

STATYSTYKA W FINANSACH I UBEZPIECZENIACH
LABORATORIUM

SPRAWOZDANIE

OPRACOWAŁA:
ALEKSANDRA GRZESZCZUK
NUMER ALBUMU: 255707

SPIS TREŚCI

1	CEL	2
2	ZADANIE 1	2
3	ZADANIE 2	10
4	ZADANIE 3	13
5	ZADANIE DODATKOWE 1	17

1 CEL

W sprawozdaniu zbadane zostaną rozkłady stabilne. Porównane zostaną różne metody estymacji ich parametrów i kwantyli. Następnie otrzymane rezultaty posłużą do estymacji punktowej i przedziałowej miar ryzyka VaR oraz ES dla cen otwarcia akcji trzech spółek giełdowych - BMW, APPLE oraz TESLA. Sprawozdanie będzie opierać się głównie na bibliotece `libstableR`, `quantmod` oraz `stats`.

2 ZADANIE 1

Pakiet `libstableR` służy między innymi do generowania liczb losowych oraz estymacji parametrów stabilnych rozkładów skośnych. Zapewnia on funkcje do pracy ze stabilnymi dystrybucjami skośnymi w szybki i dokładny sposób.

Zadanie pierwsze opiera się na rozkładach stabilnych, więc rozpoczniemy od definicji. Zmienna losowa X ma rozkład stabilny $S(\alpha, \beta, \sigma, \mu)$, jeżeli jej funkcja charakterystyczna jest postaci:

$$\varphi(t) = \exp\{-\sigma^\alpha |t|^\alpha \left(1 - i\beta \left(\tan \frac{\pi\alpha}{2}\right) (\text{sign}t)\right) + i\mu t\} \quad \alpha \neq 1$$

$$\varphi(t) = \exp\{-\sigma |t| \left(1 - i\beta \frac{2}{\pi} (\text{sign}t) \ln|t|\right) + i\mu t\} \quad \alpha = 1$$

Podkreślmy, że $0 < \alpha \neq 2$, $-1 \neq \beta \neq 1$, $\sigma > 0$ oraz $\mu \in \mathbf{R}$

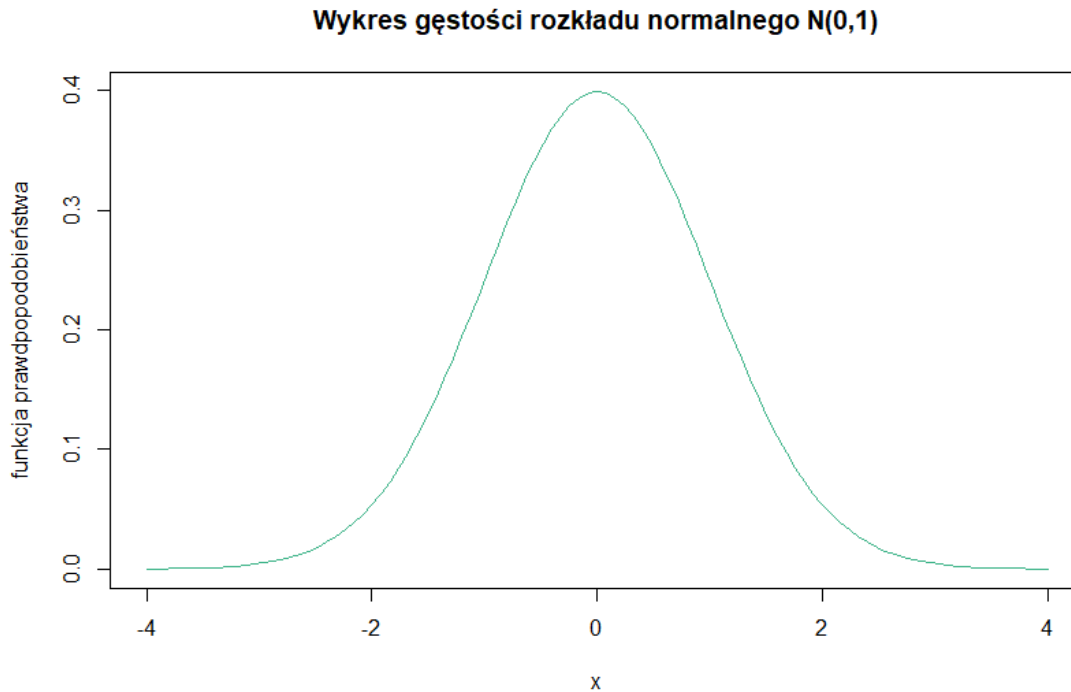
Parametr α nazywamy współczynnikiem stabilności. Określa on szybkość z jaką zanikają ogony. Gdy $\alpha = 2$ uzyskuje się rozkład normalny, w przypadku $\alpha < 2$ wariancja rozkładu jest nieskończoną oraz rozkłady posiadają grubsze ogony od rozkładu normalnego. Dla $\alpha > 1$ średnia rozkładu jest skończona i wynosi μ .

Parametr β określa skośność rozkłady. Jeżeli $\beta = 0$ to mamy do czynienia z rozkładem symetrycznym. Jeżeli $\beta > 0$ to rozkład jest prawostronnie skośny, zaś gdy $\beta < 0$, rozkład jest lewostronnie skośny.

Ogólnie, gęstości rozkładów stabilnych nie mają jawnej postaci, za wyjątkiem rozkładu

- normalnego ($\alpha = 2$)
- Cauch'ego ($\alpha = 1, \beta = 0$)
- Levy'ego ($\alpha = \frac{1}{2}, \beta = 1$)

Rozkłady stabilne charakteryzują się jednomodalnością, to znaczy, że mają tylko jedną wartość maksymalną. Poniżej przedstawiamy wykres gęstości rozkładu normalnego $\mathcal{N}(0, 1)$, który charakteryzuje się jednomodalnością.



Rysunek 1: Przykładowy wykres jednomodalny - opracowanie własne

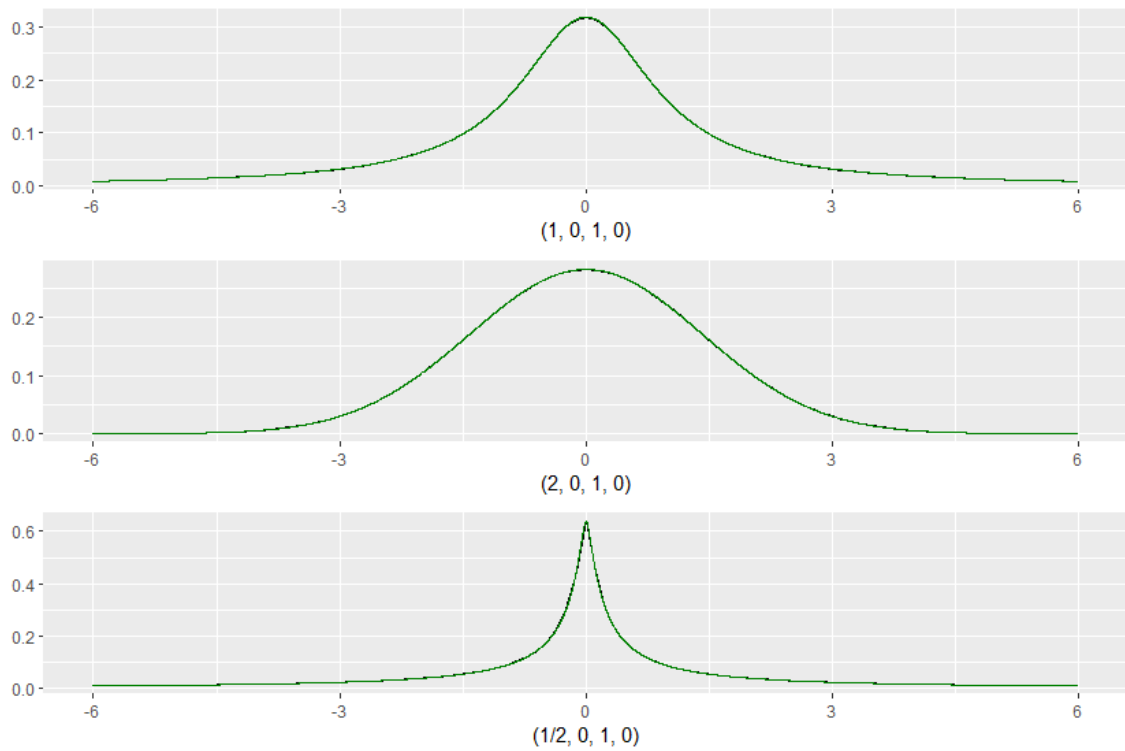
W omówionej bibliotece `libstableR` zaimplementowane są cztery metody estymacji parametrów rozkładów stabilnych:

