

Kamil Szaryniak, Piotr Stawarski

2024-11-03

Hide

```
library("Tidyverse")
library("moment")
library("forecast")
library("rugarch")
library("xts")
```

Hide

Dane wykorzystywane w projekcie pochodzą z firmy "Izola" - Jarocin S.A., które pochodzą z okresu od września 2023 do 4 października 2024.

Hide

```
ito <- read.csv("C:/Visual_Studio/Ekonometria finansowa i dynamiczna/Projekt/Projekt_2/Izo_d.csv")
```

Hide

Stopa logaryczna

Hide

```
it1 <- function(a,b) {
  return(log(a/b))
}
```

Hide

Zaczynanie danych

Hide

```
date <- iso[1:151]
indexy <- date %>% str_extract("%Y-%m-%d")
```

Hide

Wykresanie stop logarycznych

Hide

```
df <- c()
for (i in 1:(nrow(indexy)-1)) {
  df <- c(df,as.numeric(it1(indexy[i,2],indexy[i+1,2])))
}
```

Hide

Podział na zbiór uczący i odczytane dane

Hide

```
stopy <- nrow(date)-41
oeny <- unname(head(log(date[21], -4))
stopy_0ct <- rev(head(df,4))
for(i in 1:(nrow(stopy)-1)) {
  oeny_n <- unname(head(log(date[21], 4))
  oeny_n <- unname(head(date[21, 4])
}
```

Hide

Test ADF

Hide

```
adf.test(stopy)
```

Hide

Warning in adf.test(stopy): p-value smaller than printed p-value

Hide

```
Augmented Dickey-Fuller Test:

data:      stopy
Dickey-Fuller = -6.8519, Lag order = 6, p-value = 0.01
alternative hypothesis: stationary
```

Hide

Wartość statystyki p-value -> szereg stacjonarny

Hide

```
ARIMA

model <- auto.arima(stopy)
summary(model)
```

Hide

Series: stopy
ARIMA(3,0,2) with zero mean

Coefficients:

ar1	ar2	ar3	ma1	ma2
-0.1058	-0.1272	0.2997	0.7809	0.8234
s.e.	0.0922	0.0655	0.0791	0.0747

sigma^2 = 0.00116; log likelihood = 538.4
AIC=-1044.8 AICc=-1044.48 BIC=-1043.18

Training set error measures:

ME	RMSE	MAE	MPE	MAPE	MAE	ACF1
Training set	0.00247395	0.03999918	0.005467	NAN	1st	0.610391

Hide

```
#diagproponowany model: ARMA(3,2)
```

Hide

Rezulty modelu

Hide

```
res <- residuals(model)
Box.test(res, type = "Ljung-Box")
```

Hide

Box-Ljung test

data: res
X-squared = 1.2019, df = 1, p-value = 0.2717

Hide

```
#p-value większe od 0.05 -> brak autokorelacji
jarque.bera.test(res)
```

Hide

Jarque-Bera Test

data: res
X-squared = 2287.2, df = 2, p-value < 2.2e-16

Hide

```
#p-value mniejsze od 0.05 -> reszty nie są rozkładu normalnego
```

Hide

Hide

```
qqnorm(res)
```

Hide

Normal Q-Q Plot

Prognoza logarycznych stóp dla 4 przyszłych notowań i porównanie z odczytanymi wartościami

Hide

```
stopy_wyniki <- data.frame(
  Odczytane_Wartosci = stopy_0ct,
  Początek_Przedziału = forecast(model,4)$lower[, 2],
  Koniec_Przedziału = forecast(model,4)$upper[, 2],
  Średnia_Proгноza = forecast(model,4)$mean
)
print(stopy_wyniki)
```

Hide

Odczytane_Wartosci	Początek_Przedziału	Koniec_Przedziału	Średnia_Proгноza
-0.021394142	-0.07150475	0.05959546	-0.006024150
0.000000000	-0.08020952	0.05257170	-0.013641770
0.000000000	-0.05263535	0.07575583	0.008707442
-0.002699557	-0.05797036	0.07621352	0.009121580

4 rows

Chcę prognozy na pierwszy i trzeci notowań wykazując zgodność z rzeczywistością, kolejne notowania pokazują ujemne wartości średniej prognozy, co sugeruje przewidywanie spadków w tych okresach.

