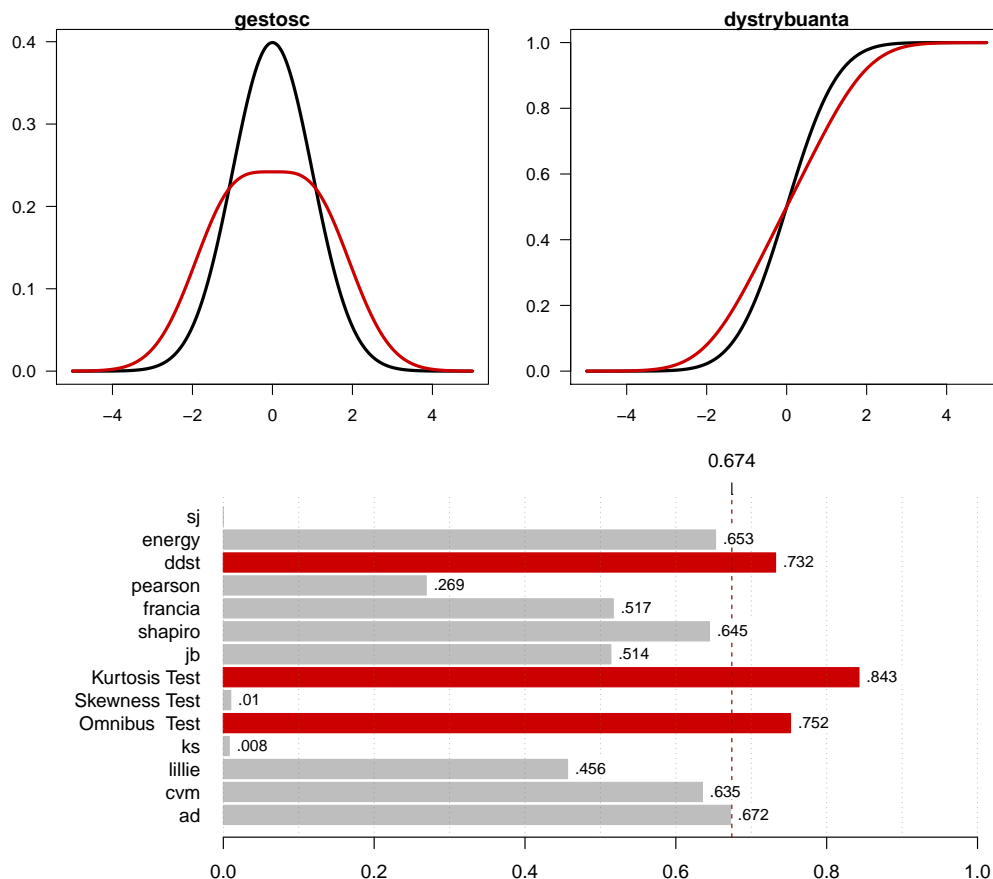
$$\begin{aligned}
E\{b_2\} &= 3(n-1)/(n+1), \\
\text{var}\{b_2\} &= \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}, \\
z &= (b_2 - E\{b_2\})/\sqrt{\text{var}\{b_2\}}, \\
\sqrt{\beta_1(b_2)} &= \frac{6(n^2-5n+2)}{(n+7)(n+9)} \sqrt{\frac{6(n+3)(n+5)}{n(n-2)(n-3)}}, \\
A &= 6 + \frac{8}{\sqrt{\beta_1(b_2)}} \left(\frac{2}{\sqrt{\beta_1(b_2)}} + \sqrt{1 + \frac{4}{\beta_1(b_2)}} \right), \\
Z(b_2) &= \left(\left(1 - \frac{2}{9A}\right) - \left(\frac{1-2/A}{1+z\sqrt{2/(A-4)}} \right)^{1/3} \right) / \sqrt{\frac{2}{9A}},
\end{aligned} \tag{11}$$

gdzie $E\{\cdot\}$, $\text{var}\{\cdot\}$ to obciążenie i wariancja oceny kurtozy dla rozkładu normalnego.

Statystyką testową jest $Z(\sqrt{b_2})$. Przy założeniu prawdziwej hipotezy zerowej ta statystyka ma asymptotyczny rozkład normalny.

W programie R ten test jest zaimplementowany w funkcji `dagoTest{fBasics}`.

```
set.seed(1313); x <- rnorm(100)
dagoTest(x)@test$p.value[3]
#Kurtosis Test
# 0.4417257
```



Rysunek 7: Moc vs. rozkład mieszaniny $N(-1, 1)$, $N(1, 1)$, $n=500$, $\alpha = 0.05$. Na górnym panelu gęstość i dystrybuantę rozkładu dla którego badano moc (na czerwono) porównano z gęstością i dystrybuantą rozkładu normalnego (na czarno). Na dolnym panelu przedstawiono moc wybranych testów. Na czerwono zaznaczono wyniki nie gorsze niż 80% mocy najlepszego z testów (patrz pionowa czerwona linia). Z wykryciem spłaszczenia rozkładu najlepiej poradził sobie test kurtozy D'Agostino. Moc oszacowana na bazie 10 000 powtórzeń.

2.3 Test typu omnibus D’Agostino-Pearsona oparty o kurtozę i skośność

Łącząc dwa powyższe testy otrzymuje się test czuły na odstępstwa od normalności zarówno w postaci niezerowej skośności jak i kurtozy istotnie różnej od 3 [Agostino1973].

Statystyką testową jest

$$K^2 = \left(Z \left(\sqrt{b_1} \right) \right)^2 + \left(Z \left(b_2 \right) \right)^2, \quad (12)$$

gdzie $Z(\sqrt{b_1})$ to statystyka testowa testu opartego o skośność a $Z(b_2)$ to statystyka testowa testu opartego o kurtozę.

Asymptotyczny rozkład tej statystyki to oczywiście rozkład χ^2_2 .

W programie R ten test jest zaimplementowany w funkcji `dagoTest{fBasics}`.

```
set.seed(1313); x <- rnorm(100)
dagoTest(x)@test$p.value[1]
#Omnibus Test
# 0.7408977
```

Rozszerzeniem testu D’Agostino-Pearsona na przypadek testowania zgodności z wielowymiarowym rozkładem normalnym, dla statystyki testowej opartej o kurtozę i skośność, jest test Doornika-Hansena [Doornik1994]. Jest on zaimplementowany w funkcji `normality.test1{normwhn.test}`, a ta implementacja pozwala też na testowanie kierunkowych alternatyw dla skośności i kurtozy, czyli że rozkład jest bardziej/mniej lewoskośny/splaszczony niż rozkład normalny. W funkcji `normality.test2{normwhn.test}` zaimplementowana jest modyfikacja pozwalająca na testowanie zgodności z rozkładem normalnym przy osłabionym założeniu dotyczącym niezależności. Tzn. jest to test normalności dopuszczający słabą strukturę zależności pomiędzy obserwacjami, patrz [Velasco2004].

Więcej o teście Doornika-Hansena i innych wielowymiarowych testach normalności (np. teście Mardii, wielowymiarowym Shapiro-Francia) przeczytać można w [BiecekTrajkowski2011].

2.4 Test typu omnibus Jarque-Bera oparty o kurtozę i skośność

Innym testem opartym o kurtozę i skośność jest test Jarque-Bera [Jarque1987]. Statystyka testowa w przypadku tego testu ma łatwiejszą postać niż dla testu D’Agostino-Pearsona. Traci się jednak na niedokładnym oszacowaniu wartości krytycznych przy niewielkich wielkościach próby. Asymptotycznie ten test jest tak samo mocny jak test D’Agostino-Pearsona, ale na asymptotykę można liczyć jedynie w przypadku dużych prób.