- metoda największej wiarygodności (*MI*)
- zmodyfikowana metoda największej wiarygodności (*MII*)
- metoda kwantyli rozpaprywana przez McCullocha (*MIII*)
- metoda Koutrouvellisa oparta na funkcji charakterystycznej (*MIV*)

Jak już zostało wyżej wspomniane, badane próby pochodzą z rozkładów o parametrach

- $(1, 0, 1, 0)$ - rozkład Cauchy'ego z parametrem położenia $\alpha = 1$ oraz parametrem skali $\beta = 0$
- $(2, 0, 1, 0)$ - rozkład normalny z parametrem położenia $\alpha = 2$ oraz parametrem skali $\beta = 0$
- $(\frac{1}{2}, 0, 1, 0)$ - rozkład Levy'ego $\alpha = \frac{1}{2}$ oraz $\beta = 1$

Poniżej przedstawiamy wykresy gęstości kolejno rozkładu Cauchy'ego, normalnego oraz Levy'ego.



Rysunek 2: Wykresy odpowiednich zadanych rozkładów gęstości - opracowanie własne

Widzimy, że zgodnie z założeniem, zadane rozkłady stabilne cechują się jednomodalnością, to znaczy, mają jeden widoczny garb.

Dla każdego z trzech wektorów parametrów generujemy 100 prób zadanych dwóch rozmiarów - $n_1 = 100$, $n_2 = 1000$. W tabelach poniżej przedstawiamy podstawowe parametry z zadanych rozkładów - wartość minimalną, maksymalną, pierwszy oraz trzeci kwantyl.

Min.	1st Qu.	3rd Qu.	Max.
9.64	25.83	143.79	2447.38

Tabela 1: Podstawowe parametry rozkładu stabilnego $(1, 0, 1, 0)$ dla próby $n = 100$

Min.	1st Qu.	3rd Qu.	Max.
2.29	3.24	4.10	5.16

Tabela 2: Podstawowe parametry rozkładu stabilnego $(2, 0, 1, 0)$ dla próby $n = 100$

Min.	1st Qu.	3rd Qu.	Max.
71.89	1106.01	43834.36	9671967.14

Tabela 3: Podstawowe parametry rozkładu stabilnego $(\frac{1}{2}, 0, 1, 0)$, dla próby $n = 100$

Min.	1st Qu.	3rd Qu.	Max.
81.35	244.03	972.59	34929.24

Tabela 4: Podstawowe parametry rozkładu stabilnego $(1, 0, 1, 0)$, dla próby $n = 1000$

Min.	1st Qu.	3rd Qu.	Max.
3.64	4.28	4.74	6.06

Tabela 5: Podstawowe parametry rozkładu stabilnego $(2, 0, 1, 0)$, dla próby $n = 1000$

Min.	1st Qu.	3rd Qu.	Max.
13798.95	86663.06	1981607.64	2020167278.84

Tabela 6: Podstawowe parametry rozkładu stabilnego $(\frac{1}{2}, 0, 1, 0)$, dla próby $n = 1000$

Przeanalizujemy najpierw tabele (1), (2), (3). Widzimy, że zarówno rozkład $(1, 0, 1, 0)$ jak i $(2, 0, 1, 0)$ mają w miarę typowe, normalne wartości podstawowych parametrów, zaś rozkład $(\frac{1}{2}, 0, 1, 0)$ już dla tak małej próby posiada te wartości bardzo duże.

Przyglądając się tabelom (4), (5), (6) dochodzimy do podobnych wniosków jak wcześniej, jednakże tym razem wartość maksymalna rozkładu stabilnego $(\frac{1}{2}, 0, 1, 0)$ jest rzędu 10^9 . Spodziewamy się, że może wystąpić problem z tym rozkładem w dalszej analizie.

Ważną rzeczą wartą omówienia jest także rozstęp międzykwartyłowy. Dokładnie widzimy, że najmniejszy IQR daje nam rozkład stabilny $(2, 0, 1, 0)$. Spodziewamy się zatem, że wystąpi tutaj najmniejszy rozrzut parametrów.

	α_1	β_1	σ_1	μ_1	α_2	β_2	σ_2	μ_2
(n_1, p_1)	1.035	0.009	1.000	-0.056	0.986	0.107	1.018	-0.122
(n_1, p_2)	2.000	0.000	0.943	0.011	1.962	0.022	0.989	-0.082
(n_1, p_3)	0.512	0.019	0.956	-0.032	0.505	0.102	1.038	-0.117

Tabela 7: Oszacowania parametrów rozkładu stabilnego dla metody *I* oraz *II*

	α_3	β_3	σ_3	μ_3	α_4	β_4	σ_4	μ_4
(n_1, p_1)	0.991	0.098	1.015	-0.053	1.002	0.000	0.977	-0.051
(n_1, p_2)	1.950	0.000	0.990	-0.070	2.000	0.000	0.962	0.007
(n_1, p_3)	0.500	0.117	0.861	-0.070	0.449	-0.076	0.923	-0.041

Tabela 8: Oszacowania parametrów rozkładu stabilnego dla metody *III* oraz *IV*

W powyższych tabelach (7) oraz (8) przedstawiliśmy oszacowania parametrów rozkładu stabilnego dla metody estymatora największej wiarygodności, zmodyfikowanej metody największej wiarygodności, McCullocha oraz Koutrouvelisa. Oznaczenie n_1 oznacza, że badamy próbkę wielkości 100. Oznaczenia p_1, p_2, p_3 oznaczają kolejno próbki z rozkładu Cauchy'ego $(1, 0, 1, 0)$, normalnego $(2, 0, 1, 0)$ oraz Levy'ego $(\frac{1}{2}, 0, 1, 0)$.

