

# Statystyka w finansach i ubezpieczeniach

## Sprawozdanie 1

Justyna Niedźwiedzka 229877

18 listopada 2020

# Część 1

Na podstawie symulacji dokonamy porównania czterech metod estymacji parametrów rozkładów stabilnych. Przyjmujemy rozmiar próby  $n = 100$  i  $n = 250$  oraz parametry  $(\alpha, \beta, \sigma, \mu) = (1, 0, 1, 0)$ ,  $(\alpha, \beta, \sigma, \mu) = (2, 0, 1, 0)$  i  $(\alpha, \beta, \sigma, \mu) = (1, 0, 3, 0)$ . Parametry definiujemy następująco:

```
pars1 <- c(1, 0, 1, 0)
pars2 <- c(2, 0, 1, 0)
pars3 <- c(1, 0, 3, 0)
```

Poniżej przedstawiono funkcje, które zwracają oszacowane parametry rozkładów stabilnych dla każdej z czterech metod.

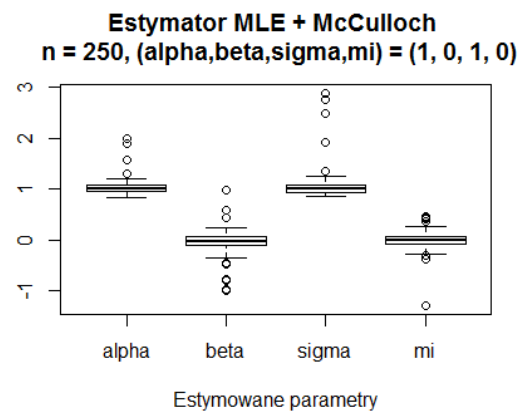
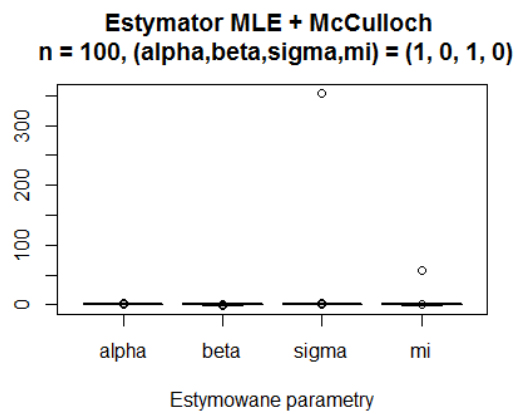
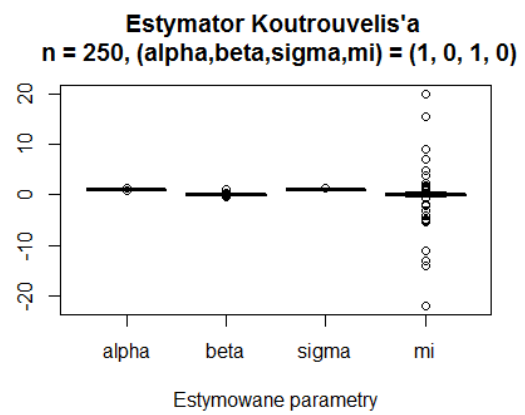
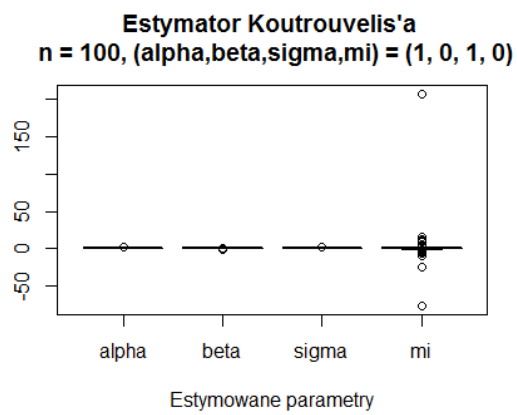
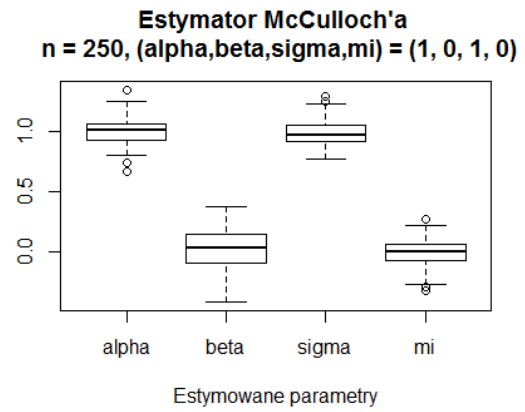
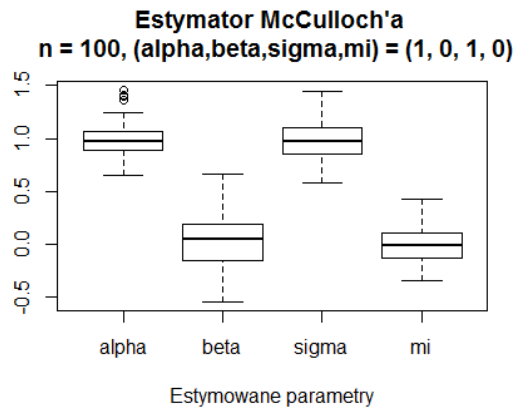
```
McCulloch <- function(n,pars){
  rnd <- stable_rnd(n,pars)
  pars_est_M <- stable_fit_init(rnd)
  return(pars_est_M)
}

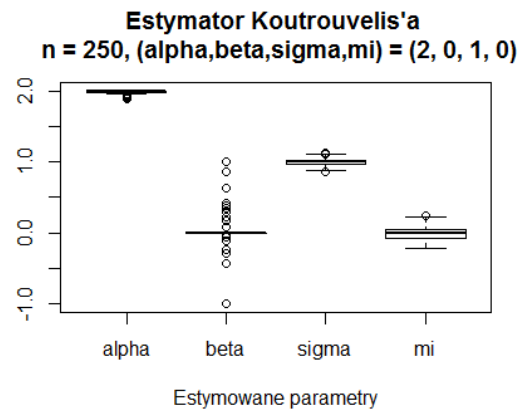
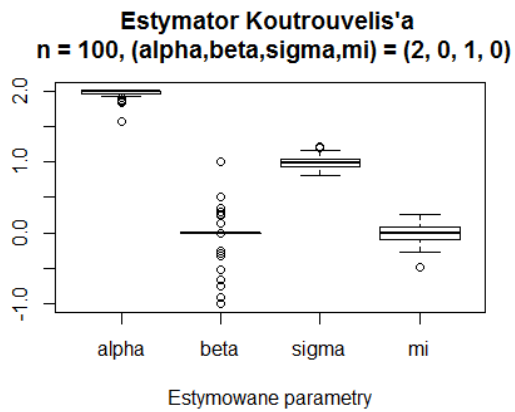
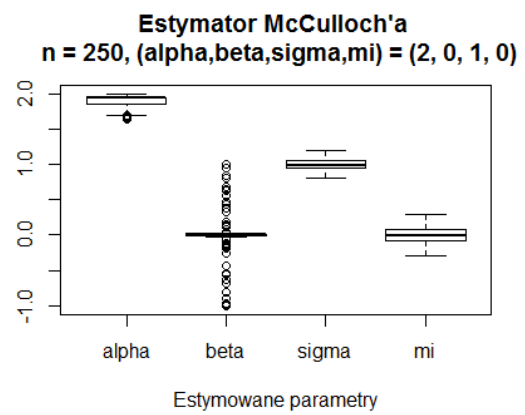
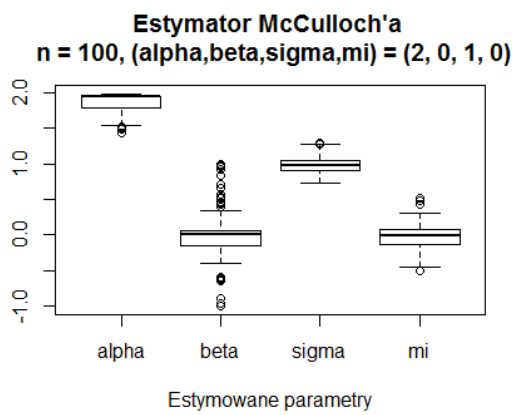
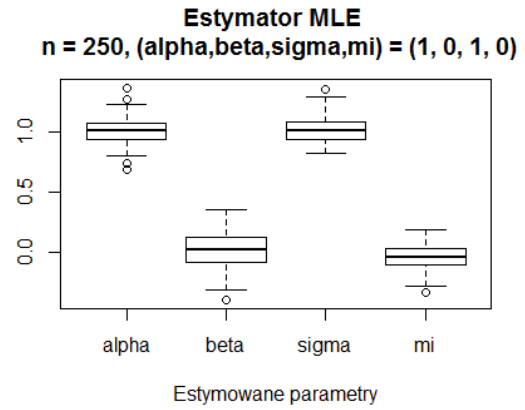
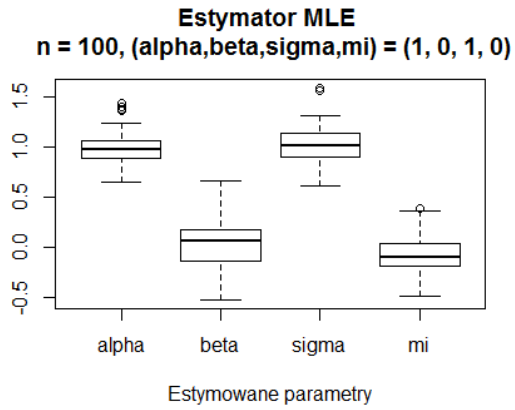
Koutrouvelis <- function(n,pars){
  rnd <- stable_rnd(n,pars)
  pars_est_M <- stable_fit_init(rnd)
  pars_est_K <- stable_fit_koutrouvelis(rnd, pars_est_M)
  return(pars_est_K)
}

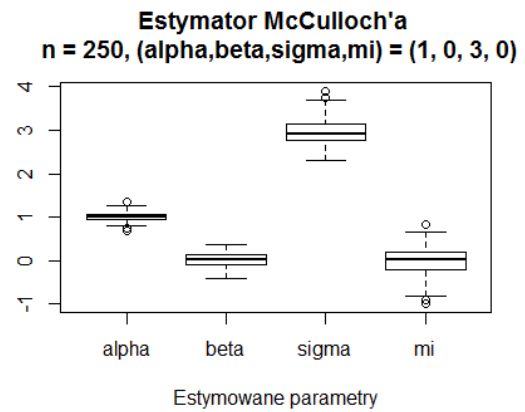
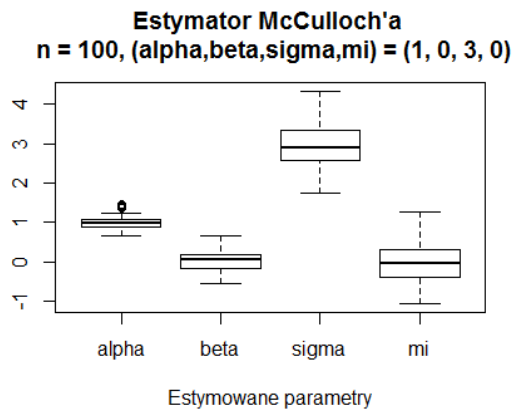
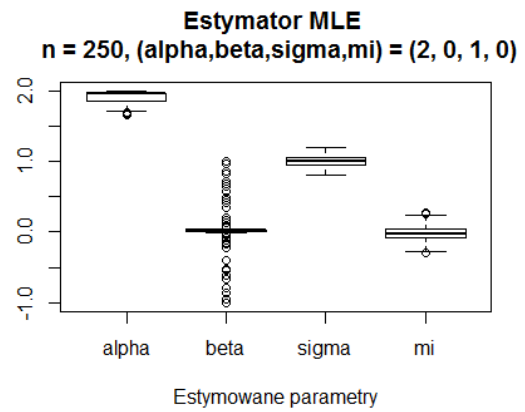
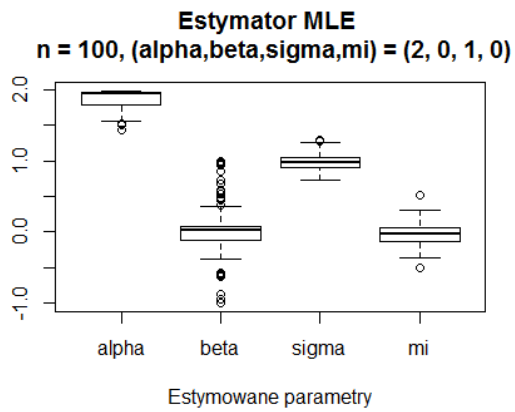
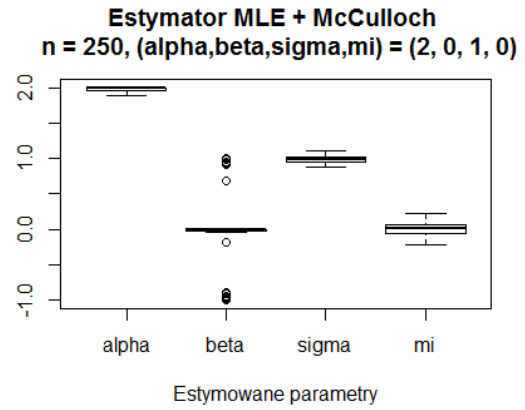
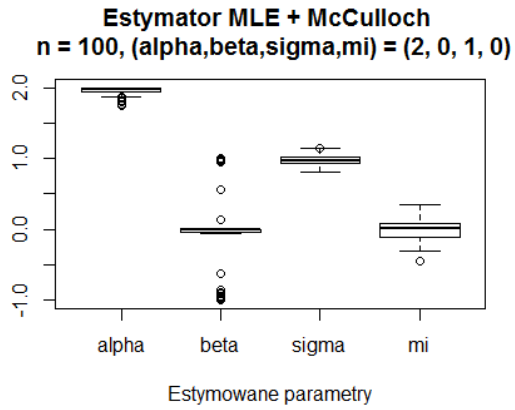
MLE <- function(n,pars){
  rnd <- stable_rnd(n,pars)
  pars_est_M <- stable_fit_init(rnd)
  pars_est_ML <- stable_fit_mle(rnd, pars_est_M)
  return(pars_est_ML)
}

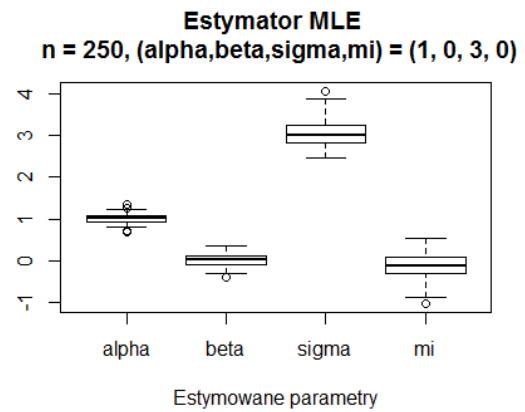
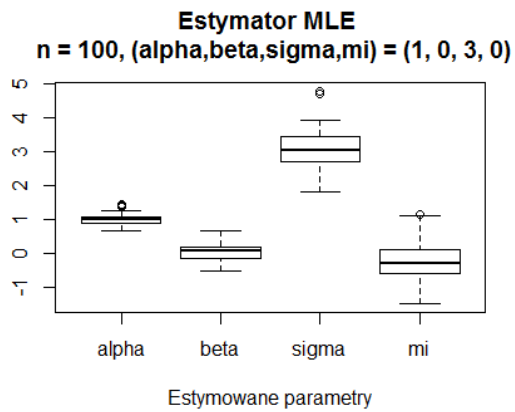
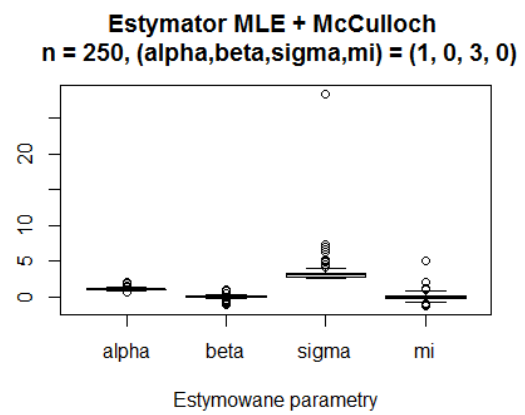
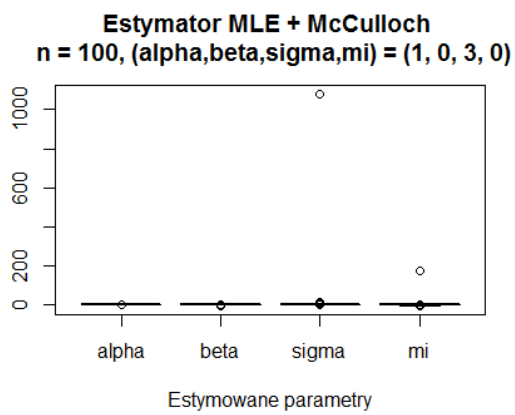
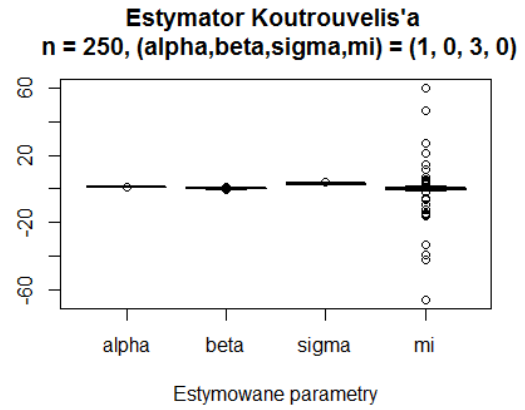
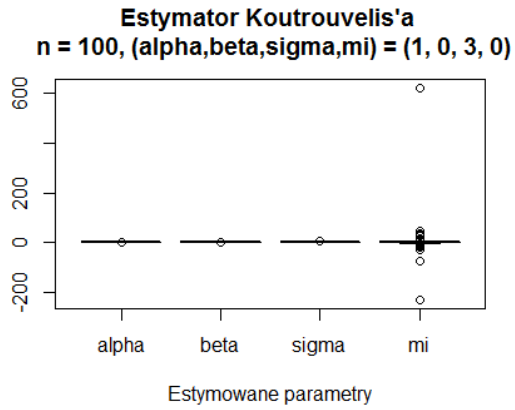
modified_MLE <- function(n,pars){
  rnd <- stable_rnd(n,pars)
  pars_est_M <- stable_fit_init(rnd)
  pars_est_ML2 <- stable_fit_mle2d(rnd, pars_est_M)
  return(pars_est_ML2)
}
```