Statystyka testowa ma postać

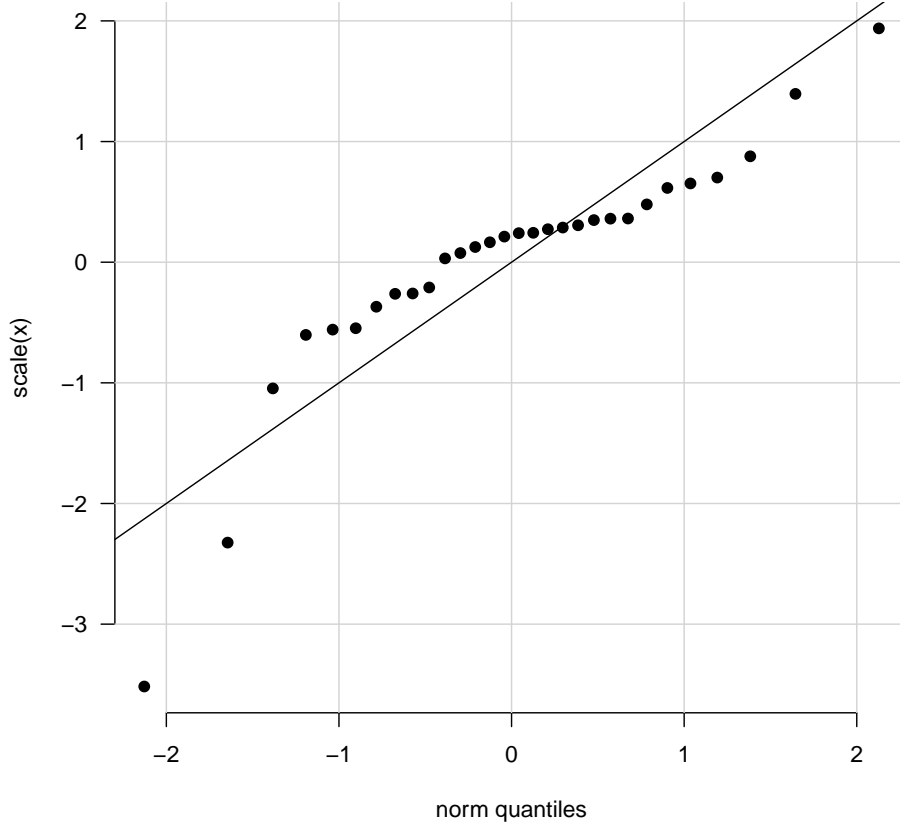
$$JB = \frac{n}{6} \left(\left(\sqrt{b_1} \right)^2 + \frac{1}{4} (b_2 - 3)^2 \right). \quad (13)$$

Asymptotyczny rozkład statystyki testowej JB to rozkład χ^2_2 . W funkcji `jbTest{fBasics}` zaimplementowane są też dwie inne metody liczenia p-wartości i szacowania kwantyli rozkładu statystyki JB oparte o mnożniki Lagrange, przy czym raportowane przez tę funkcję p-wartości są zaokrąglone do trzeciego miejsca po przecinku dziesiętnym. P-wartość wyznaczane przez funkcję `jarqueberaTest{fBasics}` nie jest zaokrąglana, jest więc podawana z większą dokładnością, ale jedynie dla przybliżenia rozkładu statystyki testowej rozkładem asymptotycznym. Zaokrąglanie p-wartości do tysięcznych uniemożliwia testowanie na poziomie istotności rzędu $\alpha = 0.001$ lub mniejszym.

```
set.seed(1313); x <- rnorm(100)
jbTest(x)@test$p.value
# LM p-value ALM p-value Asymptotic
# 0.894 0.810 0.900
```

3 Testy oparte o statystyki pozycyjne

Kolejne dwa testy oparte są o statystyki pozycyjne z próby. Ideę ich działania najłatwiej uchwycić przyglądając się wykresowi kwantylowemu QQ-plot.



Rysunek 8: Wykres kwantylowy QQ-plot, na którym prezentowane są statystyki pozycyjne i oczekiwane wartości statystyk pozycyjnych. Dla rozkładu normalnego spodziewamy się, że kwantyle empiryczne (czyli statystyki porządkowe) i kwantyle teoretyczne (czyli oczekiwane wartości statystyk porządkowych) ułożą się wzdłuż prostej $y = x$. Statystyki testowe poniżej przedstawionych testów będą badały tę współliniowość.

3.1 Test Shapiro-Wilka

Jednym z bardziej popularnych testów klasycznych jest test Shapiro-Wilka [Shapiro1965]. Statystyka testowa jest oparta o iloczyn skalarny unormowanych uporządkowanych statystyk porządkowych z unormowanymi wartościami oczekiwanymi statystyk porządkowych. Innymi słowy

$$\begin{aligned} W &= \frac{(\sum_{i=1}^n a_i x_{(i)})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}, \\ (a_1, \dots, a_n) &= \frac{m^T V^{-1}}{(m^T V^{-1} V^{-1} m)^{1/2}}, \end{aligned} \quad (14)$$

gdzie $m = (m_1, \dots, m_n)^T$ to wektor wartości oczekiwanych uporządkowanych statystyk pozycyjnych w rozkładzie normalnym, a V to macierz kowariancji tych statystyk testowych.

Wartości wektora m i macierzy V oraz rozkładu statystyki testowej pracownicy stabilizowano [Sarhan1956]. W kolejnych latach proponowano różne aproksymacje tych wartości [Davis1978, Royston1982,

Royston1992]. Wykorzystując tę ostatnią aproksymację wzór na statystykę testową można wyznaczyć następującym algorytmem.

W pierwszym kroku wyznaczane są wartości oczekiwane statystyk porządkowych i wektor wag

$$\begin{aligned}\tilde{m}_i &= \Phi^{-1}\left\{\frac{i-3/8}{n+1/4}\right\}, \\ c_i &= \tilde{m}_i \sqrt{(\tilde{\mathbf{m}}^T \tilde{\mathbf{m}})}.\end{aligned}\tag{15}$$

W zależności od rozmiaru próby stosuje się dwa różne przybliżenia.

Dla $4 \leq n < 12$ używa się przybliżenia

$$\begin{aligned}\tilde{a}_n &= -2.706056u^5 + 4.434685u^4 - 2.071190u^3 - 0.147981u^2 + 0.221157u + c_n, \\ \tilde{a}_i &= \phi^{-1/2}\tilde{m}_i, \text{ dla } i = 2, \dots, n-1, \\ u &= 1/\sqrt{n}, \\ \phi &= (\tilde{\mathbf{m}}^T \tilde{\mathbf{m}} - 2\tilde{m}_n^2)/(1 - 2\tilde{a}_n^2).\end{aligned}\tag{16}$$

Dla $n \geq 12$ używa się przybliżenia

$$\begin{aligned}\tilde{a}_n &= -2.706056u^5 + 4.434685u^4 - 2.071190u^3 - 0.147981u^2 + 0.221157u + c_n, \\ \tilde{a}_{n-1} &= -3.582633u^4 + 5.682633u^4 - 1.752461u^3 - 0.293762u^2 + 0.042981u + c_{n-1}, \\ \tilde{a}_i &= \phi^{-1/2}\tilde{m}_i, \text{ dla } i = 3, \dots, n-2, \\ u &= 1/\sqrt{n}, \\ \phi &= (\tilde{\mathbf{m}}^T \tilde{\mathbf{m}} - 2\tilde{m}_n^2 - 2\tilde{m}_{n-1}^2)/(1 - 2\tilde{a}_n^2 - 2\tilde{a}_{n-1}^2).\end{aligned}\tag{17}$$

Wyznaczywszy te współczynniki, statystykę testową wyznaczamy ze wzoru

$$W = \left(\sum_{i=1}^n \tilde{a}_i x_i \right)^2 / \sum_{i=1}^n (x_i - \bar{x})^2.\tag{18}$$

Statystyka W po przekształceniu przez funkcję $g(W)$ ma rozkład bliski rozkładowi normalnemu o średniej μ i odchyleniu standardowym σ [Royston1993].

