Projekt Egzaminacyjny

Jan Milewczyk, Maciej Wojciechowski, Kajetan Lach

December 3, 2024

1 Analizy log-zwrotów spółek

1.1 Vigo Photonics

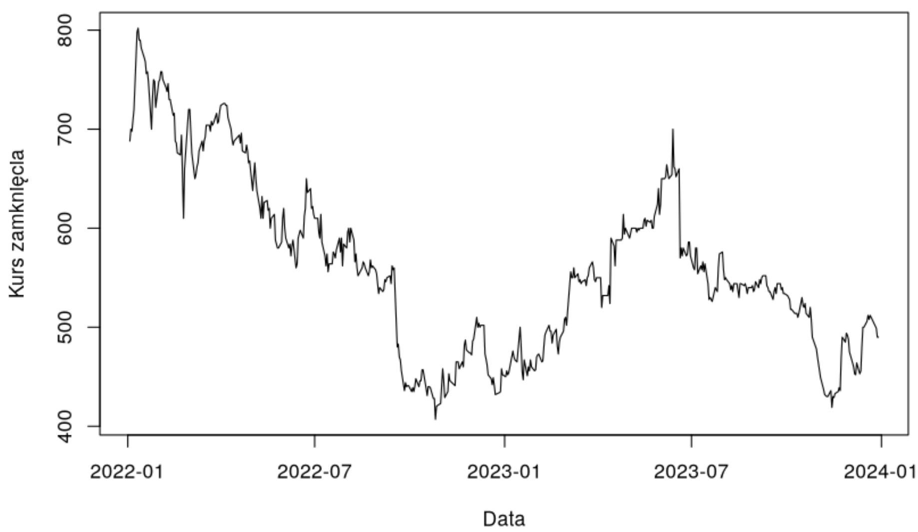
1.1.1 Wstęp

Vigo Photonics to polskie przedsiębiorstwo specjalizujące się w wytwarzaniu materiałów i przyrządów półprzewodnikowych do zastosowań fotonicznych i mikroelektronicznych. Spółka jest liderem na światowym rynku fotonowych detektorów średniej podczerwieni, a wszystkie produkty opiera na własnej, unikalnej technologii.

1.1.2 Analiza log-zwrotów spółki (Vigo Photonics)

Pierwszy rozdział zawiera analizę log-zwrotów spółki.

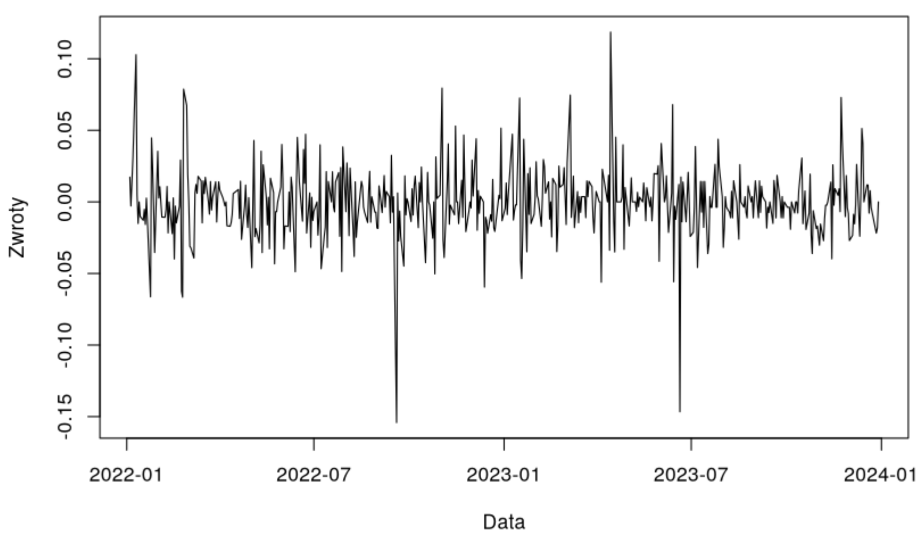
Wykresy kursów zamknięcia oraz log-zwrotów: Poniższy wykres ilustruje zmianę cen zamknięcia akcji w czasie.



Zmianę log-zwrotów, wyliczonych według wzoru

$$r_1 = \ln \frac{S_0}{S_1}, r_2 = \ln \frac{S_2}{S_1}, ..., r_n = \ln \frac{S_n}{S_{n-1}}$$

gdzie $s_0, s_1, ..., s_n$ są kursami zamknięcia z kolejnych dni, na osi czasu ilustruje wykres poniżej.