Przeprowadzimy symulację ze 100 powtórzeniami, aby dokładniej oszacować parametry rozkładów stabilnych. Wysymulowane parametry zostały przedstawione na boxplotach z podziałem na rozmiar próby i metodę estymacji.









Dla danego rozmiaru próby  $n$  i parametrów  $(\alpha, \beta, \sigma, \mu)$  zaimplementowana funkcja *estimated\_pars* zwraca listę oszacowanych parametrów dla każdej z metod. Jako estymator przyjmujemy medianę zamiast średniej, ponieważ z boxplotów widzimy, że w niektórych metodach estymacji dostajemy spore wartości odstające. Wzięcie średniej zaburzyłoby wyniki w kilku przypadkach.

```
estimated_pars <- function(n,pars){
  method1 <- data.frame()
  method2 <- data.frame()
  method3 <- data.frame()
  method4 <- data.frame()
  for (i in 1:100){
    set.seed(i)
    method1 <- rbind.data.frame(method1, McCulloch(n,pars))
    method2 <- rbind.data.frame(method2, Koutrouvelis(n,pars))
    method3 <- rbind.data.frame(method3, MLE(n,pars))
    method4 <- rbind.data.frame(method4, modified_MLE(n,pars))
  }
  estimated_pars_method1 <- c(median(method1[,1]),median(method1[,2]),
                             median(method1[,3]),median(method1[,4]))
  estimated_pars_method2 <- c(median(method2[,1]),median(method2[,2]),
                             median(method2[,3]),median(method2[,4]))
  estimated_pars_method3 <- c(median(method3[,1]),median(method3[,2]),
                             median(method3[,3]),median(method3[,4]))
  estimated_pars_method4 <- c(median(method4[,1]),median(method4[,2]),
                             median(method4[,3]),median(method4[,4]))
  return(list(estimated_pars_method1,estimated_pars_method2,
             estimated_pars_method3,estimated_pars_method4))
}
```

W poniższych tabelach przedstawiono oszacowane parametry.