Dla $4 \leq n < 12$ używa się przybliżenia

$$\begin{aligned}g(W) &= -\ln(\gamma - \ln(1 - W)), \\ \gamma &= 0.459n - 2.273, \\ \mu &= -0.0006714n^3 + 0.025054n^2 - 0.39978n + 0.5440, \\ \sigma &= \exp\{-0.0020322n^3 + 0.062767n^2 - 0.77857n + 1.3822\}, \\ u &= \ln(n).\end{aligned}\tag{19}$$

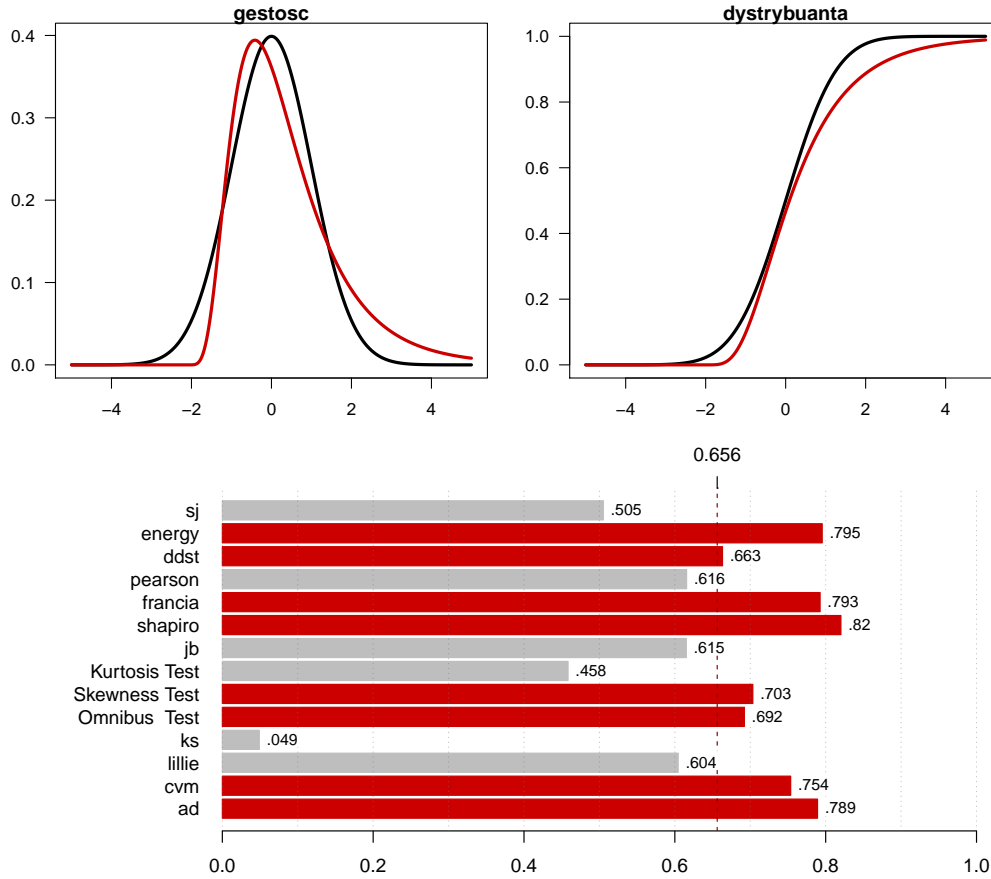
Dla $n \geq 12$ używa się przybliżenia

$$\begin{aligned}g(W) &= \ln(1 - W), \\ \mu &= 0.0038915u^3 - 0.083751u^2 - 0.31082u - 1.5851, \\ \sigma &= \exp\{0.0030302u^2 - 0.082676u - 0.4803\}, \\ u &= \ln(n).\end{aligned}\tag{20}$$

Statystyka $Z = (g(W) - \mu) / \sigma$ ma w przybliżeniu rozkład $\mathcal{N}(0, 1)$.

W programie R ten test jest zaimplementowany w funkcji `shapiro.test{stats}`. Ta implementacja pozwala na testowanie normalności wektorów o długości od 3 do 5000 obserwacji.

```
set.seed(1313); x <- rnorm(100)
shapiro.test(x)$p.value
# [1] 0.4879372
```



Rysunek 9: Moc vs. rozkład lognormalny, $n=20$, $\alpha = 0.01$. Na górnym panelu gęstość i dystrybuantę rozkładu dla którego badano moc (na czerwono) porównano z gęstością i dystrybuantą rozkładu normalnego (na czarno). Na dolnym panelu przedstawiono moc wybranych testów. Na czerwono zaznaczono wyniki nie gorsze niż 80% mocy najlepszego z testów (patrz pionowa czerwona linia). Skośność i ciężkie ogony rozkładu log normalnego najlepiej wykrywa test Shapiro-Wilka. Moc oszacowana na bazie 10 000 powtórzeń.

3.2 Test Shapiro-Francia

Test Shapiro-Francia jest uproszczoną wersją testu Shapiro-Wilka, w którym macierz kowariancji statystyk pozycyjnych V zastąpiono przez macierz identycznościową [Shapiro1972]. Dla dużych prób ten test ma podobne zachowanie jak test Shapiro Wilka.

Statystyką testową jest

$$\begin{aligned} W' &= \frac{(\sum_{i=1}^n a_i x_{(i)})^2}{\sum_{i=1}^n (X_i - \bar{X})^2}, \\ (a_1, \dots, a_n) &= \frac{\mathbf{m}^T}{(\mathbf{m}^T \mathbf{m})^{1/2}}. \end{aligned} \quad (21)$$

Warto zauważyć, że statystyka W' to korelacja punktów na wykresie kwantylowym QQ-plot. Do obliczeń wykorzystuje się wzory

$$\begin{aligned} \tilde{m}_i &= \Phi^{-1}\{(i - 3/8)/(n + 1/4)\}, \\ \tilde{a}_i &= \tilde{m}_i \sqrt{(\tilde{\mathbf{m}}^T \tilde{\mathbf{m}})}, \\ W' &= (\sum_i \tilde{a}_i x_i)^2 / \sum_i (x_i - \bar{x})^2. \end{aligned} \quad (22)$$

Podobnie jak w teście Shapiro-Wilka statystyka $g(W') = \ln(1 - W')$ jest przybliżana rozkładem normalnym o parametrach

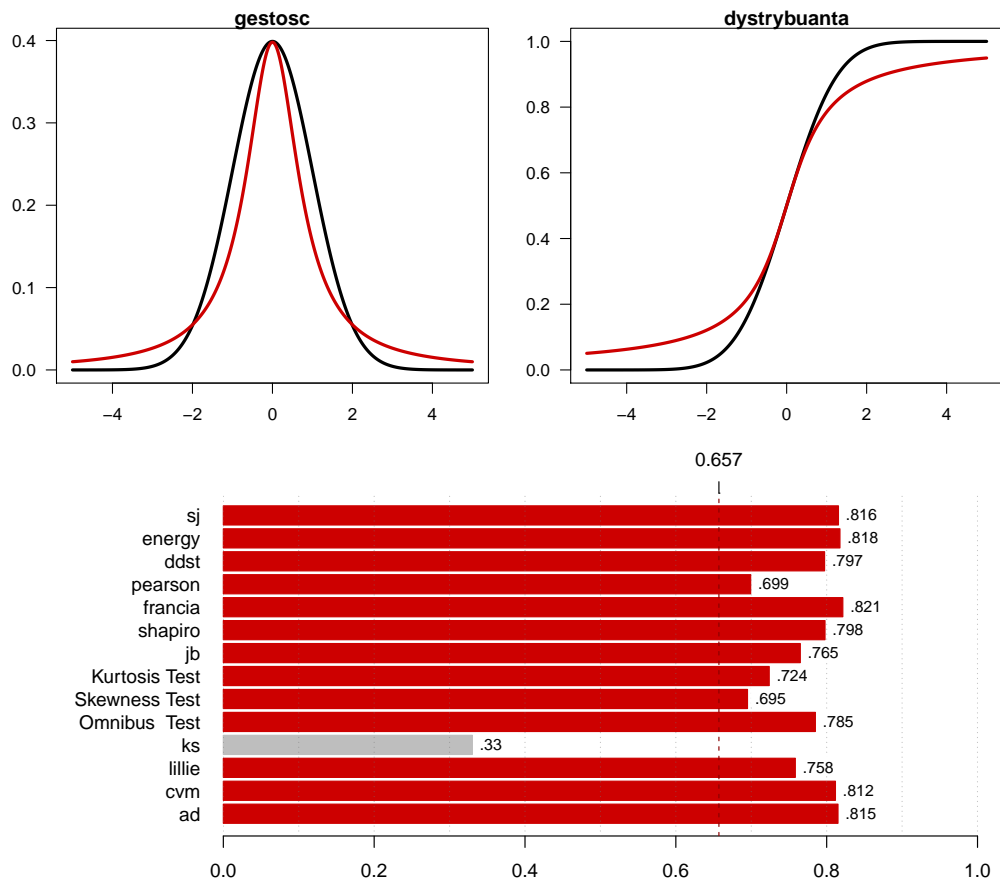
$$\begin{aligned}
\mu &= 1.0521u - 1.2725, \\
u &= \ln(\ln(n)) - \ln(n), \\
\sigma &= -0.26758v + 1.030, \\
v &= \ln(\ln(n)) + 2/\ln(n),
\end{aligned}
\tag{23}$$

dla prób z przedziału 5 do 5000 obserwacji (w R nie można testować tym testem większych prób).