	α_1	β_1	σ_1	μ_1	α_2	β_2	σ_2	μ_2
n_2p_1	1.003	0.002	1.026	0.003	1.003	-0.007	1.022	-0.022
n_2p_2	2.000	0.000	0.998	-0.003	1.962	0.022	1.002	-0.008

Tabela 9: Oszacowania parametrów rozkładu stabilnego dla metody I oraz II

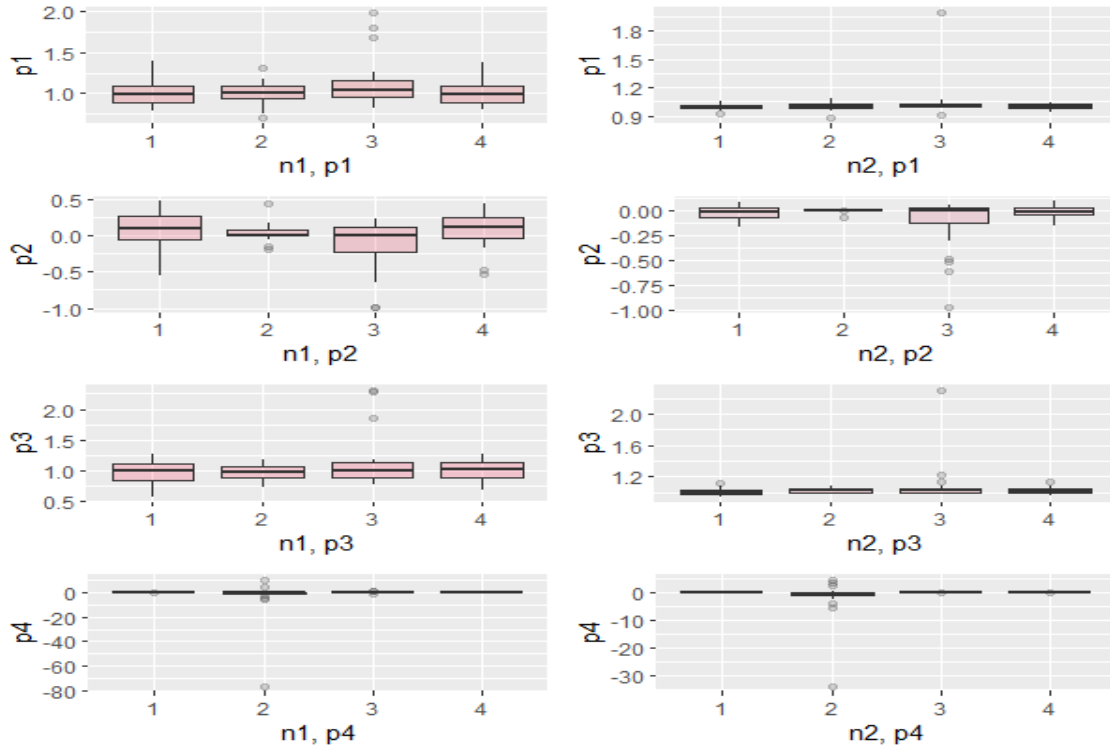
	α_3	β_3	σ_3	μ_3	α_4	β_4	σ_4	μ_4
n_2p_1	0.999	-0.006	1.004	0.007	1.000	0.000	1.025	-0.379
n_2p_2	1.950	0.000	1.002	-0.008	2.000	0.000	1.002	-0.001

Tabela 10: Oszacowania parametrów rozkładu stabilnego dla metody III oraz IV

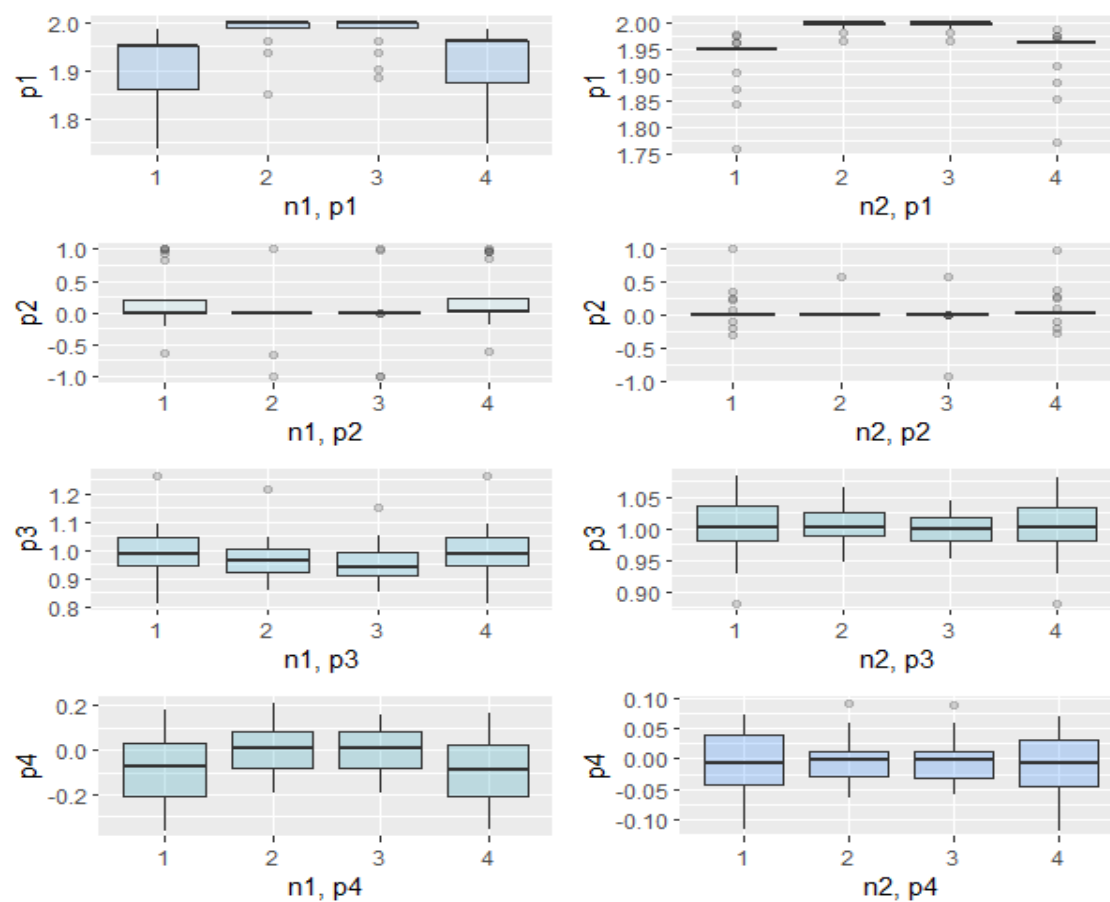
W tabelach (9) oraz (10) przedstawiliśmy oszacowania parametrów ponownie dla rozkładu stabilnego dla metody estymatora największej wiarygodności, zmodyfikowanej metody największej wiarygodności, McCullocha oraz Koutrouvelisa. Oznaczenie n_2 oznacza, że badamy próbkę wielkości 1000. Dla rozkładu p_3 , czyli rozkładu Levy'ego $(\frac{1}{2}, 0, 1, 0)$ nie dokonaliśmy oszacowania parametrów dla tak dużej próby, ponieważ wyskakuje wtedy błąd.

Wykres pudełkowy (**boxplot**) jest to forma graficznej prezentacji rozkładu cechy statystycznej, spotykana często w pakietach komputerowych wspomagających proces analizy i interpretacji danych statystycznych. Pozwala ująć na jednym rysunku wiadomości dotyczące położenia, rozproszenia i kształtu rozkładu empirycznego badanej cechy.