Tabela 1: Estymowane parametry dla  $(\alpha, \beta, \sigma, \mu) = (1, 0, 1, 0)$  i  $n = 100$

estymator	$\alpha$	$\beta$	$\sigma$	$\mu$
McCulloch'a	0.97572948	0.04219561	0.96805905	-0.01504473
Koutrouvelis'a	0.94558501	0.00000000	0.96245410	-0.03882562
MLE	1.00000000	-0.01174912	1.02612218	-0.01540433
zmodyfikowany MLE	0.9759686	0.0556235	1.0129781	-0.1025912

Tabela 2: Estymowane parametry dla  $(\alpha, \beta, \sigma, \mu) = (1, 0, 1, 0)$  i  $n = 250$

estymator	$\alpha$	$\beta$	$\sigma$	$\mu$
McCulloch'a	1.0084711883	0.0235014681	0.9700422129	0.0008645172
Koutrouvelis'a	0.990066740	0.0000000000	0.983573308	-0.003533669
MLE	1.01479013	-0.02813582	1.01153010	0.00127482
zmodyfikowany MLE	1.00550432	0.01486948	1.00052274	-0.04262371

Tabela 3: Estymowane parametry dla  $(\alpha, \beta, \sigma, \mu) = (2, 0, 1, 0)$  i  $n = 100$

estymator	$\alpha$	$\beta$	$\sigma$	$\mu$
McCulloch'a	1.948431101	0.000000000	0.981521401	-0.005855395
Koutrouvelis'a	2.000000000	0.000000000	0.99049880	-0.00764978
MLE	1.984432073	0.000000000	0.971382826	0.003481256
zmodyfikowany MLE	1.95766752	0.02192405	0.98157039	-0.02998675

Tabela 4: Estymowane parametry dla  $(\alpha, \beta, \sigma, \mu) = (2, 0, 1, 0)$  i  $n = 250$

estymator	$\alpha$	$\beta$	$\sigma$	$\mu$
McCulloch'a	1.950000000	0.000000000	0.992901054	-0.008325884
Koutrouvelis'a	2.000000000	0.000000000	0.997978952	-0.001581204
MLE	2.000000000	0.000000000	0.98658520	0.01678597
zmodyfikowany MLE	1.96201399	0.02192405	0.99346613	-0.02228577



Tabela 5: Estymowane parametry dla  $(\alpha, \beta, \sigma, \mu) = (1, 0, 3, 0)$  i  $n = 100$ 

estymator	$\alpha$	$\beta$	$\sigma$	$\mu$
McCulloch'a	0.97572948	0.04219561	2.90417715	-0.04513419
Koutrouvelis'a	0.9436065	0.0000000	2.8913129	-0.1164769
MLE	1.027908775	0.002973643	3.108505359	-0.062910103
zmodyfikowany MLE	0.9759686	0.0556235	3.0389343	-0.3077737

Tabela 6: Estymowane parametry dla  $(\alpha, \beta, \sigma, \mu) = (1, 0, 3, 0)$  i  $n = 250$ 

estymator	$\alpha$	$\beta$	$\sigma$	$\mu$
McCulloch'a	1.008471188	0.023501468	2.910126639	0.002593552
Koutrouvelis'a	0.99006674	0.00000000	2.95398849	-0.01060101
MLE	1.02946950	-0.01962468	3.08489177	0.01320759
zmodyfikowany MLE	1.00550432	0.01486948	3.00156823	-0.12787112

Aby porównać wszystkie 4 metody, obliczymy błędy średniokwadratowe dla zestawów parametrów  $(\alpha, \beta, \sigma, \mu)$ , a nie każdego parametru osobno. W ten sposób sprawdzimy, która z metod daje najlepsze oszacowanie, a także czy zwiększenie rozmiaru próby ze 100 do 250 ma sens. Skorzystamy z zaimplementowanej funkcji.

```
MSE_4parametry <- function(n,pars,method){
  MSE <- mean((pars - c(estimated_pars(n,pars)[[method]]))^2)
  return(MSE)
}
```

Tabela 7: Błędy średniokwadratowe dla  $(\alpha, \beta, \sigma, \mu) = (1, 0, 1, 0)$ 

estymator	$n = 100$	$n = 250$
McCulloch'a	0.0009040240	0.0003805741
Koutrouvelis'a	0.0014695290	0.0000952482
MLE	0.0002644259	0.0002862352
zmodyfikowany MLE	0.0035912180	0.0005171131

Tabela 8: Błędy średniokwadratowe dla  $(\alpha, \beta, \sigma, \mu) = (2, 0, 1, 0)$

estymator	$n = 100$	$n = 250$
McCulloch'a	0.0007587739	0.0006549288
Koutrouvelis'a	0.0000371980	0.0000016462
MLE	0.0002683555	0.0001154314
zmodyfikowany MLE	0.0008778898	0.0006157370

Tabela 9: Błędy średniokwadratowe dla  $(\alpha, \beta, \sigma, \mu) = (1, 0, 3, 0)$

estymator	$n = 100$	$n = 250$
McCulloch'a	0.0033971600	0.0021770070
Koutrouvelis'a	0.0071399920	0.0005820276
MLE	0.0041297090	0.0021586580
zmodyfikowany MLE	0.0249780000	0.0041512200

Możemy zauważyć, że zwiększenie rozmiaru próby daje lepsze oszacowanie z wyjątkiem parametrów  $(\alpha, \beta, \sigma, \mu) = (1, 0, 1, 0)$  i estymatora MLE, dla którego otrzymujemy nieco większy błąd średniokwadratowy, gdy próba zawiera więcej danych. W przypadku  $n = 100$  nie możemy jednoznacznie stwierdzić, która metoda jest najlepsza, ponieważ dla każdego zestawu parametrów najmniejszy błąd średniokwadratowy otrzymujemy dla innej metody. Natomiast gdy zwiększymy rozmiar próby do  $n = 250$  od razu widzimy, że estymator Koutrouvelis'a wypada najlepiej. Zatem zwiększenie rozmiaru próby było konieczne w celu wyłonienia najlepszej metody.

## Część 2

W tej części przeprowadzimy symulacje, które mają na celu porównanie metod parametrycznych i nieparametrycznych szacowania kwantyli rozkładu, z którego pochodzą dane. Będziemy generować dane z rozkładów stabilnych z parametrami  $(\alpha, \beta, \sigma, \mu) = (1, 0, 1, 0)$ ,  $(\alpha, \beta, \sigma, \mu) = (2, 0, 1, 0)$  i  $(\alpha, \beta, \sigma, \mu) = (1, 0, 3, 0)$ . Przyjmujemy rozmiar próby  $n = 100$  i  $n = 250$  oraz rzędy kwantyli 0.95 i 0.99. Dla przypadku parametrycznego rozważamy 4 metody, tj. McCulloch'a, Koutrouvelis'a, MLE oraz zmodyfikowany MLE. Natomiast w przypadku nieparametrycznym przyjmujemy typ kwantyla 1, 6 i 8.

Poniżej przedstawiono funkcję szacującą kwantyle dla każdej z 7 metod dla zadanego rzędu kwantyla, rozmiaru próby  $n$  oraz parametrów  $(\alpha, \beta, \sigma, \mu)$ . Tak samo jak w 1 części sprawozdania jako szacowaną wartość kwantyla przyjmujemy medianę spośród 100 wyestymowanych wartości.

```
estimated_quantiles <- function(kwantyl,n,pars){
  method1 <- data.frame()
  method2 <- data.frame()
  method3 <- data.frame()
  method4 <- data.frame()
  method5 <- data.frame()
  method6 <- data.frame()
  method7 <- data.frame()
  for (i in 1:100){
    set.seed(i)
    rnd <- stable_rnd(n,pars)
    method1 <- rbind.data.frame(method1, stable_q(kwantyl,
                                                    McCulloch(n,pars)))
    method2 <- rbind.data.frame(method2, stable_q(kwantyl,
                                                    Koutrouvelis(n,pars)))
    method3 <- rbind.data.frame(method3, stable_q(kwantyl,MLE(n,pars)))
    method4 <- rbind.data.frame(method4, stable_q(kwantyl,
                                                    modified_MLE(n,pars)))
    method5 <- rbind.data.frame(method5, quantile(rnd,kwantyl,type=1))
    method6 <- rbind.data.frame(method6, quantile(rnd,kwantyl,type=6))
    method7 <- rbind.data.frame(method7, quantile(rnd,kwantyl,type=8))
  }
  estimated_quantile_method1 <- c(median(method1[,1]))
  estimated_quantile_method2 <- c(median(method2[,1]))
  estimated_quantile_method3 <- c(median(method3[,1]))
  estimated_quantile_method4 <- c(median(method4[,1]))
  estimated_quantile_method5 <- c(median(method5[,1]))
  estimated_quantile_method6 <- c(median(method6[,1]))
  estimated_quantile_method7 <- c(median(method7[,1]))
  return(list(estimated_quantile_method1,estimated_quantile_method2,
              estimated_quantile_method3,estimated_quantile_method4,
              estimated_quantile_method5,estimated_quantile_method6,
              estimated_quantile_method7))
}
```

Teraz zaprezentujemy wyniki uzyskane po przeprowadzeniu 100 symulacji. W poniższych tabelach jako  $q_{0.95}$  oznaczamy kwantyl teoretyczny rzędu 0.95 dla odpowiednich parametrów  $(\alpha, \beta, \sigma, \mu)$ . Analogicznie jako  $q_{0.99}$  oznaczamy kwantyl teoretyczny rzędu 0.99.