Wartość oczekiwana: Zakładając, że log-zwroty $r_1, r_2, ..., r_n$ są niezależnymi realizacjami zmiennej losowej X, wyliczamy μ przy użyciu nieobciążonego estymatora wartości oczekiwanej:

$$E(\overline{X}_n) = \frac{1}{n}\sum_{i=1}^{n} X_i = \mu$$

w wyniku czego otrzymujemy wynik

$$\mu = -0.0006787669$$

Wariancja i odchylenie standardowe: Korzystając ze wzoru na nieobciążony estymator wariancji

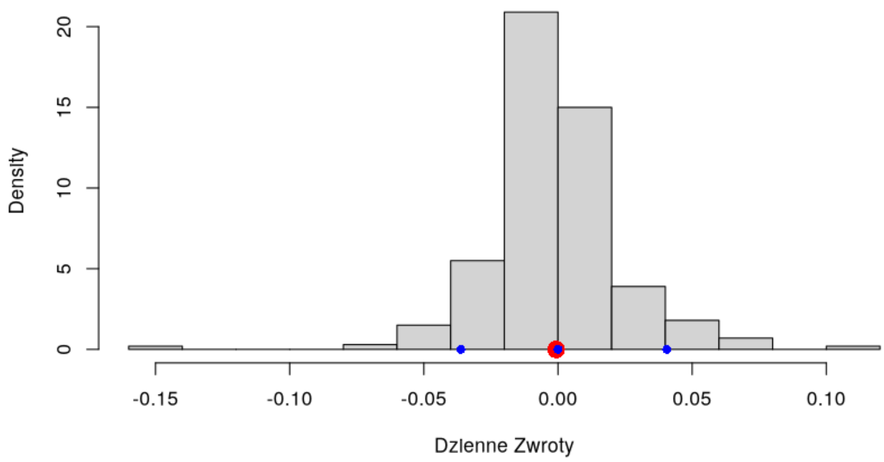
$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$$

otrzymujemy $\sigma^2 = 0.0006189792$ oraz $\sigma = 0.02487929$.

Kwantyle: Z wykorzystaniem klasycznego estymatora kwantyli wyestymowano kwantyle rzędu $\alpha = 5\%, 50\% i 95\%$. Wyniki przedstawione zostaly w tabeli poniżej:

$\overline{x}_n$	s_n^2	s_n	$q(5\%)$	$q(50\%)$	$q(95\%)$
-0.0006787669	0.0006189792	0.02487929	-0.03620640	0	0.04055989

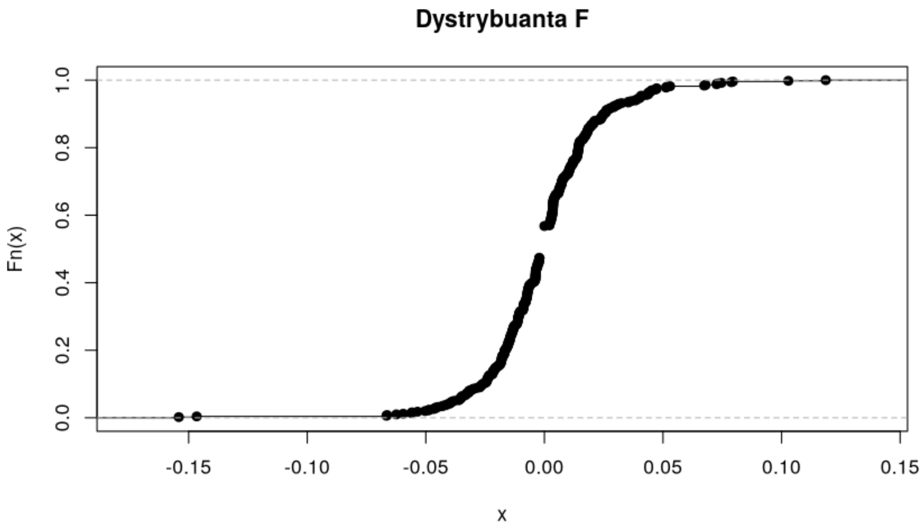
Histogram log-zwrotów: Na histogramie dziennych log-zwrotów cen akcji na czerwono i niebiesko oznaczono odpowiednio wartość wyestymowanej średniej oraz wartości kwantyli przedstawionych wcześniej.



Dystrybuanta: Z wykorzystaniem dystrybuanty empirycznej jako nieobciążonego estymatora

$$\hat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{x_i \leq x\}} = \left(\frac{\text{liczba elementów w próbie} \leq x}{n}\right)$$