Statystyka $Z = (g(W^T) - \mu) / \sigma$ ma w przybliżeniu rozkład normalny $\mathcal{N}(0, 1)$.

W programie R ten test jest zaimplementowany w funkcji `sf.test{nortest}`. Ta implementacja pozwala na testowanie normalności wektorów o długości od 5 do 5000 obserwacji.

```
set.seed(1313); x <- rnorm(100)
sf.test(x)$p.value
# [1] 0.3414609
```

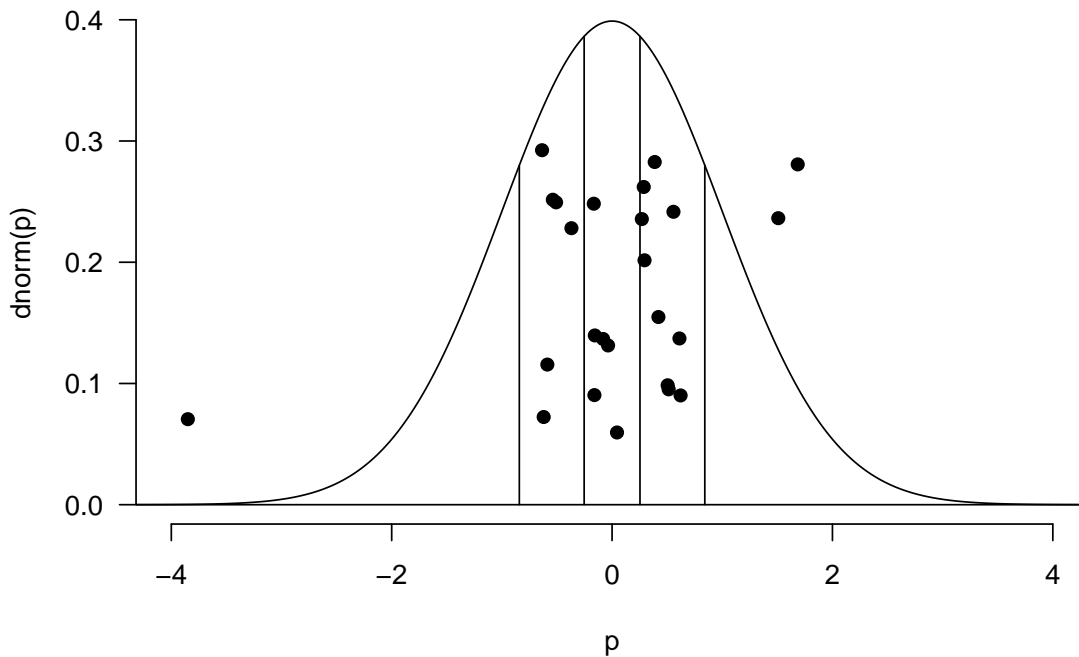


Rysunek 10: Moc vs. rozkład Cauchy’ego, $n=20$, $\alpha = 0.01$. Na górnym panelu gęstość i dystrybuantę rozkładu dla którego badano moc (na czerwono) porównano z gęstością i dystrybuantą rozkładu normalnego (na czarno). Na dolnym panelu przedstawiono moc wybranych testów. Na czerwono zaznaczono wyniki nie gorsze niż 80% mocy najlepszego z testów (patrz pionowa czerwona linia). Ciężkie ogony rozkładu Cauchego łatwo wykryć, ale najlepiej sobie z tym zadaniem poradził test Shapiro-Francia. Moc oszacowana na bazie 10 000 powtórzeń.

4 Testy oparte o statystykę χ^2

Dla rozkładów o dyskretnym nośniku do badania zgodności można wykorzystać test zgodności χ^2 . W postaci ogólnej jest on zaimplementowany w programie R w funkcji `chisq.test(stats)`.

W przypadku testowania zgodności z rozkładem ciągłym, można też wykorzystać ten test, dzieląc uprzednio nośnik rozkładu na k przedziałów, domyślnie w R liczbę przedziałów wyznacza się z wzoru $k = 2n^{2/5}$. Następnie bada się zgodność z takim zdyskretyzowanym rozkładem.



Rysunek 11: Gęstość rozkładu normalnego z zaznaczonymi brzegami pięciu przedziałów zawierających tę samą ilość masy rozkładu normalnego. Odpowiadają one przedziałom, przy których zdyskretyzowany będzie rozkład normalny. Czarne kropki przedstawiają standaryzowaną próbę, w tym przypadku wylosowaną z rozkładu o cięższych ogonach niż rozkład normalny.

4.1 Test Pearsona

Test Pearsona jest oparty o statystykę X^2

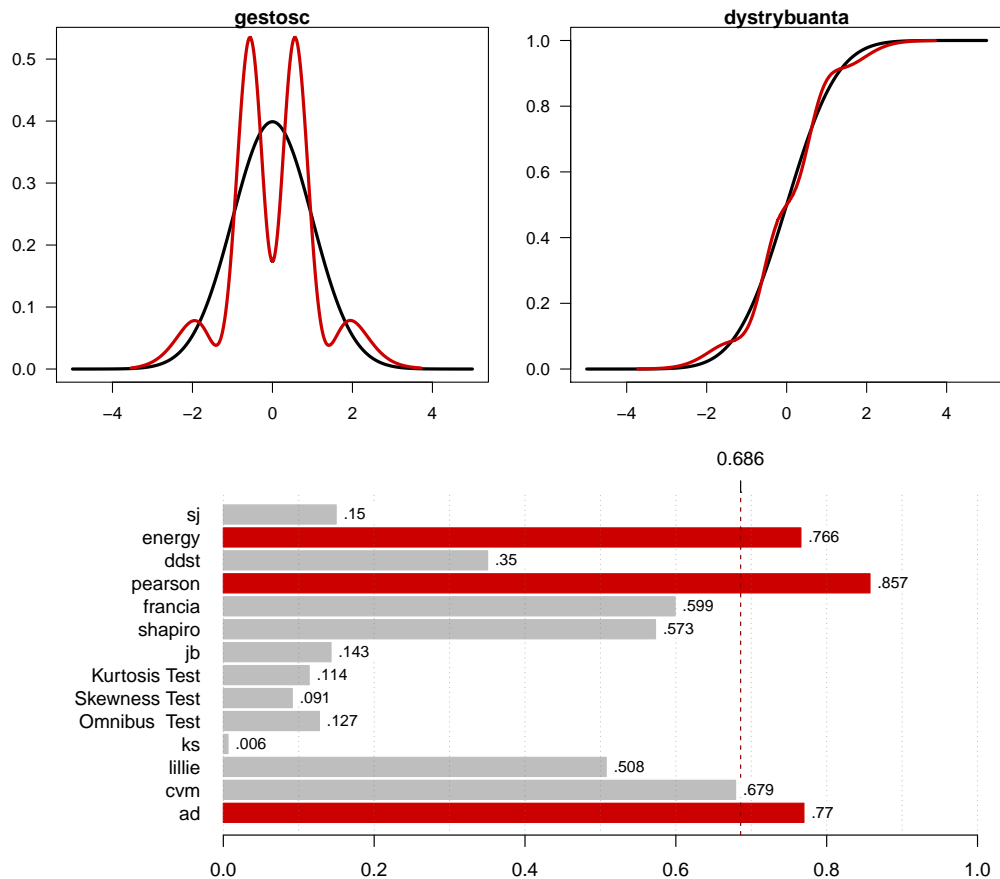
$$X^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i},$$

która dla prawdziwej hipotezy zerowej ma asymptotyczny rozkład χ^2 . Symbol E_i oznacza oczekiwaną liczebność obserwacji w klasie i , a O_i oznacza obserwowaną liczebność obserwacji w klasie i . Przybliżenie rozkładem asymptotycznym ma sens gdy liczebności O_i nie są bardzo małe, zaleca się przynajmniej 10 obserwacji w klasie, choć niektóre źródła podają za wymóg 5 lub 20 obserwacji.