Tabela 10: Estymowane kwantyle dla  $(\alpha, \beta, \sigma, \mu) = (1, 0, 1, 0)$  i rzędu kwantyla 0.95

metoda	$q_{0.95} = 6.313752$	
	n = 100	n = 250
McCulloch'a	6.298181	6.210598
Koutrouvelis'a	7.131925	6.579636
MLE	5.967933	6.087376
zmodyfikowany MLE	6.457683	6.318334
funkcja <i>quantile</i> , typ = 1	5.868180	6.252249
funkcja <i>quantile</i> , typ = 6	6.755325	6.491395
funkcja <i>quantile</i> , typ = 8	6.502285	6.328709

Tabela 11: Estymowane kwantyle dla  $(\alpha, \beta, \sigma, \mu) = (1, 0, 1, 0)$  i rzędu kwantyla 0.99

metoda	$q_{0.99} = 31.82052$	
	n = 100	n = 250
McCulloch'a	34.01206	31.92091
Koutrouvelis'a	36.84321	32.39510
MLE	28.75545	30.11599
zmodyfikowany MLE	34.71994	32.28473
funkcja <i>quantile</i> , typ = 1	19.00253	31.18943
funkcja <i>quantile</i> , typ = 6	40.94991	40.84192
funkcja <i>quantile</i> , typ = 8	33.81386	34.98203

Tabela 12: Estymowane kwantyle dla  $(\alpha, \beta, \sigma, \mu) = (2, 0, 1, 0)$  i rzędu kwantyla 0.95

metoda	$q_{0.95} = 2.326174$	
	n = 100	n = 250
McCulloch'a	2.455348	2.389219
Koutrouvelis'a	2.311871	2.316073
MLE	2.273009	2.305292
zmodyfikowany MLE	2.443775	2.378271
funkcja <i>quantile</i> , typ = 1	2.236318	2.292318
funkcja <i>quantile</i> , typ = 6	2.362347	2.325463
funkcja <i>quantile</i> , typ = 8	2.322350	2.309314

Tabela 13: Estymowane kwantyle dla  $(\alpha, \beta, \sigma, \mu) = (2, 0, 1, 0)$  i rzędu kwantyla 0.99

metoda	$q_{0.99} = 3.289953$	
	n = 100	n = 250
McCulloch'a	3.745223	3.585837
Koutrouvelis'a	3.295148	3.288846
MLE	3.261190	3.285654
zmodyfikowany MLE	3.731159	3.551832
funkcja <i>quantile</i> , typ = 1	2.984206	3.237051
funkcja <i>quantile</i> , typ = 6	3.362337	3.342588
funkcja <i>quantile</i> , typ = 8	3.284247	3.289523

Tabela 14: Estymowane kwantyle dla  $(\alpha, \beta, \sigma, \mu) = (1, 0, 3, 0)$  i rzędu kwantyla 0.95

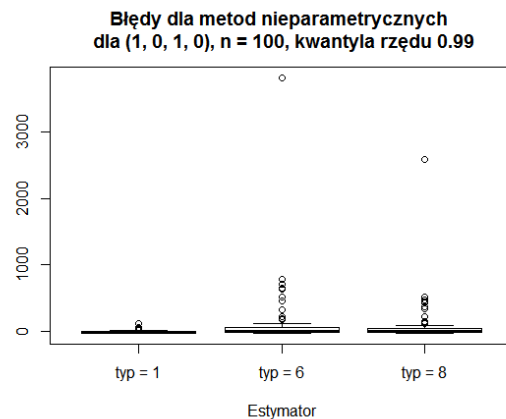
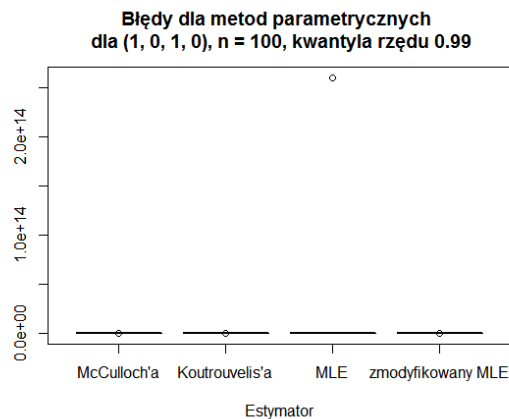
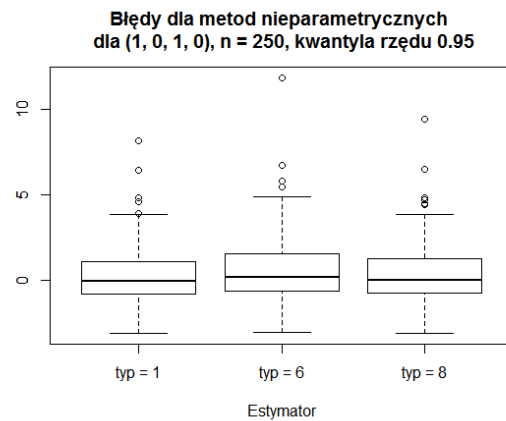
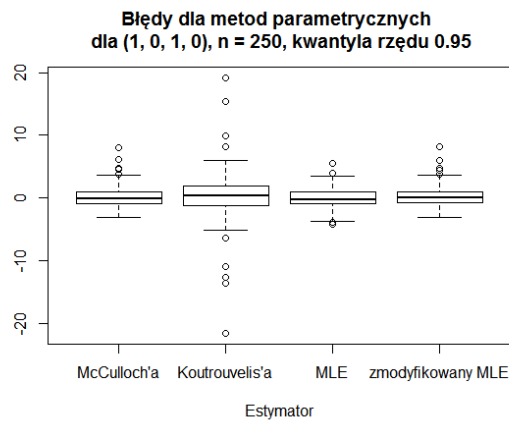
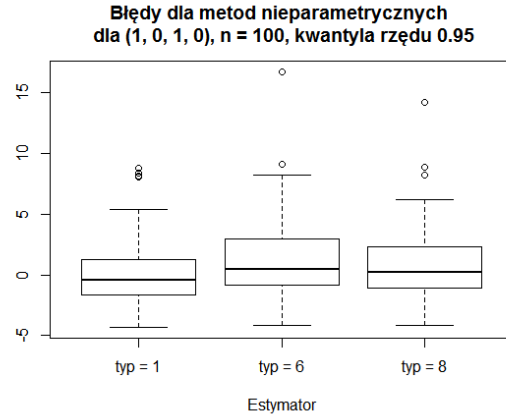
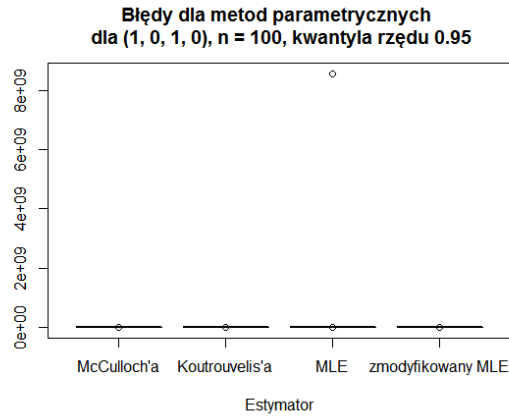
metoda	$q_{0.95} = 18.94125$	
	n = 100	n = 250
McCulloch'a	18.89454	18.63179
Koutrouvelis'a	21.67852	19.89624
MLE	18.83473	18.62748
zmodyfikowany MLE	19.37305	18.95500
funkcja <i>quantile</i> , typ = 1	17.60454	18.75675
funkcja <i>quantile</i> , typ = 6	20.26597	19.47419
funkcja <i>quantile</i> , typ = 8	19.50686	18.98613

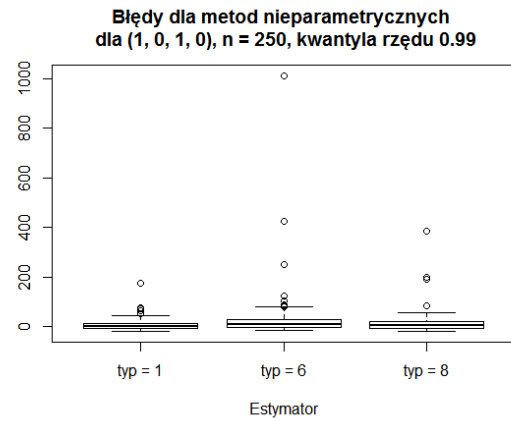
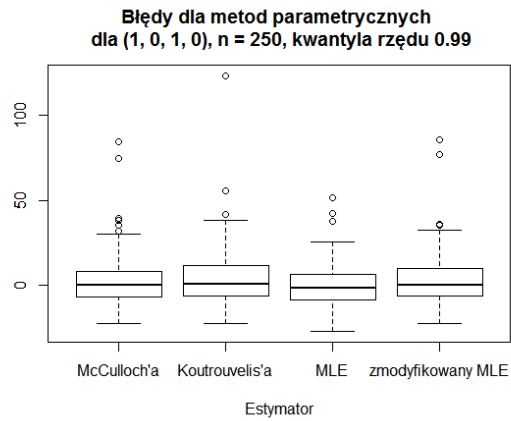
Tabela 15: Estymowane kwantyle dla  $(\alpha, \beta, \sigma, \mu) = (1, 0, 3, 0)$  i rzędu kwantyla 0.99

metoda	$q_{0.99} = 95.46155$	
	n = 100	n = 250
McCulloch'a	102.0362	95.76273
Koutrouvelis'a	111.2935	97.20296
MLE	89.36854	87.44290
zmodyfikowany MLE	104.1598	96.85419
funkcja <i>quantile</i> , typ = 1	57.00758	93.56828
funkcja <i>quantile</i> , typ = 6	122.8497	122.5258
funkcja <i>quantile</i> , typ = 8	101.4416	104.9461