estymujemy dystrybuantę F zaprezentowaną na wykresie poniżej

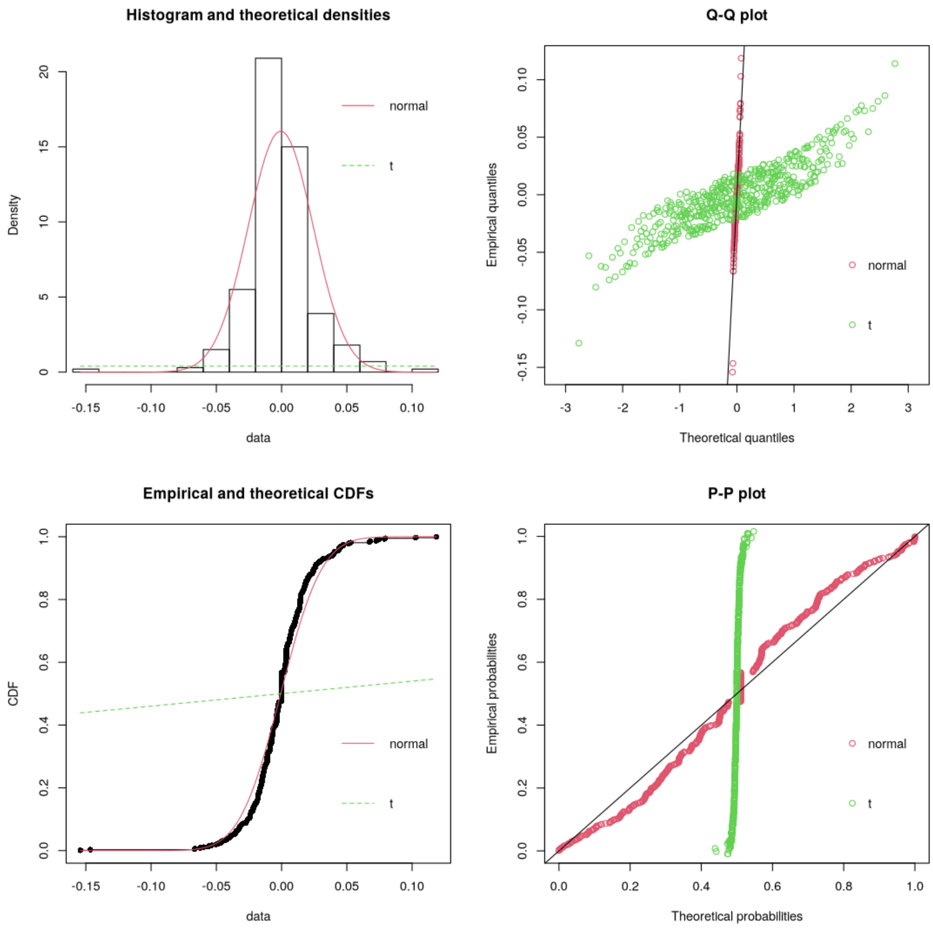


Analiza dobroci dopasowania rozkładu normalnego i t-Studenta: Korzystając z **estymatora największej wiarygodności (MLE)** parametru estymujemy parametry rozkładu normalnego i t-Studenta.

m	s	df
-0.0006787669	0.0248544	314.3698

1.1.3 Wykresy diagnostyczne

Poniższe wykresy prezentują kolejno porównania: histogram-gęstość wybranych rozkładów, kwantyl-kwantyl, dystrybuanta empiryczna-dystrybuanta teoretyczna wybranych rozkładów oraz prawdopodobieństwo teoretyczne wybranych rozkładów-prawdopodobieństwo empiryczne.



Na ich podstawie można stwierdzić, że rozkład normalny lepiej dopasowuje się do danych.

Weryfikacja wyboru z wykorzystainem statystyk oraz kryteriów informacyjnych: Poniższe tabele prezentują wyniki testów statystycznych oraz kryteriów informacyjnych dla rozkładów normalnego i t-Studenta.

Statystyki	Rozkład normalny	Rozkład t-Studenta
Kolmogorov-Smirnov	0.08251589	0.469477
Cramer-von Mises	1.18780300	39.177553
Anderson-Darling	6.95700925	183.184800

Kryteria	Rozkład normalny	Rozkład t-Studenta
AIC	-2271.782	922.0439
BIC	-2263.353	926.2585

Powyższe wyniki potwierdzają, że rozkład normalny lepiej dopasowuje się do danych.

Test hipotezy równości rozkładów metodą Monte Carlo: Testujemy hipotezę zerową o równości dystrybuant

$$H_0 : F = F_0$$

przeciwko hipotezie alternatywnej (kontrhipotezie)

$$H_1 : F \neq F_0$$

gdzie F_0 jest dystrybuantą rozkładu normalnego wybranego w poprzedniej podsekcji, a F dystrybuantą nieznaną.
Do przetestowania hipotezy zerowej o równości dystrybuant wykorzystamy metodę Monte Carlo przy użyciu statystyki Kołmogorowa-Smirnowa.
W tym celu generujemy $N = 10000$ prób z rozkładu normalnego o parametrach

$$m = -0.0006787669 \text{ oraz } s = 0.0248544$$

wyestymowanych wcześniej oraz obliczamy wartość statystyki Kołmogorowa-Smirnowa dla każdej z nich.
Następnie obliczamy prawdopodobieństwo

$$p = P(D_n > d_n)$$

gdzie D_n to wartość statystyki dla wygenerowanych prób a d_n to jej wartość dla log-zwrotów spółki.
Ustalamy poziom istotności $\alpha = 0.05$ i sprawdzamy czy $p < \alpha$. W naszym przypadku

$$p = 0.0025 < \alpha = 0.05$$

co pozwala nam odrzucić hipotezę zerową na rzecz hipotezy alternatywnej.