Do przybliżenia rozkładu statystyki testowej używa się rozkładu χ_{k-1}^2 lub χ_{k-3}^2 . W przypadku nieznanymi parametrów μ i σ żaden z tych rozkładów nie jest dokładnym rozkładem statystyki testowej ale zaleca się wyznaczanie p-wartości z użyciem rozkładu χ_{k-3}^2 .

W programie R ten test jest zaimplementowany w funkcji `pearson.test{nortest}` oraz `pchiTest{fBasics}`. W pierwszej z tych funkcji za wybór rozkładu asymptotycznego odpowiada opcja `adjust=TRUE` (χ^2_{k-3} , domyślnie) lub `adjust=FALSE` (χ^2_{k-1}). W drugiej z tych funkcji liczone są dwie p-wartości, dla rozkładu χ^2_{k-1} i dla χ^2_{k-3} .

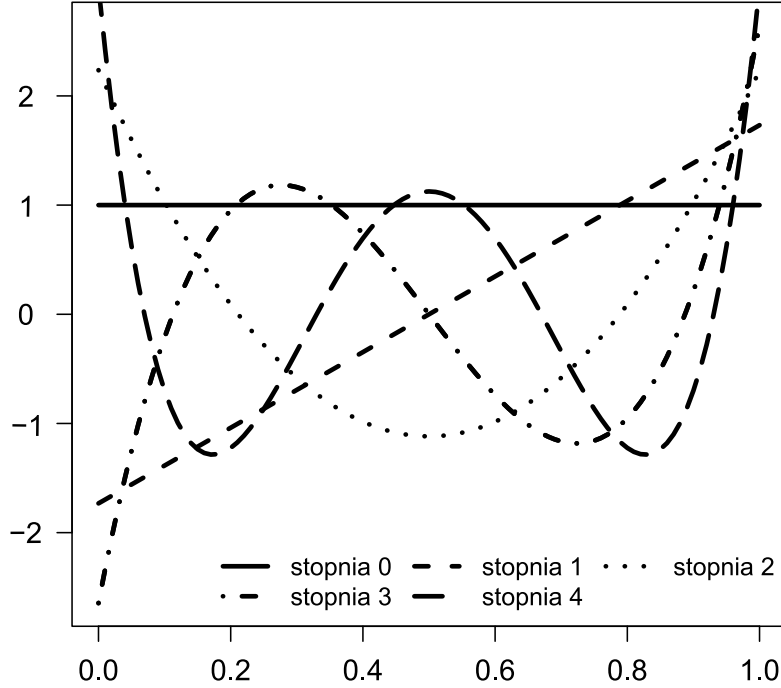
```
set.seed(1313); x <- rnorm(100)
pearson.test(x)$p.value
# [1] 0.5878833
pchiTest(x)$test$p.value
# Adjusted Not adjusted
# 0.5878833 0.7515082
```



Rysunek 12: Moc vs. alternatywa Legendre'a rzędu 7 (opis tej klasy alternatyw jest w rozdziale 5), $n=100$, $\alpha = 0.05$. Na górnym panelu gęstość i dystrybuantę rozkładu dla którego badano moc (na czerwono) porównano z gęstością i dystrybuantą rozkładu normalnego (na czarno). Na dolnym panelu przedstawiono moc wybranych testów. Na czerwono zaznaczono wyniki nie gorsze niż 80% mocy najlepszego z testów (patrz pionowa czerwona linia). Z wykryciem wielomodalności najlepiej poradził sobie test Pearsona. Moc oszacowana na bazie 10 000 powtórzeń. Test ddst był wywołany z domyślnym ograniczeniem na liczbę statystyk $k = 5$, stąd też niska moc w wykrywaniu takich złożonych odstępstw, gdyby pozwolić temu testowi na eksplorację większych k to prawdopodobnie miałby zdecydowanie większą moc.

5 Inne ciekawe testy

Poza wymienionymi klasycznymi testami zgodności, w literaturze znaleźć można też „nowoczesne” testy w znacznym stopniu oparte na możliwościach obliczeniowych współczesnych komputerów. Poniżej przedstawiona będzie modyfikacja gładkiego testu Neymana, ciekawa metodologicznie i łatwo dostępna w programie R oraz dwa inne ciekawe testy w tym „kierunkowy” test na ciężkie ogony.



Rysunek 13: Pierwsze pięć ortogonalnych wielomianów Legendre’a na odcinku $[0,1]$. Alternatywą Legendre’a rzędu k nazywam rozkład zmiennej Y , który powstał jako transformacja $Y = \Phi^{-1}(X)$ zmiennej X określonej na odcinku $[0,1]$, której gęstość zapisać można jako ważoną sumę pierwszego i k -tego wielomianu Legendre’a. Ta klasa alternatyw pozwala na badanie ciekawych scenariuszy odstępstw od rozkładu normalnego.

5.1 Gładki test Neymana zgodności z rozkładem normalnym

Poniższy opis statystyki testowej dotyczy badania zgodności z rodzinami rozkładów z parametrem położenia i skali, a w szczególności z rodziną rozkładów normalnych [Janic2008]. Więcej o gładkich testach znaleźć można w [Kallenberg1997].

Statystyką testową jest

$$W_k = \left[\frac{1}{\sqrt{n}} \sum_{i=1}^n \ell(x_i) \right] I^{-1} \left[\frac{1}{\sqrt{n}} \sum_{i=1}^n \ell(x_i) \right]^T, \quad (24)$$

gdzie $\ell(x_i)$ to k -wymiarowy wektor statystyk pomocniczych a I to macierz kowariancji tych statystyk dla prawdziwej hipotezy zerowej, $I = \text{Cov}_{\theta_0}[\ell(x_1)]^T[\ell(x_1)]$. Aby pracować z jednostkową macierzą kowariancji, za statystyki wybierane są $\ell(x_i) = (\phi_1(F(x_i)), \dots, \phi_k(F(x_i)))$, gdzie $F()$ jest dystrybuantą rozkładu z którym badamy zgodność a $\phi_i()$ to ortonormalne wycelowane wielomiany na odcinku

$[0, 1]$. Dla hipotez złożonych i badania zgodności z rodzinami rozkładów, bada się zachowanie statystyk $\ell(x_i; \hat{\gamma})$ wyznaczonych dla rozkładu o parametrach $\hat{\gamma}$.

Wybór statystyk i ich liczba wpływa na to, na jakie alternatywy ten test będzie „wyczulony”. W poniżej przedstawionej implementacji, liczba k jest wybierana z użyciem pewnego kryterium wyboru modelu, a ϕ to pierwsze k funkcji z bazy funkcji ortonormalnych, np. bazy kosinusów lub bazy wielomianów Legendre’a. Użycie kryterium wyboru modelu do wyboru liczby k powoduje, że im większa próba, tym na większą liczbę alternatyw ten test będzie czuły.

Parametry rozkładu normalnego ocenia się estymatorami

$$\begin{aligned}\tilde{\mu} &= \frac{1}{n} \sum_{i=1}^n x_i, \\ \tilde{\sigma} &= \frac{1}{n-1} \sum_{i=1}^{n-1} (x_{(i+1)} - x_{(i)})(H_{i+1} - H_i),\end{aligned}\tag{25}$$

gdzie $x_{(i)}$ to uporządkowane rosnąco statystyki porządkowe próby, a $H_i = \Phi^{-1}((i - 3/8)(n + 1/4))$.

Statystyka testowa $W_k(\tilde{\mu}, \tilde{\sigma})$ opisana jest szczegółowo w [Janic2008]. Wartość statystyki testowej można wyznaczyć dzięki stabilizowanym macierzom $[I^*(\tilde{\mu}, \tilde{\sigma})]^{-1}$. Do wyznaczania p-wartości stosuje się aproksymację funkcji kwantylowej.