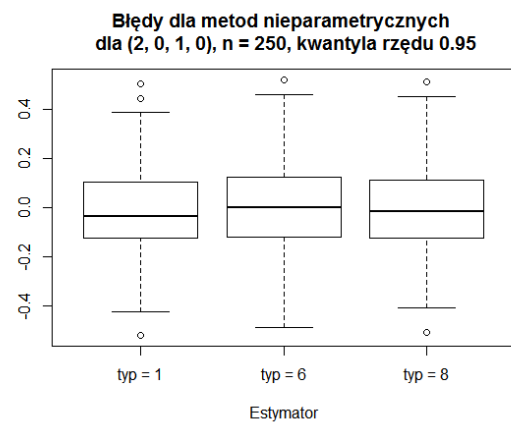
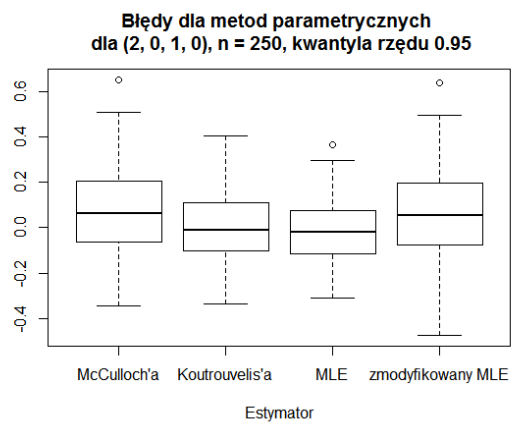
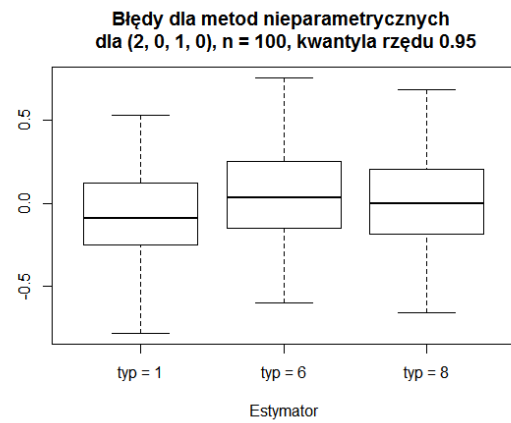
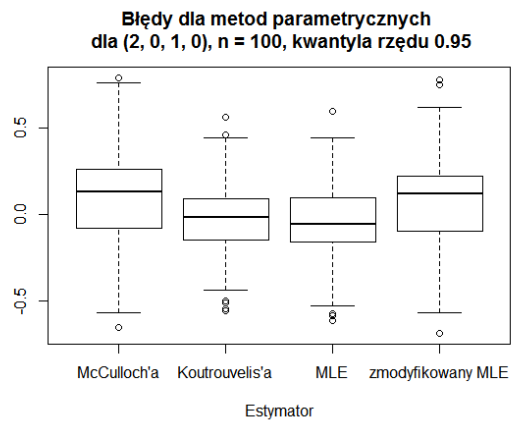
Na poniższych boxplotach przedstawiono błędy dla metod parametrycznych i nieparametrycznych dla zadanych parametrów  $(\alpha, \beta, \sigma, \mu)$ , rozmiaru próby  $n$  i rzędu kwantyla. Błąd jest szacowany jako różnica pomiędzy 100 wyestymowanymi kwantylami a kwantylem teoretycznym.

Boxploty dla  $(\alpha, \beta, \sigma, \mu) = (1, 0, 1, 0)$ :

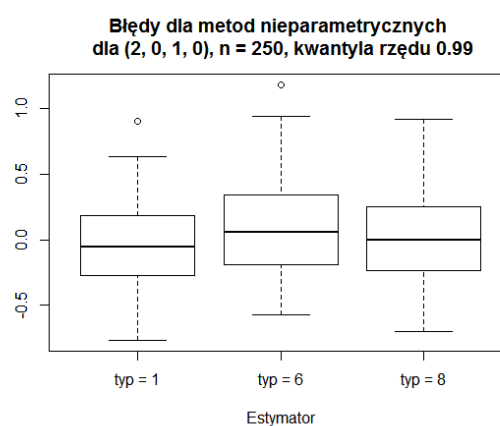
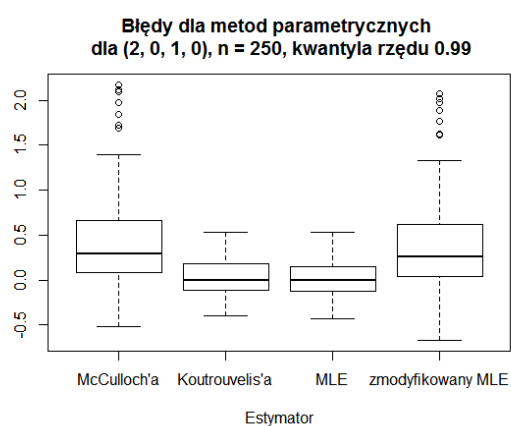
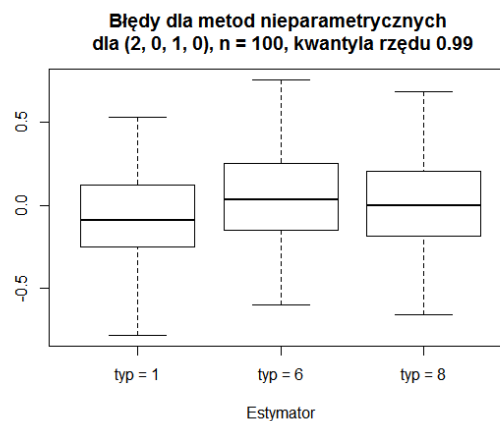
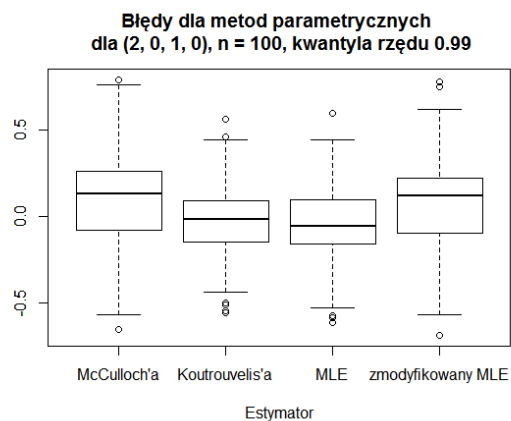




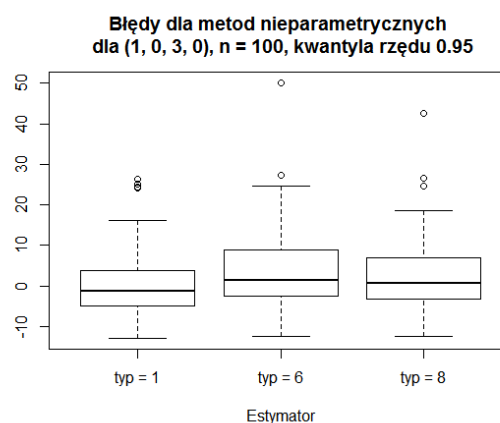
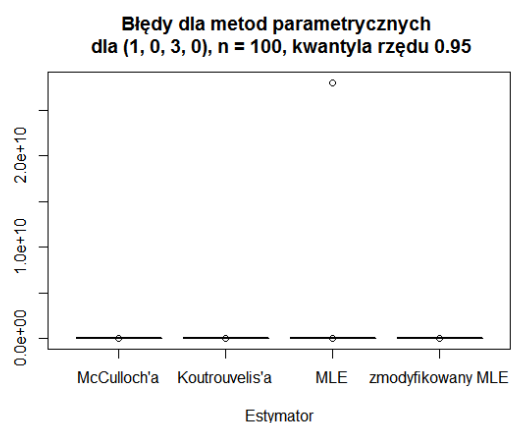
Boxploty dla  $(\alpha, \beta, \sigma, \mu) = (2, 0, 1, 0)$ :

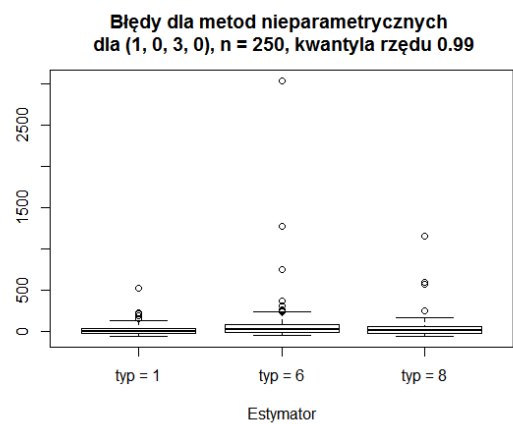
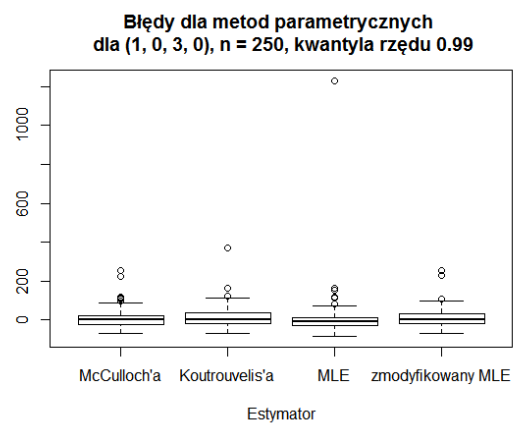
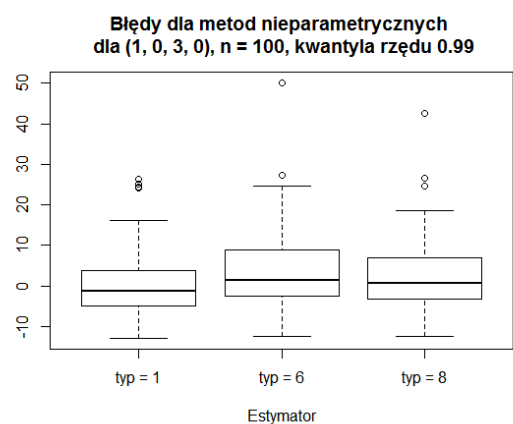
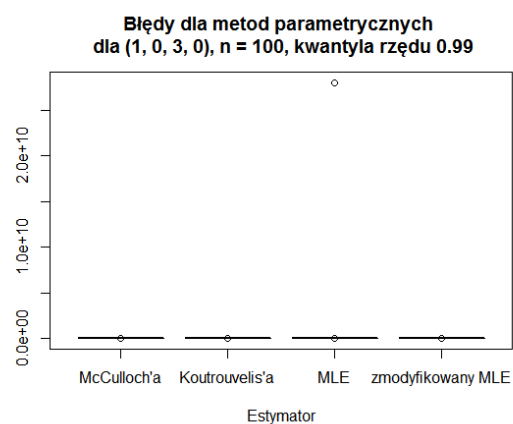
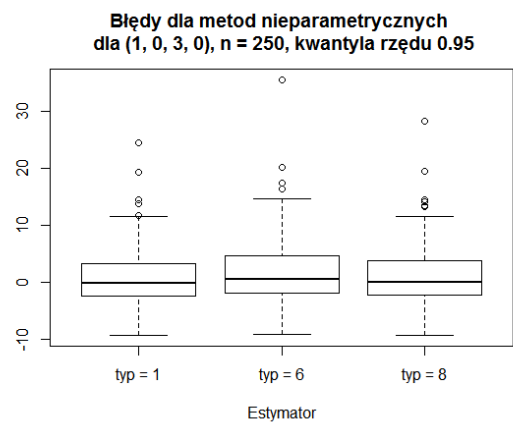
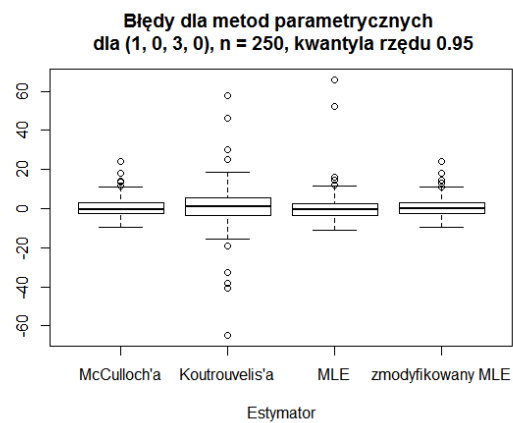






Boxploty dla  $(\alpha, \beta, \sigma, \mu) = (1, 0, 3, 0)$ :





W celu porównania wszystkich 7 metod obliczymy błędy średniokwadratowe. Skorzystamy z funkcji *mse* z pakietu **mltools**. W tabelach przedstawiono błędy średniokwadratowe zależnie od parametrów  $(\alpha, \beta, \sigma, \mu)$  i rzędu kwantyla z podziałem na rozmiar próby  $n$  i metodę szacowania.