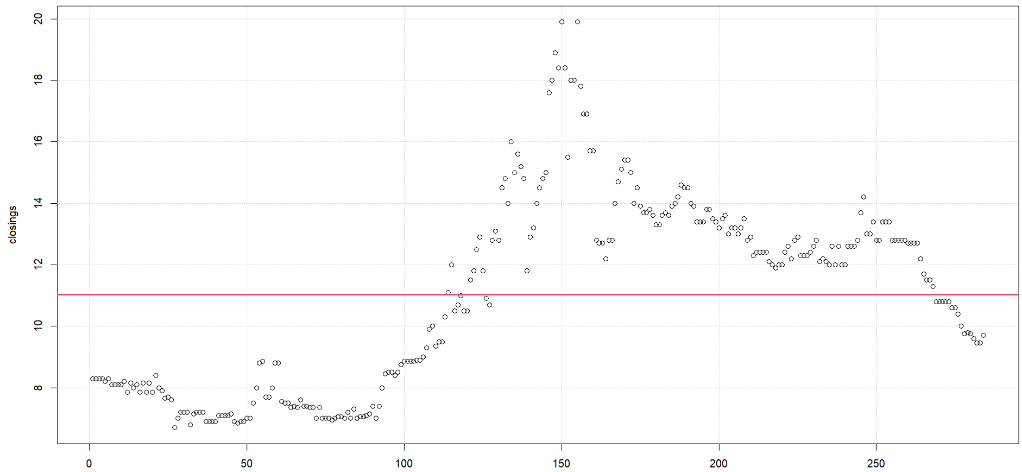
1.2 Digitree

1.2.1 Wstęp

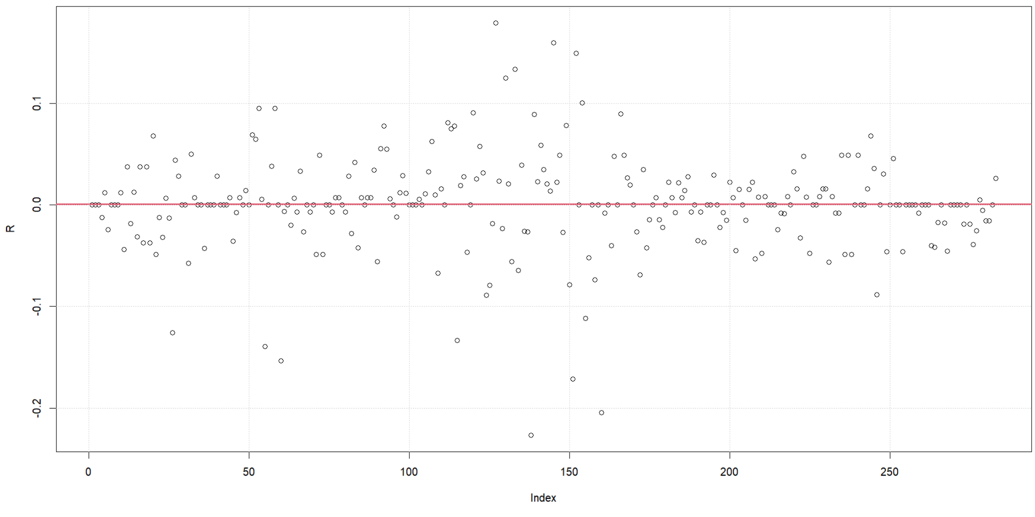
Digitree to polska spółka technologiczna specjalizująca się w kompleksowych rozwiązaniach z zakresu digital marketingu i wsparcia sprzedaży online.

1.2.2 Analiza log-zwrotów spółki DIGITREE

Spółka DIGITREE: Poniżej przedstawiam wykres kursów zamknięcia wybranej spółki



Poniżej przedstawiam wykres log-zwrotów wybranej spółki



Podstawowa analiza statystyczna log-zwrotów: Wartości log-zwrotów zostały obliczone jako różnice logarytmiczne dziennych kursów zamknięcia. Poniżej przedstawiono podstawowe statystyki log-zwrotów:

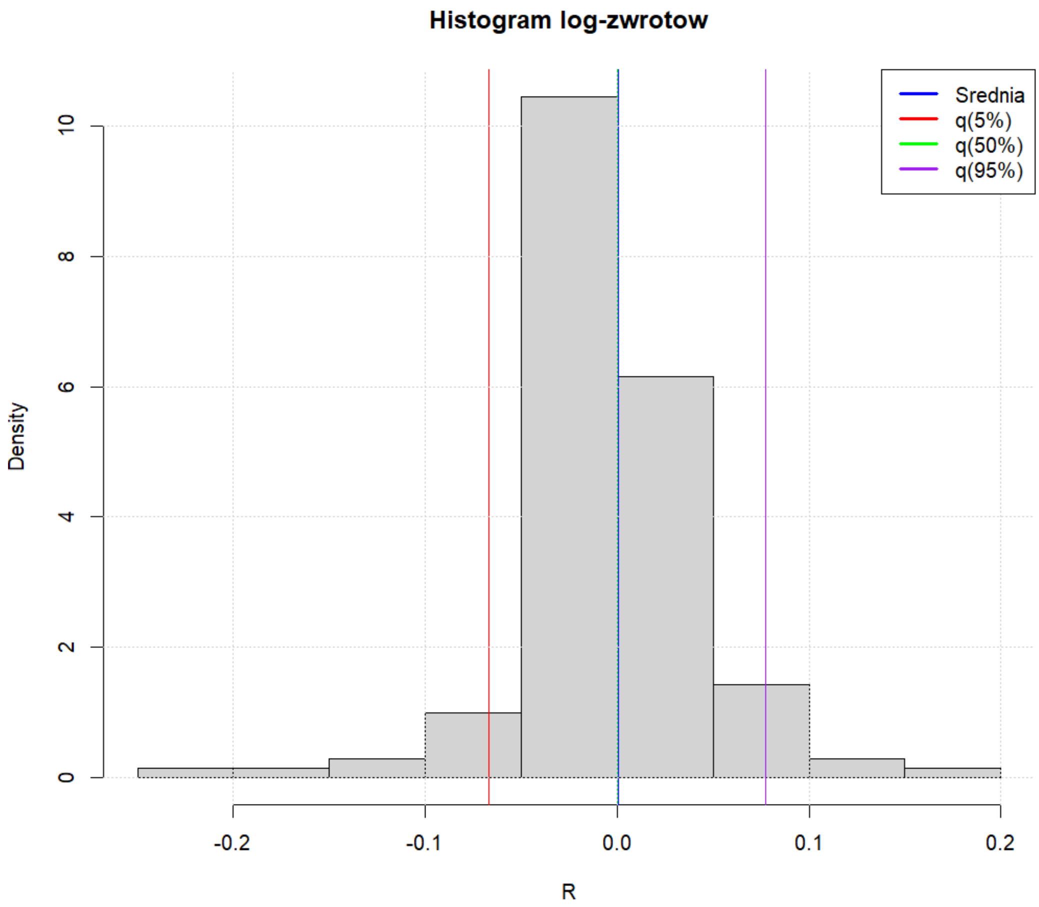
- Średnia log-zwrotów: $\overline{x}_n = 0.0005507787$
- Wariancja log-zwrotów: $s_n^2 = 0.0022445$
- Odchylenie standardowe log-zwrotów: $s_n = 0.04737616$