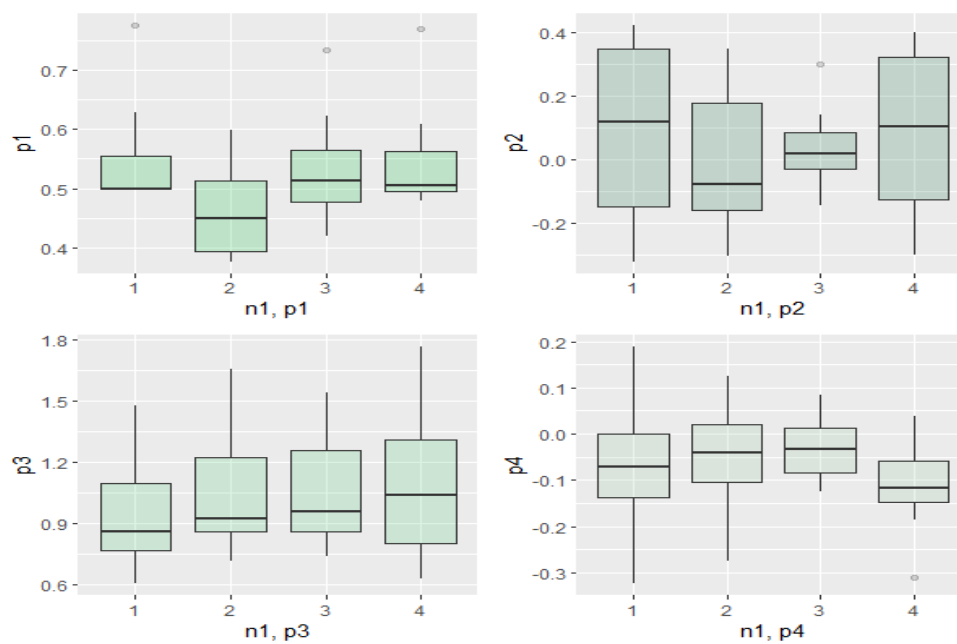
Poniższe cztery wykresy pudełkowe ((3), (4), (5)) prezentują jak rozkładały się dopasowania każdego z parametrów w próbach n_1 oraz n_2 .



Rysunek 3: Boxplot dla parametrów $(1, 0, 1, 0)$, $n_1 = 100$, $n_2 = 1000$ - opracowanie własne



Rysunek 4: Boxplot dla parametrów $(2, 0, 1, 0)$, $n_1 = 100$, $n_2 = 1000$ - opracowanie własne



Rysunek 5: Boxplot dla parametrów $(\frac{1}{2}, 0, 1, 0)$, $n_1 = 100$ - opracowanie własne

Analizując rysunek (3) widzimy, że dla większej próby $n_2 = 1000$ otrzymaliśmy zdecydowanie mniejsze "pudełka", co oznacza, że mamy mniejszy rozstęp międzykwartyłowy. To znaczy, że nasze dane są mniej rozproszone, przez co przyjmują zbliżone do siebie wartości. Podobne wnioski otrzymujemy przyglądając się wykresowi (4). W przypadku rozkładu $(\frac{1}{2}, 0, 1, 0)$ analizujemy jednak tylko mniejszą próbę, dla $n_1 = 100$. Widzimy tutaj duży rozrzut możliwych wartości. Jednakże w tym przypadku stwierdzamy, że próba mniejszego rozmiaru jest lepsza.

	α_{I-II}	α_{III-IV}	β_{I-II}	β_{III-IV}	σ_{I-II}	σ_{III-IV}	μ_{I-II}	μ_{III-IV}
n_1p_1	0.049	0.098	0.018	0.066	0.011	0.098	0.038	0.002
n_1p_2	0.038	0.022	0.045	0.096	0.050	0.000	0.029	0.077
n_1p_3	0.007	0.083	0.082	0.085	0.051	0.193	0.063	0.029

Tabela 11: Różnica oszacowań parametrów metodami I, II, III, IV

Tabela (11) zawiera wartość bezwzględną z różnicy oszacowań parametrów wszystkich rozkładów dla $n_1 = 100$. Oznaczenie α_{I-II} określa, że od α oszacowanej metodą I odjęliśmy α oszacowaną metodą II . Analogicznie ze wszystkimi innymi przypadkami. Pozwoli nam to sprawdzić, które metody są ze sobą zgodne i podają jak najbardziej zbliżone do siebie wartości.

W przypadku parametru α najmniejszą różnicę uzyskaliśmy dla parametrów $(1, 0, 1, 0)$, $(\frac{1}{2}, 0, 1, 0)$ poprzez odjęcie metody I od II , zaś dla parametrów $(2, 0, 1, 0)$ - poprzez odjęcie metody III od IV .

Dla parametru β w każdym przypadku lepszą metodą okazało się odjęcie metody I od II .

Parametr σ , dał nam identycznie wnioski jak parametr α . Najmniejszą różnicę uzyskaliśmy dla parametrów $(1, 0, 1, 0)$, $(\frac{1}{2}, 0, 1, 0)$ poprzez odjęcie metody I od II , zaś dla parametrów $(2, 0, 1, 0)$ - poprzez odjęcie metody III od IV .

Parametr μ dla rozkładów stabilnych $(1, 0, 1, 0)$, $(\frac{1}{2}, 0, 1, 0)$ najmniejszą różnicę dawał przy odjęciu metody III od IV , zaś dla rozkładu $(2, 0, 1, 0)$ - przy odjęciu metody I od II .

Na początku zostało wspomniane, że wszystkie cztery metody estymacji są zaimplementowane w programie R w bibliotece `libstableR`. Poniżej podajemy funkcje z jakich będziemy korzystać przy każdej z zadanych metod.

- Estymator McCullocha \rightarrow `stable-fit-init`
- Estymator Koutrouvelisa \rightarrow `stable-fit-koutrouvelis`
- Estymator największej wiarygodności \rightarrow `stable-fit-mle`
- Zmodyfikowany estymator największej wiarygodności \rightarrow `stable-fit-mle2d`

	α	β	σ	μ
(n_1p_1)	0.9637746	-0.18048878	1.13235583	-0.05805623
(n_1p_2)	1.95	0	1.04597614	-0.03515815
(n_1p_3)	0.5	0.4007486	0.638067	-0.1353288

Tabela 12: Oszacowania parametrów wbudowaną funkcją metody McCullocha

	α	β	σ	μ
(n_1p_1)	0.69697366	-0.18297912	0.91362797	0.02700567
(n_1p_2)	2	0	0.988873136	-0.007744301
(n_1p_3)	0.3866455	-0.0366756	0.8065543	-0.07385842

Tabela 13: Oszacowania parametrów wbudowaną funkcją metody Koutrouvelisa

	α	β	σ	μ
(n_1p_1)	1.0956980	-0.7665733	1.3906922	0.1650624
(n_1p_2)	1.96482573	-0.96547601	0.97240335	0.04516126
(n_1p_3)	0.4157755	0.0845511	0.8367212	-0.024774

Tabela 14: Oszacowania parametrów wbudowaną funkcją metody największej wiarygodności

	α	β	σ	μ
(n_1p_1)	0.9439167	-0.1730279	1.206432	-0.2204595
(n_1p_2)	1.96201399	0.02192405	1.04466828	-0.03573583
(n_1p_3)	0.4946053	0.3763376	0.6552838	-0.1245267

Tabela 15: Oszacowania parametrów wbudowaną funkcją zmodyfikowanej metody największej wiarygodności

Zbadamy teraz wartość bezwzględną różnicy z wyznaczonych wcześniej oszacowań parametrów zadanymi metodami oraz wyznaczonych teraz zadanymi funkcjami.