W programie R ten test jest zaimplementowany w funkcji `ddst.norm.test{ddst}`.

```
set.seed(1313); x <- rnorm(100)
ddst.norm.test(x, compute.p=TRUE)
#      Data Driven Smooth Test for Normality
#
#data:  x,      base: ddst.base.legendre,    c: 100
#WT* = 0.0318, n. coord = 1, p-value = 0.872
```

5.2 Test oparty o statystykę E (energii)

Do studium symulacyjnego dodano test zgodności z wielowymiarowym rozkładem normalnym (w tym przypadku jednowymiarowym), test oparty o statystykę E (patrz [Szekely2005]), czyli średnią odległość pomiędzy każdą parą obserwacji (to nie jest prawdziwa odległość w matematycznym sensie, chodzi o $h(x, y) = E\|x - Z'\| + E\|y - Z'\| - E\|Z' - Z''\| - \|x - y\|$)

$$E = -\frac{1}{n} \sum_{j,k=1}^n (y_j - y_k)^2 + 2 \sum_{j=1}^n (2y_j \Phi(y_j) + \phi(y_j) - y_j) - c,$$

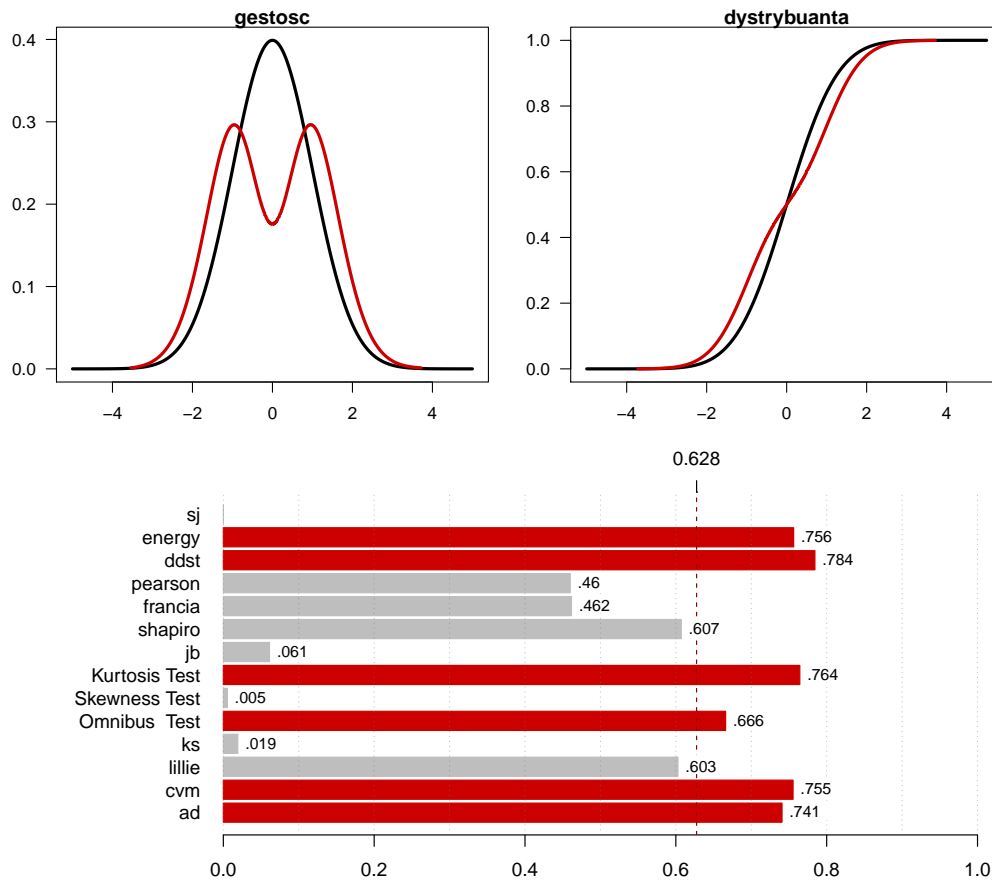
gdzie y_j to standaryzowana próba x_j , Φ i ϕ to dystrybuenta i gęstość rozkładu normalnego $\mathcal{N}(0, 1)$ a c to stała równa $c = E(Z' - Z'')^2$ gdzie $Z', Z'' \sim_{iid} \mathcal{N}(0, 1)$.

W programie R ten test jest zaimplementowany w funkcji `mvnorm.etest{energy}`.

```
set.seed(1313); x <- rnorm(100)
mvnorm.etest(x, R = 999)
#
#      Energy test of multivariate normality: estimated parameters
#
#data:  x, sample size 100, dimension 1, replicates 999
#E-statistic = 0.4521, p-value = 0.3373
```

5.3 Test oparty o statystykę SJ

W symulacjach uwzględniono również test zgodności z rozkładem normalnym oparty o statystykę SJ (S od „standard deviation” i J od „average absolute deviation from the median”) [Gel2007]. Jest to test kierunkowy badający ciężko-ogonowość rozkładu, stąd też dla pewnych alternatyw moc tego testu będzie znacznie niższa niż poziom istotności.



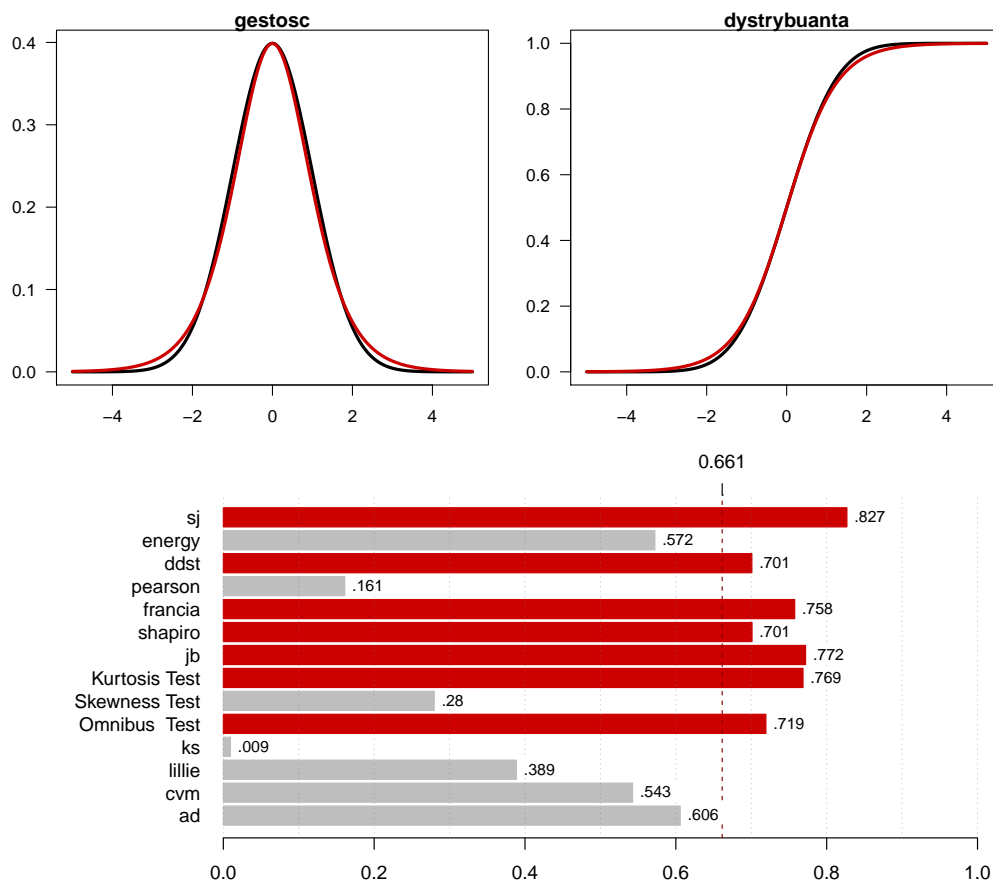
Rysunek 14: Moc vs. alternatywa Legendre'a rzędu 2, $n=100$, $\alpha = 0.05$. Na górnym panelu gęstość i dystrybuantę rozkładu dla którego badano moc (na czerwono) porównano z gęstością i dystrybuantą rozkładu normalnego (na czarno). Na dolnym panelu przedstawiono moc wybranych testów. Na czerwono zaznaczono wyniki nie gorsze niż 80% mocy najlepszego z testów (patrz pionowa czerwona linia). Z rozpoznaniem zaburzenia masy w środku rozkładu, najlepiej poradził sobie gładki test Neymana. Moc oszacowana na bazie 10 000 powtórzeń.

$$\begin{aligned}
 SJ &= \frac{S_n}{J_n}, \\
 S_n &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2, \\
 J_n &= \frac{\sqrt{\pi/2}}{n} \sum_{i=1}^n |x_i - M|,
 \end{aligned} \tag{26}$$

M to mediana próby x_i a \bar{x} to średnia z próby. Statystyka $\sqrt{n}(SJ - 1)$ ma asymptotyczny rozkład normalny.