Table 1: Estymacja parametrów log-zwrotów

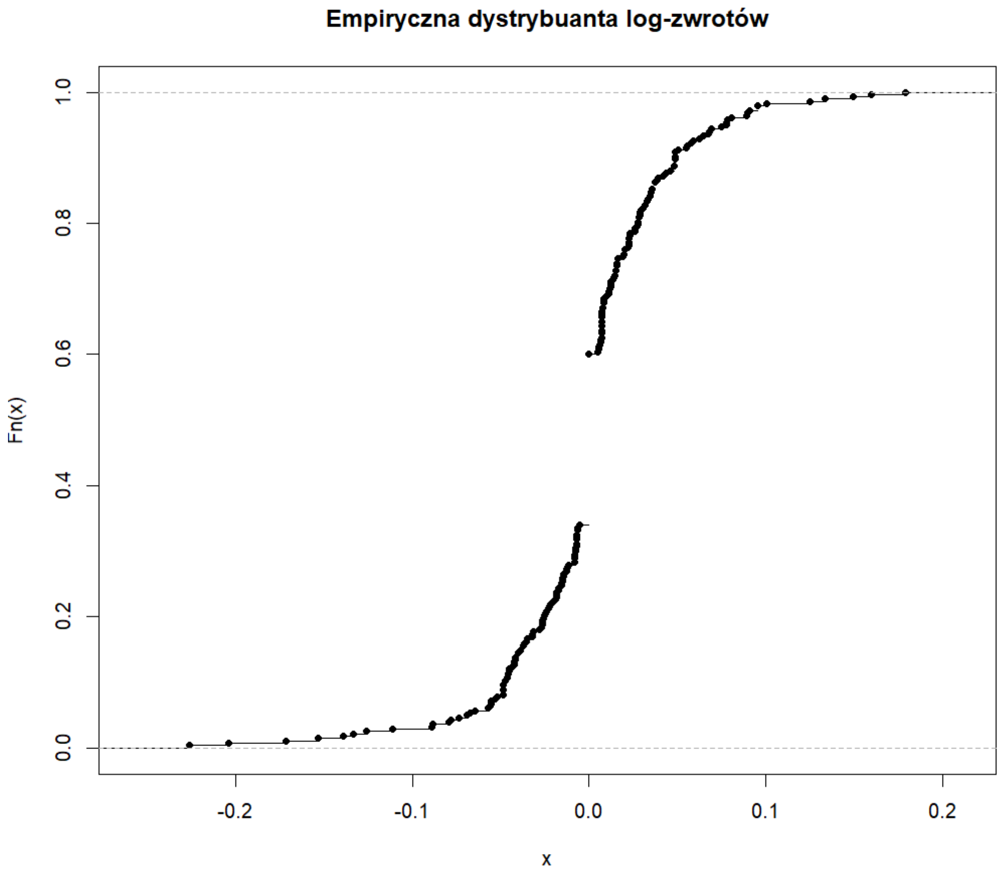
$\overline{x}_n$	s_n^2	s_n	$q(5\%)$	$q(50\%)$	$q(95\%)$
0.0005507787	0.0022445	0.04737616	-0.06694173	0.00000000	0.07764551

Kwantyle 5

Histogram log-zwrotów z zaznaczoną średnią i kwantylami: Poniżej zamieszczono histogram log-zwrotów z zaznaczonymi wartościami średniej oraz kwantyli 5



Dystrybuanta empiryczna log-zwrotów: Estymacja dystrybuanty empirycznej dla log-zwrotów została przeprowadzona przy użyciu wzoru empirycznej dystrybuanty $F_n(x)$. Wykres dystrybuanty empirycznej log-zwrotów znajduje się poniżej.