Tabela 16: Błędy średniokwadratowe dla  $(\alpha, \beta, \sigma, \mu) = (1, 0, 1, 0)$  i rzędu kwantyla 0.95

metoda	rozmiar próby	
	n = 100	n = 250
McCulloch'a	0.00024245	0.01064071
Koutrouvelis'a	0.66940720	0.07069448
MLE	0.11959070	0.05124607
zmodyfikowany MLE	0.02071638	0.00002099
funkcja <i>quantile</i> , typ = 1	0.19853430	0.00378258
funkcja <i>quantile</i> , typ = 6	0.19498710	0.03155732
funkcja <i>quantile</i> , typ = 8	0.03554499	0.00022372

Tabela 17: Błędy średniokwadratowe dla  $(\alpha, \beta, \sigma, \mu) = (1, 0, 1, 0)$  i rzędu kwantyla 0.99

metoda	rozmiar próby	
	n = 100	n = 250
McCulloch'a	4.802859	0.0100791
Koutrouvelis'a	25.22745	0.3301442
MLE	9.394624	2.9054060
zmodyfikowany MLE	8.406635	0.2154960
funkcja <i>quantile</i> , typ = 1	164.3009	0.3982732
funkcja <i>quantile</i> , typ = 6	83.34579	81.385780
funkcja <i>quantile</i> , typ = 8	3.973418	9.9951690

Tabela 18: Błędy średniokwadratowe dla  $(\alpha, \beta, \sigma, \mu) = (2, 0, 1, 0)$  i rzędu kwantyla 0.95

metoda	rozmiar próby	
	n = 100	n = 250
McCulloch'a	0.016685950	0.003974607
Koutrouvelis'a	0.000204577	0.000102045
MLE	0.002826586	0.000436092
zmodyfikowany MLE	0.013829930	0.002714082
funkcja <i>quantile</i> , typ = 1	0.008074072	0.001146247
funkcja <i>quantile</i> , typ = 6	0.001308461	0.000000507
funkcja <i>quantile</i> , typ = 8	0.000014629	0.000284283

Tabela 19: Błędy średniokwadratowe dla  $(\alpha, \beta, \sigma, \mu) = (2, 0, 1, 0)$  i rzędu kwantyla 0.99

metoda	rozmiar próby	
	n = 100	n = 250
McCulloch'a	0.207270600	0.087547470
Koutrouvelis'a	0.000026993	0.000001225
MLE	0.000827287	0.000018482
zmodyfikowany MLE	0.194663200	0.068581010
funkcja <i>quantile</i> , typ = 1	0.093480980	0.002798643
funkcja <i>quantile</i> , typ = 6	0.005239448	0.002770422
funkcja <i>quantile</i> , typ = 8	0.000032551	0.000000185

Tabela 20: Błędy średniokwadratowe dla  $(\alpha, \beta, \sigma, \mu) = (1, 0, 3, 0)$  i rzędu kwantyla 0.95

metoda	rozmiar próby	
	n = 100	n = 250
McCulloch'a	0.00218208	0.09576641
Koutrouvelis'a	7.49262900	0.91199380
MLE	0.01134690	0.09845255
zmodyfikowany MLE	0.18644740	0.00018896
funkcja <i>quantile</i> , typ = 1	1.78680900	0.03404323
funkcja <i>quantile</i> , typ = 6	1.75488400	0.28401590
funkcja <i>quantile</i> , typ = 8	0.31990490	0.00201348

Tabela 21: Błędy średniokwadratowe dla  $(\alpha, \beta, \sigma, \mu) = (1, 0, 3, 0)$  i rzędu kwantyla 0.99

metoda	rozmiar próby	
	n = 100	n = 250
McCulloch'a	43.22573	0.090712
Koutrouvelis'a	250.6513	3.032517
MLE	37.12479	64.29872
zmodyfikowany MLE	75.65971	1.939464
funkcja <i>quantile</i> , typ = 1	1478.708	3.584459
funkcja <i>quantile</i> , typ = 6	750.1121	732.4720
funkcja <i>quantile</i> , typ = 8	35.76077	89.95652

Dla parametrów  $(\alpha, \beta, \sigma, \mu) = (1, 0, 1, 0)$  i kwantyla rzędu 0.95 zwiększenie rozmiaru próby polepsza wyniki oprócz metody McCulloch'a. Dla tych parametrów najlepsze oszacowanie daje zmodyfikowany MLE. W przypadku kwantyla rzędu 0.99 otrzymujemy nieco gorsze wyniki. Zwiększenie rozmiaru próby skutkuje polepszeniem oszacowań z wyjątkiem metody nieparametrycznej z typem kwantyla 8. Najlepszy wynik uzyskujemy dla metody McCulloch'a.

Biorąc pod uwagę parametry  $(\alpha, \beta, \sigma, \mu) = (2, 0, 1, 0)$  i kwantyl rzędu 0.95 możemy zauważyć, że zwiększenie rozmiaru próby skutkuje lepszymi oszacowaniami z wyjątkiem metody nieparametrycznej z typem kwantyla 8. Najlepsze rezultaty uzyskujemy w przypadku 6 metody, czyli nieparametrycznej z typem kwantyla 6. Dla tych samych parametrów  $(\alpha, \beta, \sigma, \mu)$  i kwantyla rzędu 0.99 dostajemy lepsze oszacowania dla większej próby dla wszystkich metod. Najlepszy wynik uzyskujemy stosując metodę nieparametryczną z typem kwantyla 6.

Z kolei analizując parametry  $(\alpha, \beta, \sigma, \mu) = (1, 0, 3, 0)$  i kwantyl rzędu 0.95 widzimy, że zwiększenie rozmiaru próby polepsza oszacowania z wyjątkiem metody McCulloch'a i MLE. Najlepszą metodą okazuje się zmodyfikowany MLE, lecz także metoda nieparametryczna z typem kwantyla 8 daje dobre wyniki. Najgorsze oszacowania spośród wszystkich dostajemy dla parametrów  $(\alpha, \beta, \sigma, \mu) = (1, 0, 3, 0)$  i kwantyla rzędu 0.99. Z boxplotów widzimy, że dla tych parametrów dostajemy spore wartości odstające zarówno dla metod parametrycznych jak i nieparametrycznych. Lepsze szacowania dostajemy dla większej próby z wyjątkiem metody MLE. Najlepiej w tym zestawieniu wypada metoda McCulloch'a.

Ostatecznie możemy stwierdzić, że dla parametrów  $(\alpha, \beta, \sigma, \mu) = (1, 0, 1, 0)$  oraz  $(\alpha, \beta, \sigma, \mu) = (1, 0, 3, 0)$  lepsze są metody parametryczne. Dla kwantyla rzędu 0.95 najlepsze oszacowania uzyskujemy stosując zmodyfikowany MLE, a dla kwantyla 0.99 metoda McCulloch'a wypada najlepiej. Dla parametrów  $(\alpha, \beta, \sigma, \mu) = (2, 0, 1, 0)$  lepsze są metody nieparametryczne w przypadku dużej liczby danych.

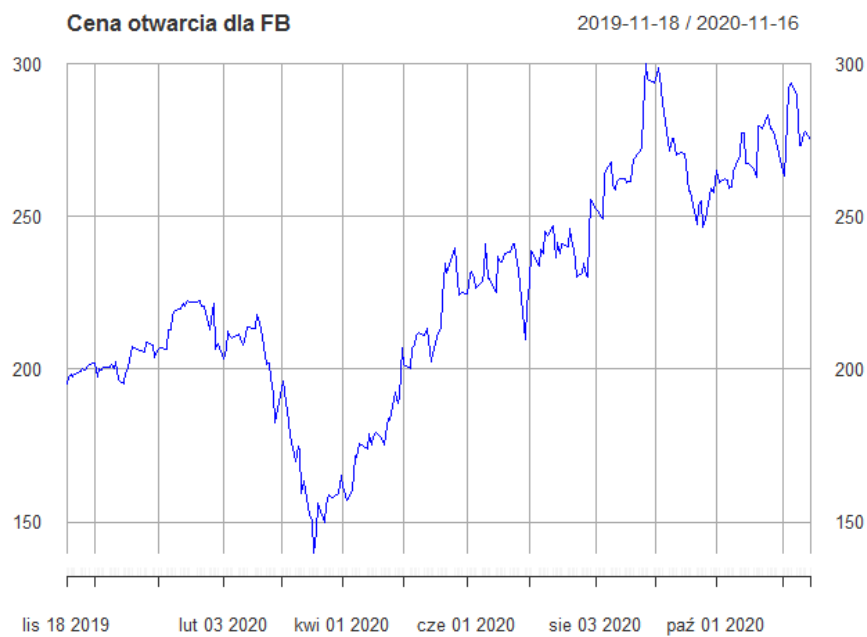
## Część 3

Za pomocą funkcji *getSymbols* z pakietu **quantmod** wczytujemy dane dotyczące akcji dwóch spółek za ostatni rok. Na ich podstawie obliczymy dzienne logarytmiczne stopy zwrotu. Będziemy rozważać akcje spółek *Facebook* oraz *Twitter*, oznaczane jako **FB**, **TWTR**.