	α	β	σ	μ
(n_1p_1)	0.027225399	0.278588775	0.117655826	0.004556227
(n_1p_2)	0.00000000	0.00000000	0.05587614	0.03534185
(n_1p_3)	0.00000000	0.28344862	0.22263296	0.06512881

Tabela 16: Wartość bezwzględna oszacowań parametrów metodą McCullocha

	α	β	σ	μ
(n_1p_1)	0.30552634	0.18297912	0.06337203	0.07820567
(n_1p_2)	0.00000000	0.00000000	0.02737314	0.01464430
(n_1p_3)	0.06275450	0.03932440	0.11684570	0.03305842

Tabela 17: Wartość bezwzględna oszacowań parametrów metodą Koutrouvelisa

	α	β	σ	μ
(n_1p_1)	0.06079804	0.77587333	0.39069217	0.22126237
(n_1p_2)	0.03517427	0.96547601	0.02910335	0.03386126
(n_1p_3)	0.096624510	0.065051101	0.118878807	0.007725999

Tabela 18: Wartość bezwzględna oszacowań parametrów metodą największej wiarygodności

	α	β	σ	μ
(n_1p_1)	0.04168328	0.27992791	0.18873197	0.09865946
(n_1p_2)	1.398833e-05	2.405265e-05	5.586828e-02	4.926417e-02
(n_1p_3)	0.010794694	0.274237617	0.382316226	0.007026691

Tabela 19: Wartość bezwzględna oszacowań parametrów zmodyfikowaną metodą największej wiarygodności

Podsumowując widzimy, że w każdym z zadanych przypadków wartość bezwzględna z różnicy oszacowań parametrów daną metodą jest bardzo mała.

Rozważając wszystkie analizy wykonane w zadaniu uważam, że najlepszą z rozważanych metod okazała się metoda McCullocha, ponieważ przy wartości bezwzględne oszacowywanych parametrów funkcją wbudowaną a funkcją napisaną były najmniejsze.

3 ZADANIE 2

Zadanie drugie polega na przeprowadzeniu symulacji w celu porównania metod parametrycznych i nieparametrycznych szacowań kwantyli z rozkładu stabilnego. Ponownie będziemy rozważać te same rozkłady stabilne co w zadaniu pierwszym, czyli $(\alpha, \beta, \sigma, \mu) = (1, 0, 1, 0) \vee (2, 0, 1, 0) \vee (\frac{1}{2}, 0, 1, 0)$. W przypadku parametrycznym zastosujemy metodę podstawień oszacowań parametrów uzyskanych czterema metodami w zadaniu pierwszym, oraz wykorzystamy funkcję `stable-q` z biblioteki `libstableR`. Jednakże w przypadku nieparametrycznym skorzystamy z funkcji `quantile` z pakietu `stats`.

Ustalmy oznaczenia. Niech

- Metoda 1P - oszacowania uzyskane metodą McCullocha z poprzedniej analizy
- Metoda 2P - oszacowania uzyskane metodą Koutrouvelisa z poprzedniej analizy
- Metoda 3P - oszacowania uzyskane metodą największej wiarygodności I z poprzedniej analizy
- Metoda 4P - oszacowania uzyskane zmodyfikowaną metodą największej wiarygodności z poprzedniej analizy

W przypadku nieparametrycznym, odpowiednie kwantyle próbkowe $Q(p) = (1-y)x[j] + yx[j+1]$ uzyskamy w postaci nieciągłej jako odwrotną funkcję dystrybuanty empirycznej lub poprzez liniową interpolację między punktami $(p[k], x[k])$. Ustalamy, że $x[k]$ jest statystyką k -tego rzędu oraz

- Metoda 1NP - odwrotność funkcji dystrybuanty empirycznej (nieciągły kwantyl próbkowy, $y = 0$)
- Metoda 2NP - wykorzystująca liniową interpolację (ciągły kwantyl próbkowy), gdzie $p[k] = E[F(x[k])]$

	1P	2P	3P	4P	1NP	2NP	3NP
$(n_1, p_1, 0.95)$	0.033	1.051	1.056	0.111	0.679	0.351	0.226
$(n_1, p_2, 0.95)$	0.005	0.089	0.084	0.004	0.092	0.037	0.055
$(n_1, p_3, 0.95)$	9.109	23.271	5.261	2.172	3.483	39.626	32.343

Tabela 20: Bezwzględny błąd oszacowania kwantyli rzędu 0.95

- Metoda 3NP - ciągły kwantyl próbkowy $p[k] = \text{mediana}[F(x[k])]$

W powyższej tabelce dostrzegamy, że w przypadku p_1 oraz p_2 nie ma znaczących różnic w oszacowywaniu parametrów metodami parametrycznymi, jak i nieparametrycznymi. W przypadku p_3 natomiast występują bardzo duże różnice w oszacowywaniu. Tak jak zostało już wcześniej wspomniane, spodziewaliśmy się, że w przypadku p_3 nasze wyniki nie będą zbyt dobre i dokładne.

	1P	2P	3P	4P	1NP	2NP	3NP
$(n_1, p_1, 0.99)$	9.652	1.443	5.118	9.971	11.218	10.972	2.275
$(n_1, p_2, 0.99)$	0.508	0.035	0.074	0.450	0.242	0.075	0.045
$(n_1, p_3, 0.99)$	200.124	1352.042	269.308	205.359	1213.726	208.204	248.853

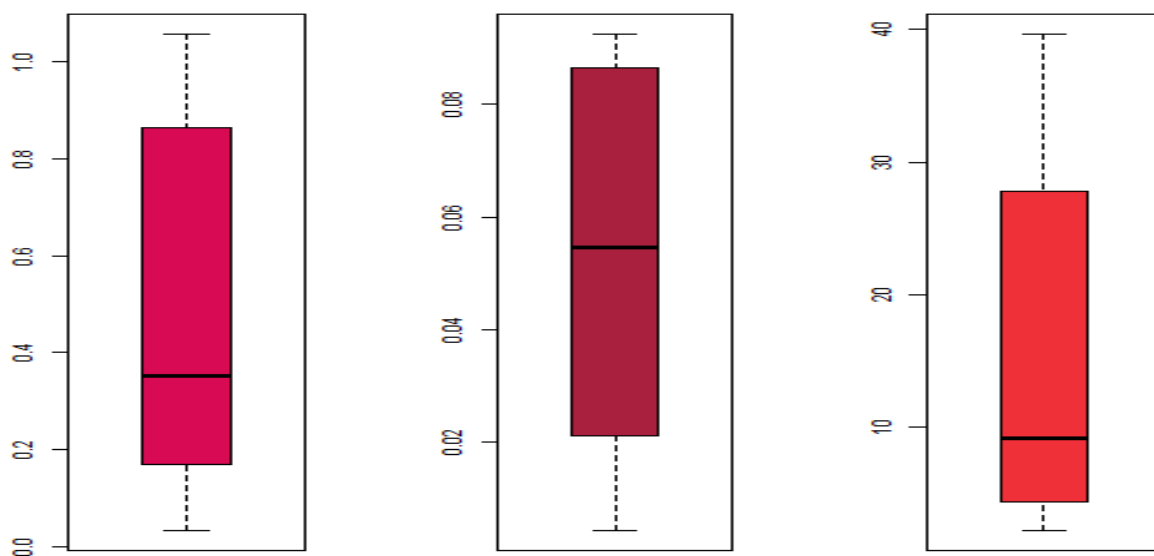
Tabela 21: Bezwzględny błąd oszacowania kwantyli rzędu 0.99

W tabelce (21) przy bezwzględnym błędzie oszacowania kwantyli rzędu 0.99 widzimy, że tak jak wcześniej, dla rozkładów stabilnych p_1 oraz p_2 oszacowania metodami parametrycznymi i nieparametrycznymi dają nam zbliżone wyniki. W przypadku rozkładu p_3 ponownie widzimy bardzo dużą rozbieżność między tymi oszacowaniami.