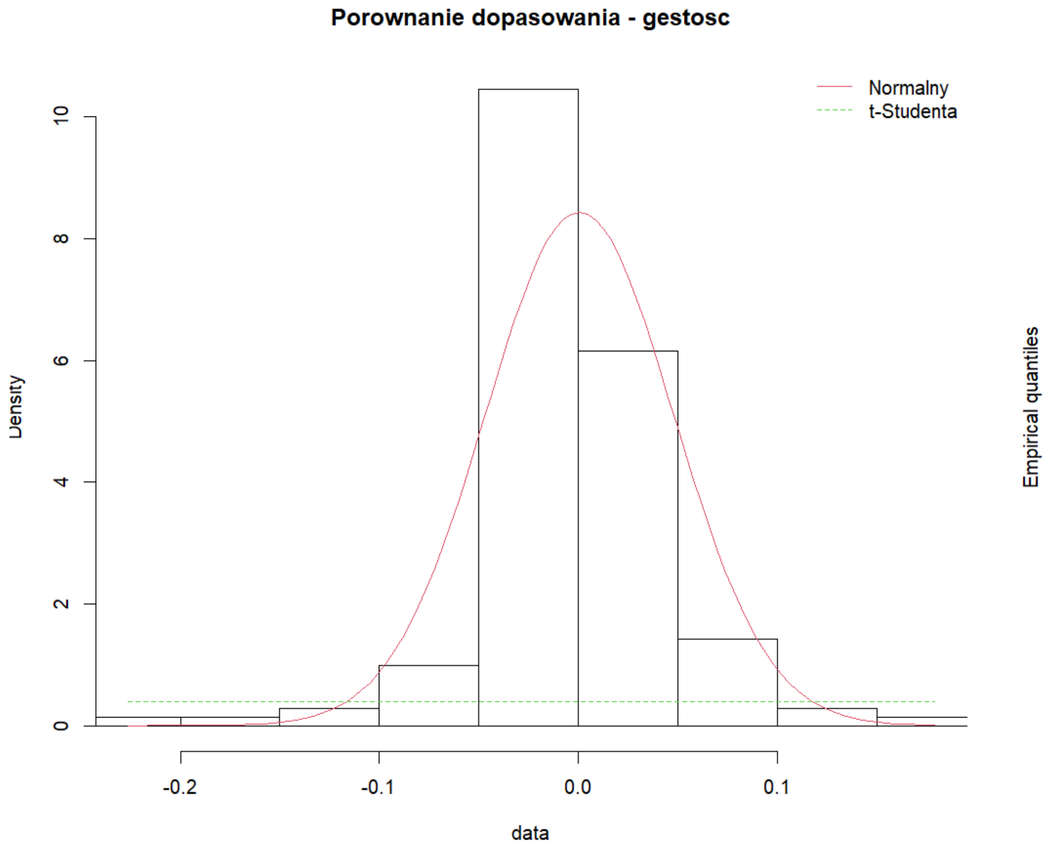


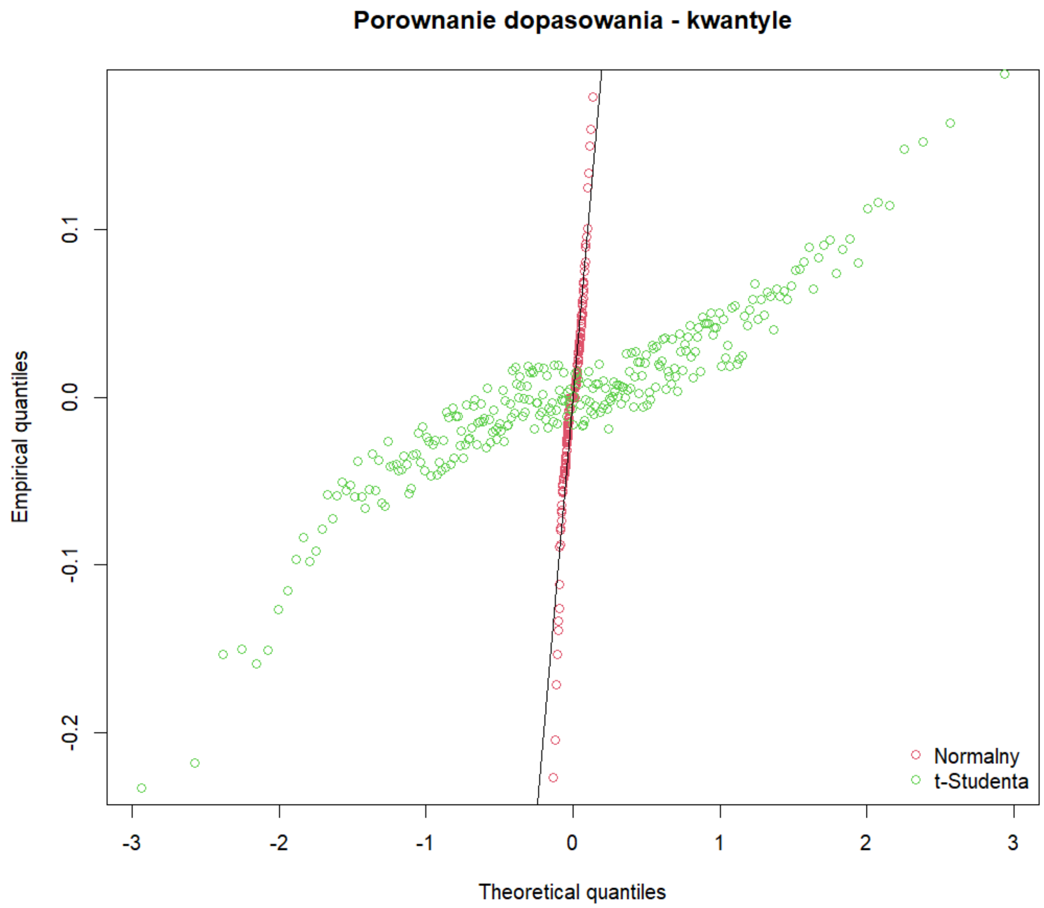
1.2.3 Analiza dobroci dopasowania rozkładów