W programie R ten test jest zaimplementowany w funkcji `sj.test{lawstat}`.

```
set.seed(1313); x <- rnorm(100)
sj.test(x)
#
#      Test of Normality - SJ Test
#
#data:  x
#Standardized SJ Statistic = 0.0353, ratio of S to J = 1.001, p-value = 0.4865
```



Rysunek 15: Moc vs. rozkład logistyczny, $n=350$, $\alpha = 0.05$. Na górnym panelu gęstość i dystrybucję rozkładu dla którego badano moc (na czerwono) porównano z gęstością i dystrybucją rozkładu normalnego (na czarno). Na dolnym panelu przedstawiono moc wybranych testów. Na czerwono zaznaczono wyniki nie gorsze niż 80% mocy najlepszego z testów (patrz pionowa czerwona linia). Aby rozpoznać rozkład logistyczny z przyzwoitą mocą potrzeba było ponad 300 obserwacji. Najlepiej z tym zadaniem poradził sobie test oparty o statystykę SJ, który częściej wykrywał cięższe ogony niż w rozkładzie normalnym. Moc oszacowana na bazie 10 000 powtórzeń.

Literatura

- [Anscombe1983] Anscombe FJ, Glynn WJ. Distribution of the Kurtosis Statistic b_2 for Normal Samples. *Biometrika*. 1983; 70(1):227-34. Blom G. Statistical Estimates and Transformed Beta-Variables. New York, NY. John Wiley & Sons, 1958.
- [Agostino1970] D'Agostino RB. Transformation to Normality of the Null Distribution of g_1 . *Biometrika*. 1970; 57(3):679-81.
- [Agostino1973] D'Agostino R, Pearson ES. Tests for Departure from Normality. Empirical Results for the Distributions of b_2 and b_1 . *Biometrika*. 1973; 60(3):613-22.
- [Agostino1990] D'Agostino RB, Belanger A, D'Agostino Jr RB. A Suggestion for Using Powerful and Informative Tests of Normality. *The American Statistician*. 1990; 44(4):316-21.
- [Anderson1954] Anderson, T.W. and Darling, D.A. (1954). A Test of Goodness-of-Fit. *Journal of the American Statistical Association* 49: 765-769.
- [BiecekTrajkowski2011] Przemysław Biecek, Krzysztof Trajkowski (2011). Na przelaj przez Data Mining, URL <http://www.biecek.pl/NaPrzelajPrzezDataMining>.
- [Davis1978] Davis C, Stephens M. Approximating the Covariance Matrix of Normal Order Statistics. *Applied Statistics*. 1978; 27(2):206-212.
- [Dickey1979] Dickey, D.A. and W.A. Fuller (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, 74, p. 427-431.
- [Doornik1994] Doornik, J.A., and H. Hansen (1994). An Omnibus Test for Univariate and Multivariate Normality, Working Paper, Nuffield College, Oxford University, U.K. Lobato, I.,
- [Gel2007] Gel, Y. R., Miao, W., and Gastwirth, J. L. (2007). Robust Directed Tests of Normality Against Heavy Tailed Alternatives. *Computational Statistics and Data Analysis* 51, 2734-2746.
- [Jarque1980] Jarque CM, Bera AK. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*. 1980; 6(3):255-259.
- [Jarque1987] Jarque CM, Bera AK. A Test for Normality of Observations and Regression Residuals. *International Statistical Review/Revue Internationale de Statistique*. 1987; 55(2):163-172.
- [Janic2008] Janic, A. and Ledwina, T. (2008). Data-driven tests for a location-scale family revisited. *J. Statist. Theory. Pract.* Special issue on Modern Goodness of Fit Methods.
- [Kallenberg1997] Kallenberg, W.C.M., Ledwina, T. (1997 a). Data driven smooth tests for composite hypotheses: Comparison of powers. *J. Statist. Comput. Simul.* 59, 101-121.
- [Lilliefors1967] Lilliefors, H. (June 1967), On the Kolmogorov-Smirnov test for normality with mean and variance unknown, *Journal of the American Statistical Association*, Vol. 62. pp. 399-402.
- [Royston1982] Royston JP. Algorithm AS 177: Expected normal order statistics (exact and approximate). *Journal of the Royal Statistical Society. Series C (Applied Statistics)*. 1982; 31(2):161-165.
- [Royston1992] Royston P. Approximating the Shapiro-Wilk W-test for non-normality. *Statistics and Computing*. 1992; 2(3):117-119.
- [Royston1993] Royston P. A Toolkit for Testing for Non-Normality in Complete and Censored Samples. *The Statistician*. 1993; 42(1):37.
- [Sarhan1956] Sarhan AE, Greenberg BG. Estimation of location and scale parameters by order statistics from singly and doubly censored samples. *The Annals of Mathematical Statistics*. 1956; 27(2):427-451.
- [Shapiro1965] Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). *Biometrika*. 1965; 52(3-4):591-611.
- [Shapiro1972] Shapiro SS, Francia R. An approximate analysis of variance test for normality. *Journal of the American Statistical Association*. 1972; 67(337):215-216.
- [Stephens1986] Stephens, M.A. (1986): Tests based on EDF statistics. In: D'Agostino, R.B. and Stephens, M.A., eds.: *Goodness-of-Fit Techniques*. Marcel Dekker, New York.
- [Szekely2005] Szekely, G. J. and Rizzo, M. L. (2005) A New Test for Multivariate Normality, *Journal of Multivariate Analysis*, 93/1, 58-80.
- [Velasco2004] C. Velasco (2004). A Simple Test of Normality of Time Series, *Econometric Theory*, 20, pp. 671-689, Cambridge University Press.

Dodatek, kody w programie R

Kod w programie R rysujący gęstości i dystrybuanty wybranych alternatyw

Rozkład Cauchego

```
x <- seq(-5,5,0.01)
plot(x, dnorm(x), type="l", lwd=3, las=1, ylab="", xlab="", main="gęstość")
lines(x, dcauchy(x,0,0.8), type="l", lwd=3, col="red3")
plot(x, pnorm(x), type="l", lwd=3, las=1, ylab="", xlab="", main="dystrybuanta")
lines(x, pcauchy(x,0,0.8), type="l", lwd=3, col="red3")
```

Rozkład Gumbela (wartości ekstremalnej)

```
plot(x, dnorm(x), type="l", lwd=3, las=1, ylab="", xlab="", main="gęstość")
lines(x, dgumbel(x,-0.3), type="l", lwd=3, col="red3")
plot(x, pnorm(x), type="l", lwd=3, las=1, ylab="", xlab="", main="dystrybuanta")
lines(x, pgumbel(x,-0.3), type="l", lwd=3, col="red3")
```

Rozkład mieszający dwóch rozkładów $N(-1,1)$, $N(1,1)$

```
plot(x, dnorm(x), type="l", lwd=3, las=1, ylab="", xlab="", main="gęstość")
lines(x, (dnorm(x,-1)+dnorm(x,1))/2, type="l", lwd=3, col="red3")
plot(x, pnorm(x), type="l", lwd=3, las=1, ylab="", xlab="", main="dystrybuanta")
lines(x, (pnorm(x,1)+pnorm(x,-1))/2, type="l", lwd=3, col="red3")
```

Rozkład logistyczny

```
plot(x, dnorm(x), type="l", lwd=3, las=1, ylab="", xlab="", main="gęstość")
lines(x, dlogis(x,0,0.627), type="l", lwd=3, col="red3")
plot(x, pnorm(x), type="l", lwd=3, las=1, ylab="", xlab="", main="dystrybuanta")
lines(x, plogis(x,0,0.627), type="l", lwd=3, col="red3")
```

Rozkład lognormalny

```
plot(x, dnorm(x), type="l", lwd=3, las=1, ylab="", xlab="", main="gęstość")
lines(x, dlnorm(x+2.2,0.83,0.5), type="l", lwd=3, col="red3")
plot(x, pnorm(x), type="l", lwd=3, las=1, ylab="", xlab="", main="dystrybuanta")
lines(x, plnorm(x+2.2,0.83,0.5), type="l", lwd=3, col="red3")
```