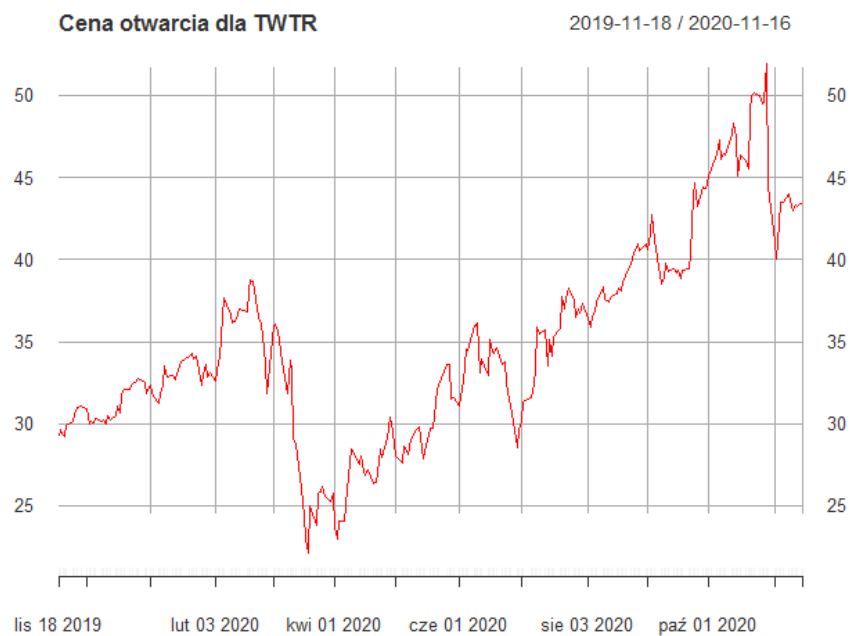
```
require("quantmod")

stock_list = c("FB", "TWTR")
start_date = Sys.Date() - 365
end_date = Sys.Date()
master_df = NULL
for(idx in seq(length(stock_list))){
  stock_index = stock_list[idx]
  getSymbols(stock_index, verbose = TRUE, src = "yahoo",
             from=start_date, to=end_date)
  temp_df = as.data.frame(get(stock_index))
  temp_df$Date = row.names(temp_df)
  temp_df$Index = stock_index
  row.names(temp_df) = NULL
  colnames(temp_df) = c("Open", "High", "Low", "Close",
                        "Volume", "Adjusted", "Date", "Index")
  temp_df = temp_df[c("Date", "Index", "Open", "High",
                      "Low", "Close", "Volume", "Adjusted")]
  master_df = rbind(master_df, temp_df)
}
```

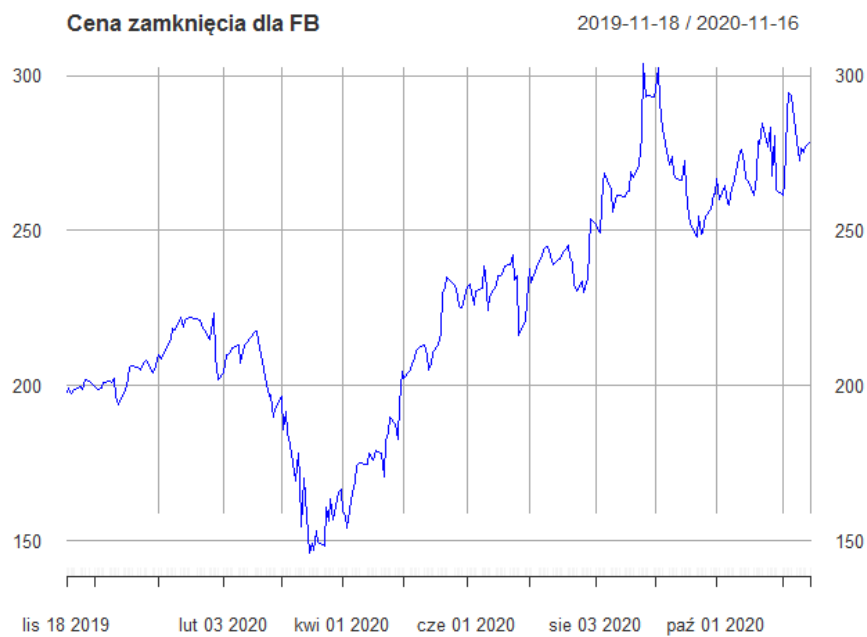
Wczytane dane zawierają informacje m.in. o cenach otwarcia, cenach zamknięcia, najniższej i najwyższej cenie notowanej w ciągu dnia. Teraz na kilku wykresach przedstawimy wizualizację pobranych danych.



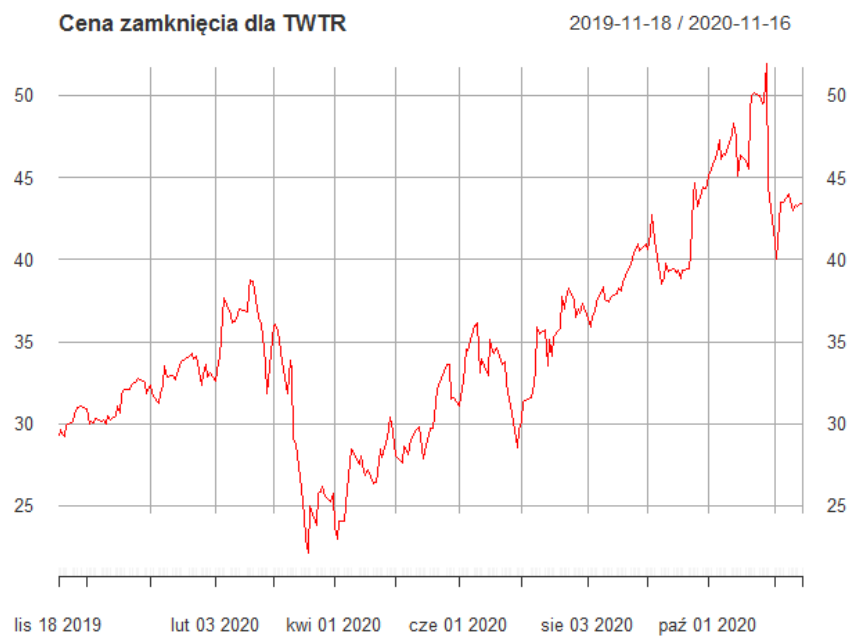
Rysunek 25: Cena otwarcia dla firmy Facebook