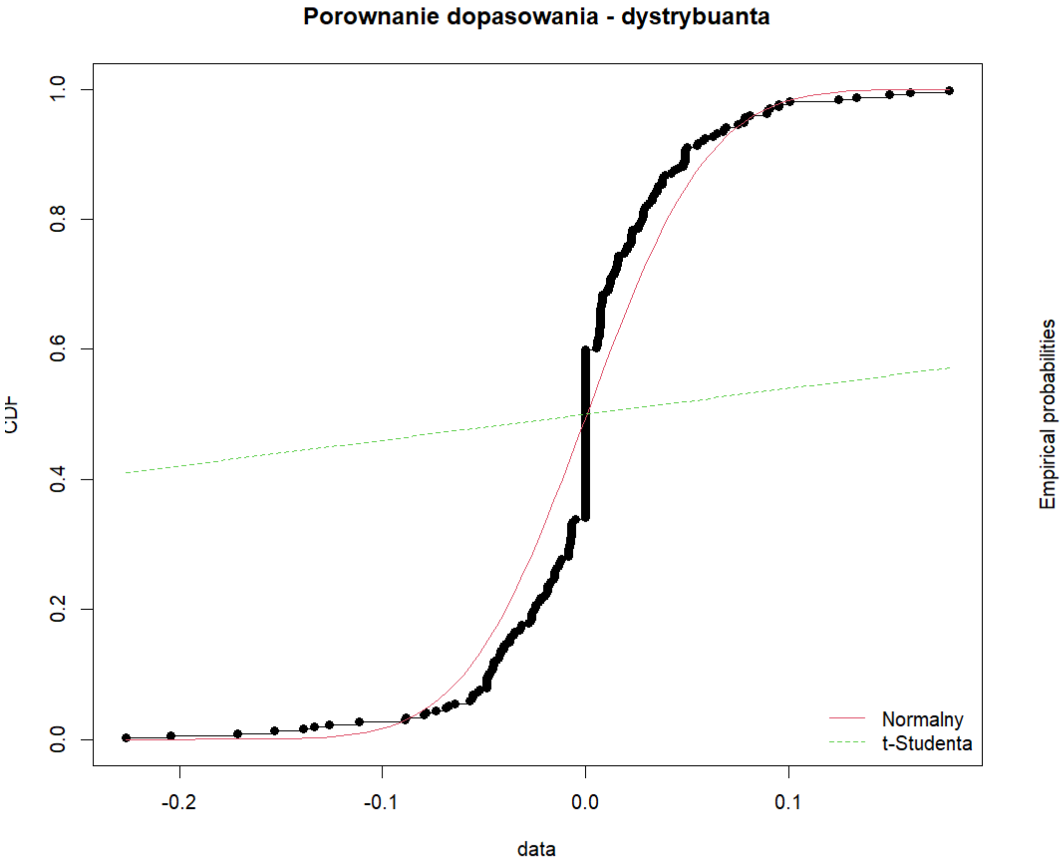
Estymacja parametrów rozkładu normalnego i t-Studenta: Parametry rozkładu normalnego i t-Studenta zostały wyestymowane przy użyciu estymatora największej wiarygodności (MLE):