Rozkład alternatywa Legendre'a rzędu 2

```
tx <- seq(0.0001,1-0.0001,0.0001)
ty <- (ddst::ddst.polynomial.fun[[3]](tx)+2)/2
plot(x, dnorm(x), type="l", lwd=3, las=1, ylab="", xlab="", main="gęstość")
lines(qnorm(tx)[-1], ty[-1]/diff(qnorm(tx))/10000, type="l", lwd=3, col="red3")
plot(x, pnorm(x), type="l", lwd=3, las=1, ylab="", xlab="", main="dystrybuanta")
lines(qnorm(tx), cumsum(ty)/10000, type="l", lwd=3, col="red3")
```

Rozkład alternatywa Legendre'a rzędu 7

```
tx <- seq(0.0001,1-0.0001,0.0001)
ty <- (ddst::ddst.polynomial.fun[[7]](tx)+2)/2
plot(x, dnorm(x), type="l", lwd=3, las=1, ylab="", xlab="", main="gęstość", ylim=c(0,0.53))
lines(qnorm(tx)[-1], ty[-1]/diff(qnorm(tx))/10000, type="l", lwd=3, col="red3")
plot(x, pnorm(x), type="l", lwd=3, las=1, ylab="", xlab="", main="dystrybuanta")
lines(qnorm(tx), cumsum(ty)/10000, type="l", lwd=3, col="red3")
```

Rozkład alternatywa Legendre'a rzędu 6

```
tx <- seq(0.0001,1-0.0001,0.0001)
ty <- (ddst::ddst.polynomial.fun[[6]](tx)+4)/4
plot(x, dnorm(x), type="l", lwd=3, las=1, ylab="", xlab="", main="gęstość", ylim=c(0,0.5))
lines(qnorm(tx)[-1], ty[-1]/diff(qnorm(tx))/10000, type="l", lwd=3, col="red3")
plot(x, pnorm(x), type="l", lwd=3, las=1, ylab="", xlab="", main="dystrybuanta")
lines(qnorm(tx), cumsum(ty)/10000, type="l", lwd=3, col="red3")
```

Rozkład alternatywa Legendre'a rzędu 4

```
tx <- seq(0.0001,1-0.0001,0.0001)
ty <- (ddst:::ddst.polynomial.fun[[4]](tx)+3)/3
plot(x, dnorm(x), type="l", lwd=3, las=1, ylab="",xlab="", main="gęstość")
lines(qnorm(tx)[-1], ty[-1]/diff(qnorm(tx))/10000, type="l", lwd=3, col="red3")
plot(x, pnorm(x), type="l", lwd=3, las=1, ylab="",xlab="", main="dystrybuanta")
lines(qnorm(tx), cumsum(ty)/10000, type="l", lwd=3, col="red3")
```

Rozkład alternatywa w której część rozkładu normalnego skupiono w punkcie

```
dd <- dnorm(x)
dd[x < 0 & x > -0.32] = 0
dd[abs(x) < 0.001] = 100
plot(x, dnorm(x), type="l", lwd=3, las=1, ylab="",xlab="", main="gęstość", ylim=c(0,0.5))
lines(x, dd, type="l", lwd=3, col="red3")
pp <- pnorm(x)
pp[x < 0 & x > -0.32] = pnorm(-0.32)
plot(x, pnorm(x), type="l", lwd=3, las=1, ylab="",xlab="", main="dystrybuanta")
lines(x, pp, type="l", lwd=3, col="red3")
```

Kod w programie R rysujący wykresy mocy

```
library(fBasics)
library(lawstat)
library(nortest)
library(energy)
library(ddst)
#
# funkcja liczy p-wartosci dla rozważanych testów
# za wyjątkiem testu E, dla którego liczona jest wartosc statystyki testowej
zbiorTestow <- function(x) {
  c(ad=ad.test(x)$p.value,
    cvm = cvm.test(x)$p.value,
    lillie=lillie.test(x)$p.value,
    ks=ks.test(scale(x),"pnorm")$p.value,
    dagoTest(x)$test$p.value,
    jbTest(x)$test$p.value,
    shapiro=shapiro.test(x)$p.value,
    francia=sf.test(x)$p.value,
    pearson=pearson.test(x)$p.value,
    ddst=ddst.norm.test(x, compute.p=TRUE)$p.value,
    normal.e(x),
    sj.test(x)$p.value)
}
#
# scenariusze alternatyw
getData <- function(k = 1) {
  switch(k,
    '1'= runif(50),
    '2'= rcauchy(20),
    '3'= rlnorm(20),
    '4'= rgumbel(60),
    '5'= ifelse(runif(500)<0.5,rnorm(500),rnorm(500,2)),
    '6'= ifelse(runif(300)<0.5,rnorm(300),rnorm(300,0,2)),
    '7'= rnorm(25),
    '8' = rchisq(45,3),
    '9' = rweibull(20,0.8),
```



```

'10' = rweibull(40,1.5),
'11' = rexp(25,0.7),
'12' = ifelse(runif(1000)<0.08,rnorm(1000),rnorm(1000,0,10)),
'13' = round(rnorm(500)*5,0),
'14' = runif(500)+runif(500),
'15' = runif(1000)+runif(1000)+runif(1000),
'16' = rlogis(350),
'17' = rt(100,3),
'18' = qnorm(sample(tx,80,replace=T,prob=wx)),
'19' = qnorm(sample(tx,100,replace=T,prob=wx2)),
'20' = qnorm(sample(tx,100,replace=T,prob=wx3)),
'21' = qnorm(sample(tx,250,replace=T,prob=wx4)),
'22' = qnorm(sample(tx,1000,replace=T,prob=wx5)),
'23' = qnorm(sample(tx,100,replace=T,prob=wx6)))
}
#
# gestosci pomocnicze
tx <- seq(0.0001,1-0.0001,0.0001)
wx <- ddst::ddst.polynomial.fun[[5]](tx)+2
wx2 <- ddst::ddst.polynomial.fun[[3]](tx)+2
wx3 <- ddst::ddst.polynomial.fun[[7]](tx)+2
wx4 <- ddst::ddst.polynomial.fun[[6]](tx)+4
wx5 <- ddst::ddst.polynomial.fun[[2]](tx)+2
wx6 <- ddst::ddst.polynomial.fun[[4]](tx)+3
#
# poziomy istotnosci dla kolejnych testow
alfy <- c(0.05,0.01,0.01,0.05,0.05,0.05,0.01,0.01,0.01,0.05,
          0.03,0.05,0.01,0.05,0.05,0.05,0.05,0.05,0.05,0.05,
          0.05,0.05,0.05,0.05,0.05,0.05,0.05,0.05,0.05,0.05)
#
# wyznacz moce dla wszystkich alternatyw
N <- 10000
for (alte in 1:23) {
  # policz p-wartosci
  res <- replicate(N, zbiorTestow(getData(alte)) )
  # dla testu E wyznacz empirycznie p-wartosci
  n <- length(getData(alte))
  res0 <- replicate(N, normal.e(rnorm(n)) )
  res["energy",] <- sapply(res["energy",], function(r) (sum(r < res0) + 1)/(N+1))
  # rysuj wykres do pliku pdf
  pdf(paste0("moc",alte,".pdf"),9,4)
  par(mar=c(3,10,2,2))
  moce <- rowMeans(res < alfy[alte])
  th <- max(moce,na.rm=T)*0.8
  bb <- barplot(moce[-(8:9)], horiz=T,las=1, xlim=c(0,1), main="",
               col=ifelse(moce[-(8:9)] > th, "red3", "grey"),
               border=ifelse(moce[-(8:9)] > th, "red3", "grey"))
  abline(v = seq(0,1,0.1), lty = 3, col = "#77777777")
  abline(v = th, lty = 2, lwd = 1, col = "red4")
  rect(moce[-(8:9)]+0.005, bb - 0.35, moce[-(8:9)]+0.08, bb + 0.4,
       col="white", border="white")
  text(moce[-(8:9)]+0.01, bb - 0.25, substr(as.character(round(moce[-(8:9)]),3)),2,10),
       adj=c(0,0), cex=0.8)
  axis(3,th,round(th,3))
  dev.off()
}

```