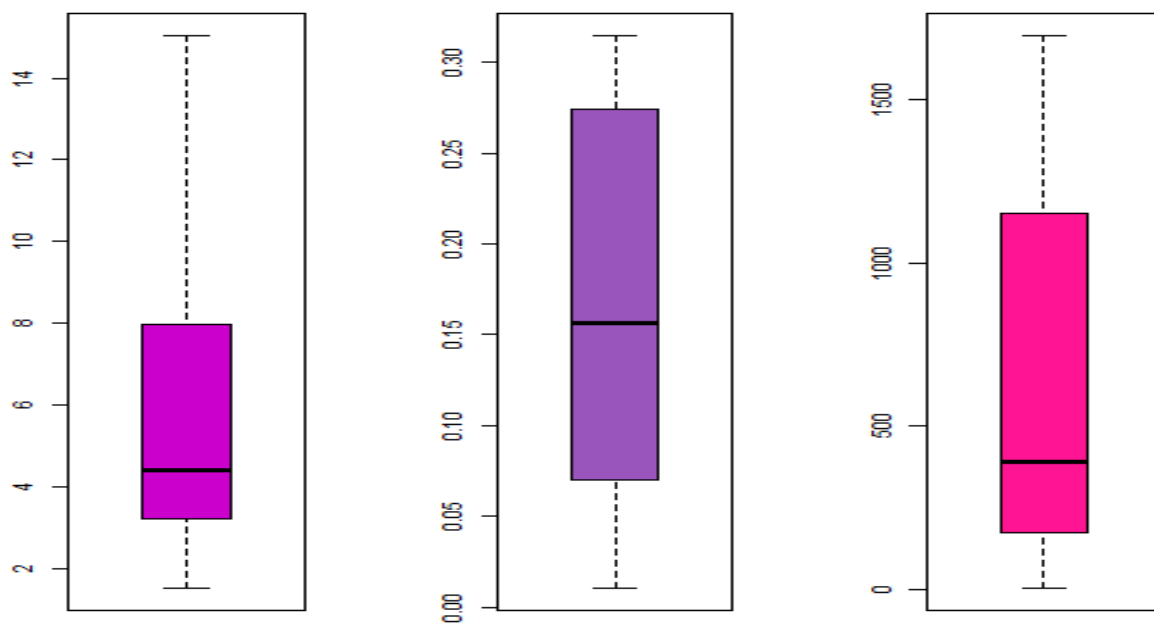
Reasumując, uważam, że metody parametryczne radzą sobie porównywalnie do metod nieparametrycznych. Nie ma między nimi dużej rozbieżności, zatem możemy używać ich obydwu bez obawy, że nasze wyniki wyjdą zakłamanie.

Dla kwantyli rzędu 0.95 najlepszą metodą parametryczną okazała się według mnie metoda 1P, czyli metoda McCullocha. Dawała ona najmniejsze bezwzględne błędy oszacowania kwantyli. Dla tego samego rzędu najlepszą nieparametryczną metodą okazała się metoda 3NP, czyli ciągły kwantyl próbkowy.

Dla kwantyli rzędu 0.99 jako najlepszą metodą parametryczną wybrałabym metodę 2P - metodę Koutrouvelisa, zaś jako nieparametryczną - ponownie metodę 3P.



Rysunek 6: Boxplot kwantyli rzędu 0.95 - opracowanie własne



Rysunek 7: Boxplot kwantyli rzędu 0.99 - opracowanie własne

Obserwując powyższe boxploty widzimy, że w każdym możliwych przypadku rozrzut danych jest bardzo duży.

4 ZADANIE 3

Pakiet `quantmod` dla R został zaprojektowany, aby pomóc handlowcom ilościowym w opracowywaniu, testowaniu i wdrażaniu modeli handlowych opartych na statystyce. `Quantmod` jest środowiskiem szybkiego prototypowania, w którym można szybko i czysto eksplorować i budować modele handlowe. Funkcja `GetSymbols` z tego pakietu łąduje oraz zarządza danymi pozyskanymi z wielu źródeł.

W oparciu o pakiet `quantmod`, w szczególności o funkcję `getSymbols` wybierzemy dane dotyczące otwarcia akcji trzech spółek (`BMW`, `APPLE`, `TESLA`) i na ich podstawie obliczymy dzienne logarytmiczne stopy zwrotu. Będziemy rozważali trzy przypadki rozkładów zwrotów logarytmicznych

- Przyjmujemy, że zwroty logarytmiczne mają rozkład normalny o nieznanych parametrach (Z1)
- Przyjmujemy, że zwroty logarytmiczne mają rozkład stabilny o nieznanych parametrach (Z2)
- Przyjmujemy, że zwroty logarytmiczne mają rozkład nieznany absolutnie ciągły (Z3)

Zanim przejdziemy do dokładnej analizy, zdefiniujemy kilka pojęć.

- Zwykłą stopą zwrotu (zyskiem netto z kapitału) nazywamy

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{P_t}{P_{t-1}} - 1$$

- Logarytmiczną stopą zwrotu nazywamy

$$r_t = \log(1 + R_t)$$

- Ustalmy, że ν jest przestrzenią liniową zmiennych losowych, określonych na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, \mathcal{P})$, nad ciałem liczb rzeczywistych. Odwzorowanie $\rho : \nu \rightarrow \mathbf{R}$ nazywamy miarą ryzyka, jeżeli jest:

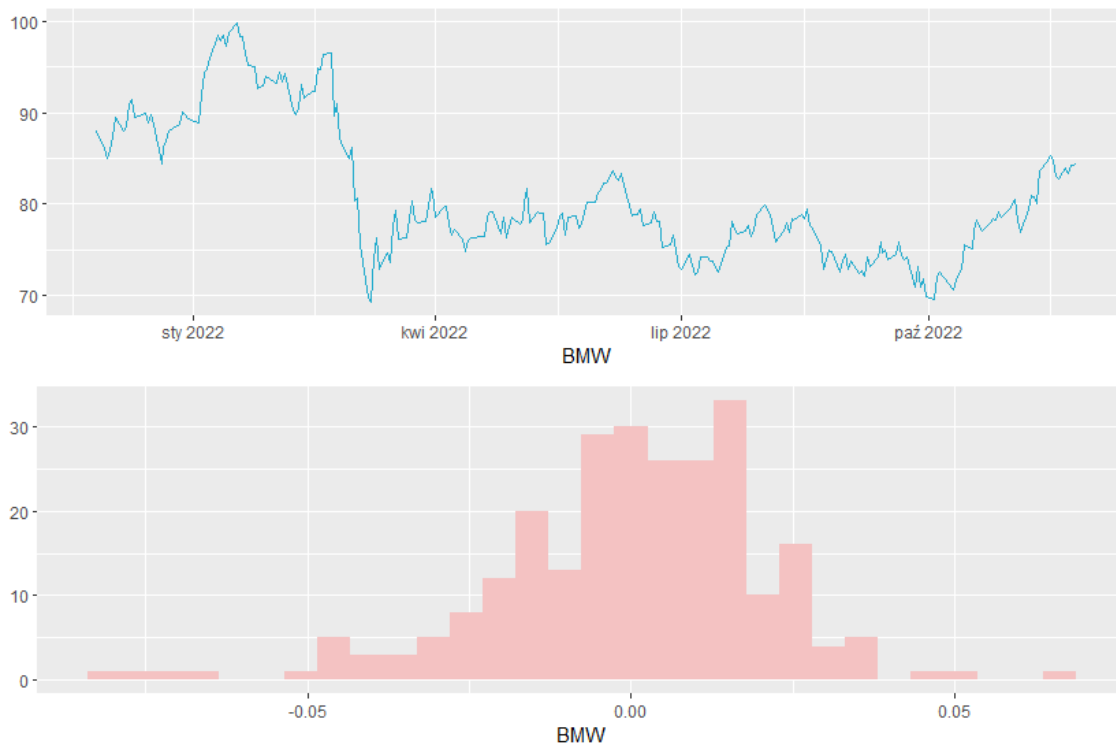
- znormalizowane
- monotoniczne
- nieimiennicze na przesunięcia