Rysunek 26: Cena otwarcia dla firmy Twitter

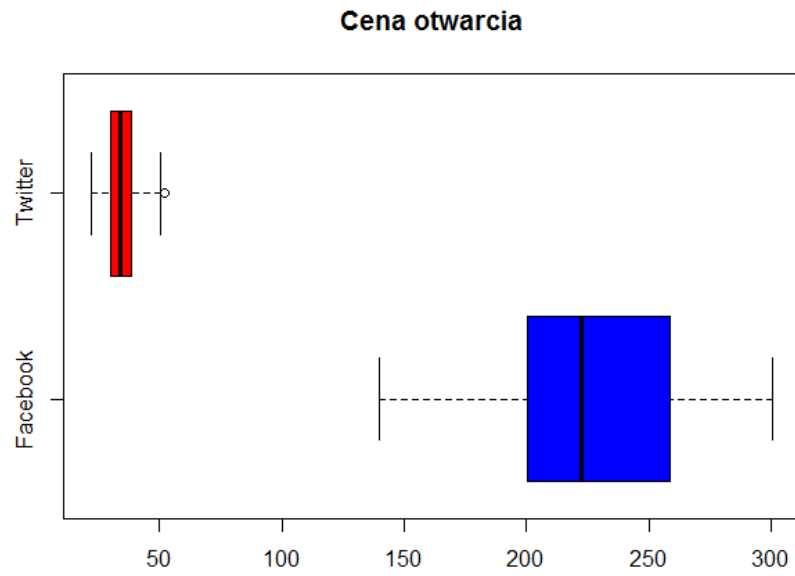


Rysunek 27: Cena zamknięcia dla firmy Facebook

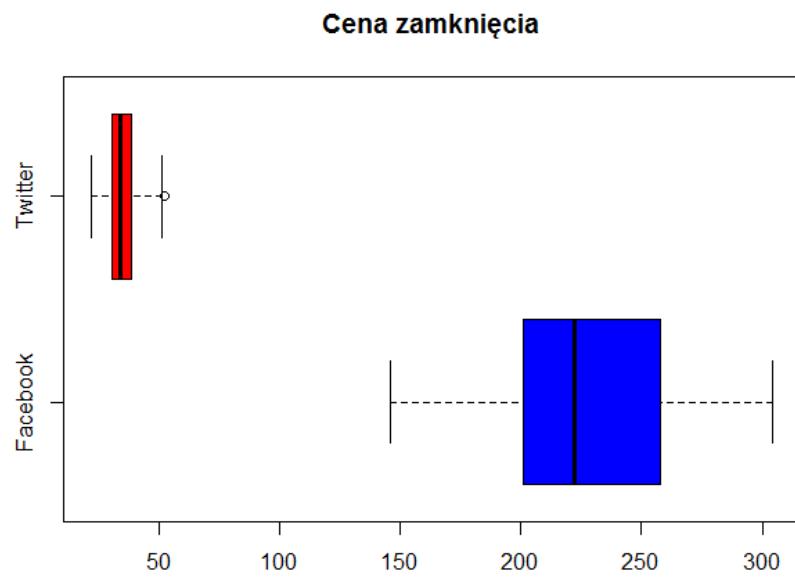


Rysunek 28: Cena zamknięcia dla firmy Twitter





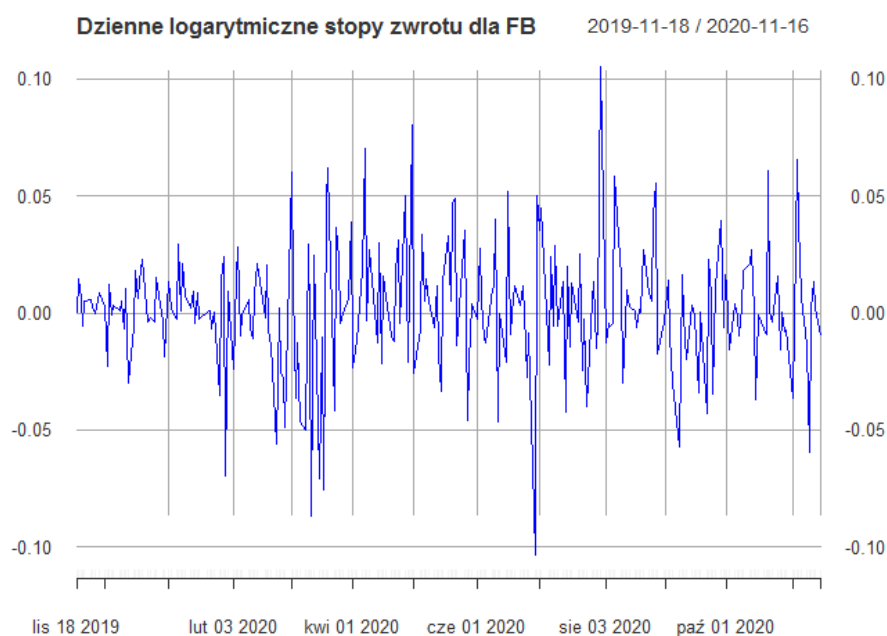
Rysunek 29: Cena otwarcia dla obu firm



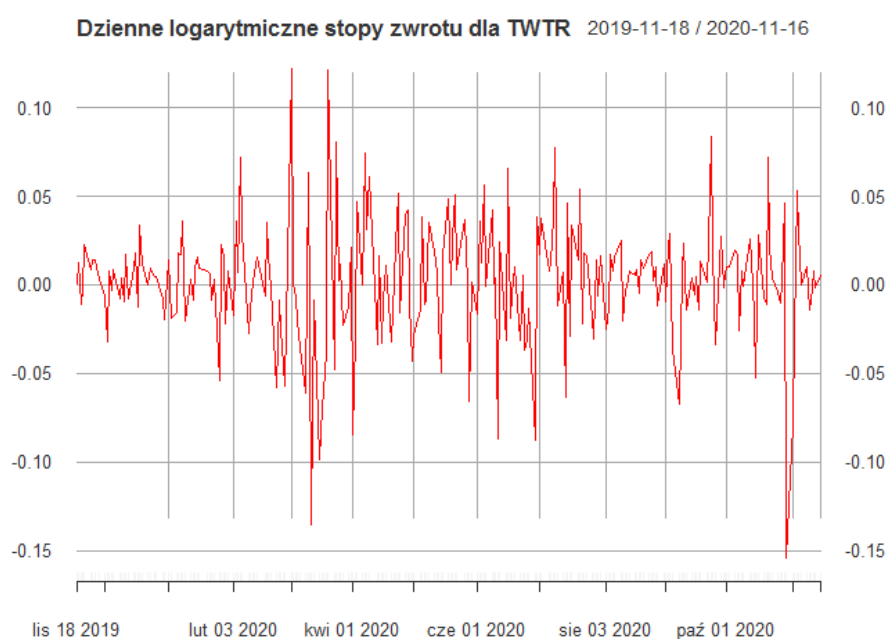
Rysunek 30: Cena zamknięcia dla obu firm

Dzienne logarytmiczne stopy zwrotu liczymy za pomocą funkcji *periodReturn*.

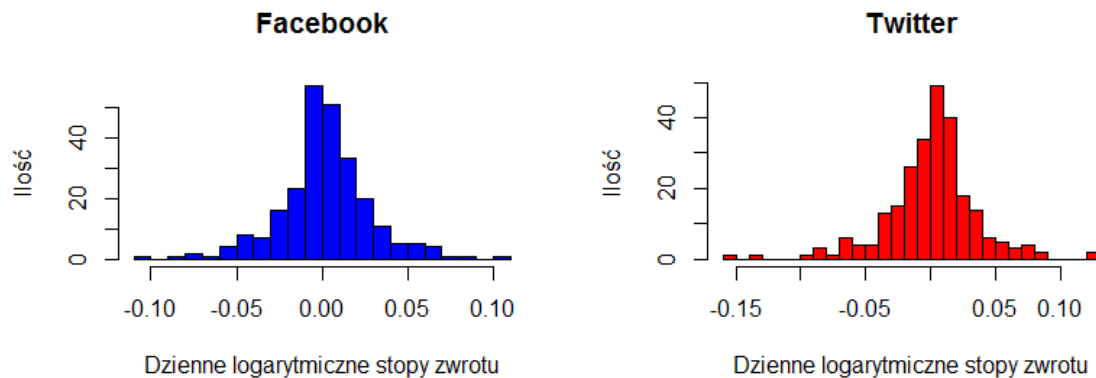
```
r.daily_FB <- periodReturn(FB$FB.Open,type = "log",period = "daily")  
r.daily_TWTR <- periodReturn(TWTR$TWTR.Open,type = "log",period = "daily")
```



Rysunek 31: Dzienne logarytmiczne stopy zwrotu dla firmy Facebook



Rysunek 32: Dzienne logarytmiczne stopy zwrotu dla firmy Twitter



Rysunek 33: Dzienne logarytmiczne stopy zwrotu dla obu firm

Przyjmując, że te zwroty logarytmiczne mają rozkład

1. normalny o nieznanym parametrach,
2. stabilny o nieznanym parametrach,
3. nieznanym absolutnie ciągły.

oszacujemy punktowo  $VaR_\alpha(L)$  oraz  $ES_\alpha(L)$ , gdzie  $L$  jest losową stratą związaną z jednodniową inwestycją  $S$  jednostek dla  $\alpha = 0.05$ . Chcemy przyjąć takie liczby jednostek dla obu spółek, aby uzyskać w przybliżeniu tę samą kwotę inwestycji. Jeśli jako  $S_1$  oznaczmy ilość kupionych akcji pierwszej spółki, to  $S_2$  wyznaczmy następująco:

```
shares <- function(S1){
  valueFB <- as.numeric(FB$FB.Open[length(FB$FB.Open)])
  valueTWTR <- as.numeric(TWTR$TWTR.Open[length(TWTR$TWTR.Open)])
  S2 <- round((S1*valueFB)/valueTWTR)
  return(S2)
}
S1 <- 10
S2 <- shares(S1)
```

W rezultacie, aby zainwestowana kwota w obie spółki była w przybliżeniu taka sama, przyjmujemy  $S_1 = 10$  oraz  $S_2 = 63$ .

Teraz obliczymy  $VaR_\alpha(L)$  oraz  $ES_\alpha(L)$  za pomocą zaimplementowanych funkcji.

```
VaR <- function(case,company,p){
  if(company=="FB"){
    P0 <- FB$FB.Open[length(FB$FB.Open)]
    a <- S1
    if(case==1){
      q <- qnorm(p,mean(r.daily_FB),sd(r.daily_FB))
    }
    if(case==2){
      pars_est_M <- stable_fit_init(r.daily_FB)
      q <- stable_q(p,pars_est_M)
    }
    if(case==3){
      pars_est_M <- stable_fit_init(r.daily_FB)
      q <- quantile(r.daily_FB,p,type=1)
    }
  }
  if(company=="TWTR"){
    P0 <- TWTR$TWTR.Open[length(TWTR$TWTR.Open)]
    a <- S2
    if(case==1){
      q <- qnorm(p,mean(r.daily_TWTR),sd(r.daily_TWTR))
    }
    if(case==2){
      pars_est_M <- stable_fit_init(r.daily_TWTR)
      q <- stable_q(p,pars_est_M)
    }
    if(case==3){
      pars_est_M <- stable_fit_init(r.daily_TWTR)
      q <- quantile(r.daily_TWTR,p,type=1)
    }
  }
  VaR <- a*as.numeric(P0)*(1-exp(q))
  VaR <- as.vector(VaR)
  return(VaR)
}
```

```

integrand <- function(case,company){
  value <- function(p){
    return(VaR(case,company,p))
  }
  return(value)
}

ES <- function(case,company,alpha){
  integral <- as.numeric(integrate(integrand(case,company),
                                   lower = 0,upper = alpha)[1])
  ES <- integral/alpha
  return(ES)
}

```

Tabela 22: Wartość narażona na ryzyko  $VaR_{0.05}(L)$

rozkład	Facebook	Twitter
normalny o nieznanych parametrach	116.2134	147.5386
stabilny o nieznanych parametrach	121.6231	159.5271
nieznany absolutnie ciągły	122.744	160.7428

Tabela 23: Oczekiwana strata  $ES_{0.05}(L)$

rozkład	Facebook	Twitter
normalny o nieznanych parametrach	145.7238	184.5504
stabilny o nieznanych parametrach	324.8475	398.6668
nieznany absolutnie ciągły	168.6859	225.7222

Wyznamy jeszcze realizacje przedziałów ufności na poziomie ufności 0.95 dla parametru  $VaR_{\alpha}(L)$ .

```

confidence_interval <- function(company){
  if (company=="FB"){
    a <- S1
    PO <- FB$FB.Open[length(FB$FB.Open)]
    r.daily <- periodReturn(FB$FB.Open,type = "log",period = "daily")
  }
  if (company=="TWTR"){
    a <- S2
    PO <- TWTR$TWTR.Open[length(TWTR$TWTR.Open)]
    r.daily <- periodReturn(TWTR$TWTR.Open,type = "log",period = "daily")
  }
  i = qbinom(0.05/2,length(r.daily),0.05)
  j = qbinom(1-0.05/2,length(r.daily),0.05) + 1
  L <- a*PO*(1-exp(sort(as.numeric(r.daily))[i]))
  L <- as.numeric(L)
}

```

```

U <- a*P0*(1-exp(sort(as.numeric(r.daily))[j]))
U <- as.numeric(U)
return(c(L,U))
}

```

Tabela 24: Przedziały ufności dla  $VaR_{0.05}(L)$

	L	U
Facebook	95.91866	159.12600
Twitter	115.7828	222.0878

$VaR_{0.05}(L)$  jest wartością narażoną na ryzyko (Value at Risk), której nie przekroczymy na 95%. Natomiast wartość oczekiwana straty (Expected Shortfall)  $ES_{0.05}(L)$  to strata którą poniesiemy, jeśli trafimy na pozostałe 5%. Możemy zauważyć, że obliczone wartości  $VaR_{0.05}(L)$  dla każdej ze spółek mieszczą się w odpowiednich przedziałach ufności. Po przeanalizowaniu wyników  $VaR_{0.05}(L)$  oraz  $ES_{0.05}(L)$  dla obu firm możemy stwierdzić, że inwestowanie w akcje TWTR jest narażone na większe ryzyko. Zarówno wartości Value at Risk oraz Expected Shortfall są większe dla każdego z rozważanych rozkładów w przypadku firmy Twitter.