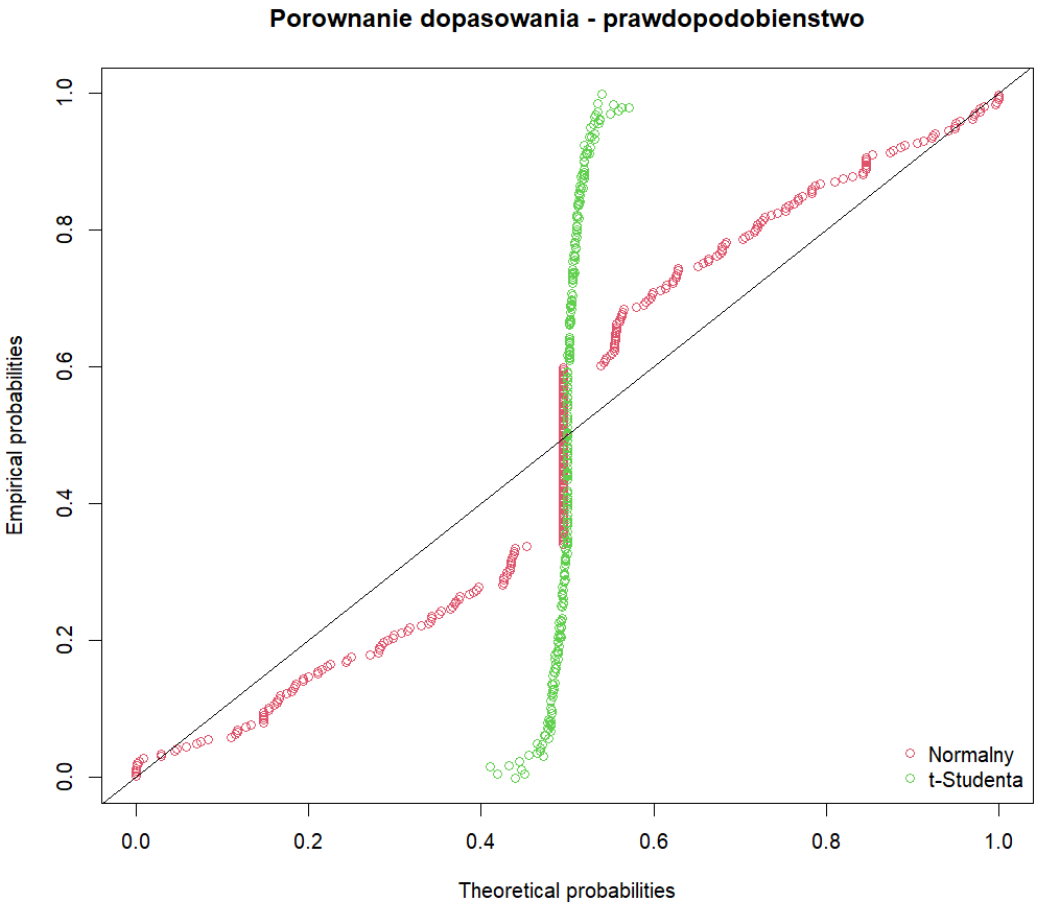
- Rozkład normalny: średnia = 0.0005507787, odchylenie standardowe = 0.0472923802
- Rozkład t-Studenta: liczba stopni swobody = 307.2041

Wykresy diagnostyczne: Poniżej przedstawiono wykresy diagnostyczne dla dopasowania rozkładów normalnego i t-Studenta do danych log-zwrotów, umożliwiające ocenę jakości dopasowania.


Porownanie dopasowania - gestosc
Normalny
t-Studenta
Density
0
2
4
6
8
10
-0.2
-0.1
0.0
0.1
data
Empirical quantiles


Porownanie dopasowania - kwantyle
Empirical quantiles
Theoretical quantiles
0.1
0.0
-0.1
-0.2
-3
-2
-1
0
1
2
3
Normalny
t-Studenta


Porownanie dopasowania - dystrybuanta
CDF
Empirical probabilities
data
Normalny
t-Studenta
-0.2
-0.1
0.0
0.1
0.0
0.2
0.4
0.6
0.8
1.0



1.3 Ocena dopasowania rozkładów

Na podstawie wykresów diagnostycznych oraz statystyk dopasowania:

- Statystyki dla rozkładu normalnego: KS = 0.1561, CM = 1.8659, AD = 9.3673, AIC = -919.9766, BIC = -912.6857
- Statystyki dla rozkładu t-Studenta: KS = 0.4424, CM = 21.0334, AD = 99.0999, AIC = 523.2149, BIC = 526.8603

Wyniki sugerują, że rozkład normalny lepiej dopasowuje się do danych log-zwrotów niż rozkład t-Studenta. Wybór rozkładu normalnego uzasadniają niższe wartości AIC i BIC oraz lepsze dopasowanie na wykresach diagnostycznych.

Test hipotezy o równości rozkładów: Przeprowadzono test hipotezy o równości rozkładów dla wybranego rozkładu normalnego, wykorzystując statystykę Kolmogorova-Smirnova (KS). Wyniki testu przedstawiono poniżej:

- **Statystyka testowa D:** Obliczona wartość wynosi $D = 0.1561313$. Statystyka D mierzy maksymalną różnicę pomiędzy dystrybuantą empiryczną danych a teoretyczną dystrybuantą rozkładu normalnego. Wysoka wartość D wskazuje na większą różnicę, co może sugerować, że dane nie pochodzą z badanego rozkładu.

- **P-wartość (p):** Obliczona p-wartość wynosi $p = 0$. Oznacza to, że przy założeniu prawdziwości hipotezy zerowej, prawdopodobieństwo uzyskania tak dużej lub większej różnicy D wynosi praktycznie zero. W rezultacie, hipoteza zerowa o zgodności rozkładu danych z rozkładem normalnym zostaje odrzucona na dowolnym poziomie istotności.

Interpretacja:

- Obliczona wartość statystyki D sugeruje, że istnieją istotne różnice pomiędzy rozkładem danych a teoretycznym rozkładem normalnym.

- Niska p-wartość ($p = 0$) wskazuje, że te różnice są statystycznie istotne. Oznacza to, że dane najprawdopodobniej nie pochodzą z badanego rozkładu normalnego.

1.4 Wawel

2 Analiza łącznego rozkładu log-zwrotów