Najpopularniejszymi miarami ryzyka związanymi z losową stratą \mathcal{L} są

- wartość narażona na ryzyko, czyli inaczej wartość zagrożona ryzykiem. Oznaczamy $Var_\alpha(\mathcal{L})$
- oczekiwany deficyt (niedobór) oznaczany jako $ES_\alpha(\mathcal{L})$

Przypomnijmy ponadto, że cena otwarcia jest to cena akcji, po jakiej inwestorzy realizują transakcje zaraz po otwarciu się sesji notowań. Na ustalenie kursu otwarcia wpływa wiele czynników, między innymi zainteresowanie danym walorem. Cenę otwarcia ustala się na podstawie liczby przyjętych zleceń jeszcze przed otwarciem notowań.

BMW, czyli *Bayerische Motoren Werke Aktiengesellschaft* jest niemieckim koncernem motoryzacyjnym produkującym od 1916 roku samochody osobowe, motocykle, skutery oraz silniki. Poniżej przedstawiamy wykres cen otwarcia oraz histogram logarytmicznych stóp zwrotu akcji tej firmy.



Rysunek 8: Wykres cen otwarcia oraz histogram logarytmicznych stóp zwrotu BMW - opracowanie własne

W analizowanym (8) widzimy, że cena otwarcia na przełomie lutego/marca znacząco spadła. Możemy się domyślać, że spadek ten spowodowało rozpoczęcie się wojny na Ukrainie (24 lutego).

	VaR (Z1)	VaR (Z2)	VaR (Z3)	ES (Z1)	ES (Z2)	ES(Z3)	Alfa	Wartość
BMW	335.74	352.79	383.23	418.56	619.64	522.65	1.83	10000.00

Tabela 22: Oszacowania punktowe VaR oraz ES dla akcji BMW

Analizując wartość narażoną na ryzyko widzimy, że najmniejszą wartość dostajemy korzystając z metody Z1, czyli zakładamy, że zwroty pochodzą z rozkładu normalnego, zaś największe *VaR* otrzymujemy, gdy korzystamy z metody Z3 zakładającej, że zwroty pochodzą z rozkładu nieznanego absolutnie ciągłego. Przyglądając się oczekiwanemu niedoborowi *ES*, widzimy, że ponownie dla metody Z1 otrzymaliśmy wartość najmniejszą. Tym razem jednak, największą

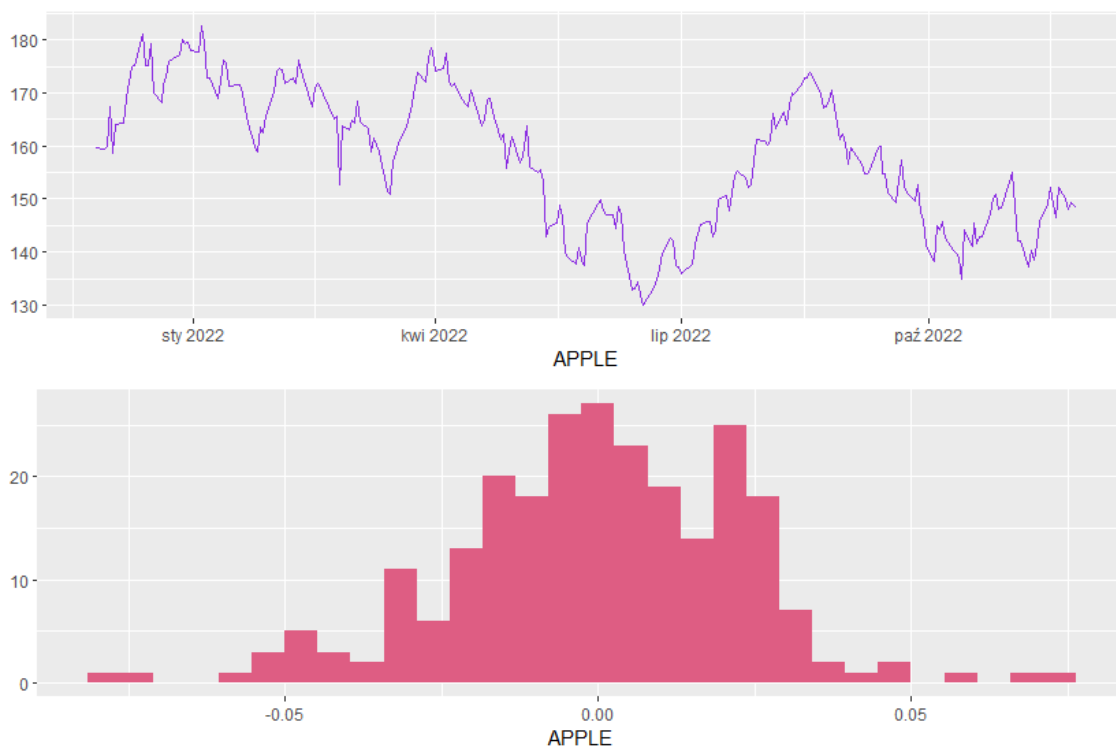
wartość otrzymaliśmy, zakładając, że zwroty pochodzą z rozkładu stabilnego o nieznanych parametrach (Z2).

	dolny BMW	górny BMW	różnica
Dokładne (Z3)	288.81	453.87	165.06
Bootstrap (Z1)	280.72	388.48	107.76
Bootstrap (Z2)	264.54	433.10	168.55
Bootstrap (Z3)	277.82	443.36	165.53

Tabela 23: Przedziały ufności BMW $\text{VaR} = 335.01$

Widzimy, że najmniejsza różnica między górnym a dolnym przedziałem ufności występuje dla Z1, zatem przy założeniu, że zwroty pochodzą z rozkładu normalnego.

Apple Inc. jest amerykańskim przedsiębiorstwem informatycznym zajmującym się projektowaniem i produkcją komputerów osobistych i mobilnych oraz oporowania i serwisów internetowych. Poniżej przedstawiamy wykres cen otwarcia oraz histogram logarytmicznych stóp zwrotu akcji tej firmy.



Rysunek 9: Wykres cen otwarcia oraz histogram logarytmicznych stóp zwrotu APPLE - opracowanie własne

	VaR (Z1)	VaR (Z2)	VaR (Z3)	ES (Z1)	ES (Z2)	ES(Z3)	Alfa	Wartość
AAPL	371.61	361.97	392.94	462.74	480.60	528.19	1.92	10000.00

Tabela 24: Oszacowania punktowe VaR oraz ES dla akcji APPLE

Analizując wartość narażoną na ryzyko widzimy, że w przypadku APPLE, najmniejszą wartość otrzymaliśmy zakładając, że zwroty pochodzą z rozkładu stabilnego o nieznanym parametrach, zaś największa - sugerując, że zwroty pochodzą z rozkładu normalnego. W przypadku oczekiwanego niedoboru, najmniejszą wartość otrzymaliśmy zakładając, że zwroty pochodzą z rozkładu normalnego, zaś największą - gdy pochodzą z rozkładu nieznanego absolutnie ciągłego.

	dolny APPLE	górny APPLE	różnica
Dokładne (Z3)	320.80	521.53	200.73
Bootstrap (Z1)	324.87	427.93	103.06
Bootstrap (Z2)	318.65	470.81	152.16
Bootstrap (Z3)	317.41	474.63	157.23

Tabela 25: Przedziały ufności APPLE VaR = 373.01

Podobnie jak dla BMW, otrzymaliśmy, że najmniejsze przedziały ufności firmy APPLE występują przy założeniu Z1, to znaczy, że kiedy zwroty pochodzą z rozkładu normalnego.

Tesla Inc. jest amerykańskim producentem samochodów elektrycznych, stacji ładowania, paneli fotowoltaicznych oraz akumulatorów. Poniżej przedstawiamy wykres cen otwarcia oraz histogram logarytmicznych stóp zwrotu akcji tej firmy.



Rysunek 10: Wykres cen otwarcia oraz histogram logarytmicznych stóp zwrotu TESLA - opracowanie własne

	VaR (Z1)	VaR (Z2)	VaR (Z3)	ES (Z1)	ES (Z2)	ES(Z3)	Alfa	Wartość
AAPL	701.16	688.55	701.98	863.86	897.31	926.00	1.96	10000.00

Tabela 26: Oszacowania punktowe VaR oraz ES dla akcji TESLA

Prawie analogicznie jak w przypadku APPLE. Najmniejsze wartości VaR i ES otrzymaliśmy zakładając kolejny zwroty pochodzące z rozkładu stabilnego o nieznanym parametrach i normalnego. Największe zaś zakładając, że zwroty pochodzą z rozkładu nieznanego absolutnie ciągłego.

	dolny TESLA	górny TESLA	różnica
Dokładne (Z3)	585.44	825.92	240.48
Bootstrap (Z1)	617.09	805.67	188.58
Bootstrap (Z2)	608.59	816.40	207.80
Bootstrap (Z3)	1348.15	1849.03	500.88

Tabela 27: Przedziały ufności TESLA $VaR = 702.69$

Identycznie jak w poprzednich przypadkach, najmniejszy przedział ufności dostaliśmy przy założeniu, że rozkładu pochodzą z rozkładu normalnego.

Reasumując, największą wartość narażoną na ryzyko otrzymujemy w każdym z przypadków przy założeniu, że rozkłady pochodzą z rozkładu nieznanego absolutnie ciągłego.

Logarytmiczne stopy zwrotu mają kilka własności, ze względu na które są częściej stosowane w ekonometrii finansowej i matematyce finansowej. Logarytmiczne stopy zwrotu są addytywne i można dla nich obliczać średnią arytmetyczną, która odpowiada kapitalizacji ciągłej. Cena akcji nie może spaść poniżej zera, dlatego zwykle stopy zwrotu mogą teoretycznie przyjmować wartości od minus jeden do plus nieskończoności. Teoretyczny rozkład prostych stóp zwrotu powinien być zatem rozkładem uciętym z lewej strony. Logarytmiczne stopy zwrotu mogą teoretycznie przyjmować wartości od minus do plus nieskończoności.

5 ZADANIE DODATKOWE 1

Najczęściej spotykaną metodą estymacji jest metoda największej wiarygodności. W procesie estymacji na podstawie prób X_i wyznaczamy parametr λ opisujący domniemany rozkład prawdopodobieństwa. Na podstawie tego rozkładu możemy określić a posteriori prawdopodobieństwo próby X_i . Parametr λ dopieramy tak, aby zmaksymalizować prawdopodobieństwo a posteriori prób, z których je wyznaczamy. Funkcją wiarygodności nazywać możemy iloczyn prawdopodobieństwa a posteriori dla N dostępnych prób

$$L = \prod_{i=1}^N f(x_i, \lambda)$$

$$l = \ln(L) = \sum_{i=1}^N \ln f(x_i, \lambda)$$

Szukamy jej maksimum, czyli zera pochodnej. Poniżej przedstawiamy przykład zaczerpnięty ze źródła [1].

Spośród studentów informatyki pewnego elitarnego wydziału wybrano losowo i niezależnie od siebie 50 osób, a następnie każdą z nich spytano, czy kiedykolwiek w trakcie studiów otrzymała ocenę niedostateczną. Okazało się, iż 14 osób odpowiedziało "TAK", natomiast pozostałe odpowiedziały "NIE". Pytamy teraz: jaki procent studentów informatyki otrzymał w trakcie swoich studiów ocenę niedostateczną.

Obserwujemy próbkę X_1, \dots, X_n , dla $n = 50$ z rozkładu dwupunktowego $(0, 1, p)$. Odpowiedź *NIE* interpretujemy jako wartość 0, zaś odpowiedź *TAK* jako wartość 1. Naszym zadaniem będzie wskazanie parametru p . Obliczamy prawdopodobieństwo wylosowania naszej próbki w zależności od p , a następnie uznajemy, że najlepszym przybliżeniem nieznanego parametru będzie taka wartość p , dla której obliczone właśnie prawdopodobieństwo jest największe.

Korzystając z niezależności zmiennych losowych otrzymujemy:

$$P(X_1 = x_1, \dots, X_n = x_n) = P(X_1 = x_1) \dots P(X_n = x_n)$$

$$P(X_i = x_i) = \begin{cases} p, & \text{dla } i = 1 \\ 1 - p, & \text{dla } i = 0 \end{cases}$$

Dalej wiemy, że 14 osób odpowiedziało *TAK*, więc $50 - 14 = 36$ osób odpowiedziało *NIE*. Stąd

$$P(X_1 = x_1, \dots, X_n = x_n) = p^{14}(1 - p)^{36}$$

$$l(p) = p^{14}(1 - p)^{36}$$

$$L(p) = \ln l(p) = 14 \ln p + 36 \ln(1 - p)$$

$$L'(p) = \frac{14}{p} - \frac{36}{1 - p}$$

$$L'(p) = 0 \rightarrow \frac{14}{p} - \frac{36}{1 - p} = 0$$

Ostatecznie otrzymujemy, że $\hat{p} = \frac{14}{50} = 0.28$, więc 28% studentów informatyki otrzymało w trakcie swoich studiów ocenę niedostateczną. Otrzymany w ten sposób estymator \hat{p} nazywamy estymatorem największej wiarygodności parametru p .

Literatura

- [1] Rachunek prawdopodobieństwa i statystyka/Wykład 12: Metoda największej wiarygodności, https://wazniak.mimuw.edu.pl/index.php?title=Rachunek_prawdopodobie%C5%84stwa_i_statystyka/Wyk%C5%82ad_12:_Metoda_najwi%C4%99kszej_wiarygodno%C5%9Bci
- [2] Alicja Jokiel - Rokita *Statystyka w finansach i ubezpieczeniach - WYKŁADY*
- [3] Piotr Fiszeder *Modele klasy Garch w empirycznych badaniach finansowych*, http://www.home.umk.pl/~piter/Modele_klasy_Garch.pdf