Model dla logarytmów cen

Hide

```
model2 <- auto.arima(ceny)
summary(model2)
```

Hide

Series: ceny
ARIMA(4,0,2) with non-zero mean

Coefficients:

ar1	ar2	ar3	ar4	ma1	ma2	mean
-0.0245	-0.2949	0.7186	0.2095	0.7919	0.8453	1.2135
s.e.	0.0972	0.0787	0.0679	0.0812	0.0782	0.0578

sigma^2 = 0.001055; log likelihood = 547.33
AIC=-1078.66 AICc=-1078.13 BIC=-1048.41

Training set error measures:

ME	RMSE	MAE	MPE	MAPE	MAE	ACF1
Training set	0.0006344601	0.0322061	0.02100185	-0.01401037	1.711285	1.011895

Hide

```
res2 <- residuals(model2)
Box.test(res2, type = "Ljung-Box")
```

Hide

Box-Ljung test

data: res2
X-squared = 0.0124, df = 1, p-value = 0.9119

Hide

```
#p-value większe od 0.05 -> brak autokorelacji
jarque.bera.test(res2)
```

Hide

Jarque-Bera Test

data: res2
X-squared = 2287.2, df = 2, p-value < 2.2e-16

Hide

```
#p-value mniejsze od 0.05 -> reszty nie są rozkładu normalnego
```

Hide

Hide

```
arima_forecast2 <- forecast(model2,4)$mean
```

Hide

Porównanie z odczytanymi wartościami

Hide

```
ceny_wyniki <- data.frame(
  Odczytane_Wartosci = ceny_0ct,
  Początek_Przedziału = exp(forecast(model2,4)$lower[, 2]),
  Koniec_Przedziału = exp(forecast(model2,4)$upper[, 2]),
  Średnia_Proгноza = exp(arima_forecast2)
)
print(ceny_wyniki)
```

Hide

Odczytane_Wartosci	Początek_Przedziału	Koniec_Przedziału	Średnia_Proгноza
273	1.311032	1.248221	1.367123
274	1.311032	1.204724	1.355944
275	1.311032	1.197584	1.372553
276	1.308333	1.193663	1.378063

4 rows

Dla każdego z analizowanych notowań, wartości dolnej i górnej granicy przedziałów ułożą się dookoła bliskie średniej prognozy, co wskazuje na umiarkowaną zmienność prognoz. Różnice między wartościami prognozowanymi a rzeczywistymi są niewielkie, co może sugerować dobrą jakość prognoz, ze modelem skutecznie uchwytując ogólny trend cenowy w krótkim okresie.

Analizując dla cen nieogrzewywalnych

Hide

```
ceny_n_wyniki <- data.frame(
  Odczytane_Wartosci = ceny_0ct,
  Początek_Przedziału = exp(forecast(model2,4)$lower[, 2]),
  Koniec_Przedziału = exp(forecast(model2,4)$upper[, 2]),
  Średnia_Proгноza = exp(arima_forecast2)
)
print(ceny_n_wyniki)
```

Hide

Odczytane_Wartosci	Początek_Przedziału	Koniec_Przedziału	Średnia_Proгноza
273	3.71	3.484139	3.959521
274	3.71	3.335839	3.915420
275	3.71	3.312104	3.945411
276	3.70	3.299144	3.971179

4 rows

W porównaniu do wcześniejszych wyników dotyczących logarytmicznych stóp zwrotu, które wskazywały na większą zmienność i niepewność prognoz, wyniki dla cen nieogrzewywalnych wydają się bardziej stabilne. Średnie prognozy cen są niewiele różnie od rzeczywistych wartości, co sugeruje, że model skutecznie przewidywał ogólny kierunek cenowy w krótkim okresie.

Metody symulacyjne

Hide

```
# Linowa symulacja
n <- 1000
# Błądki prognozy (4 notowania)
h <- 4
```

Hide

Wartość początkowa logarytmów cen (ostatnia znana wartość w cenach)

Hide

```
last_price <- as.numeric(tail(ceny, 1))
```

Hide

Symulacje Monte Carlo

Hide

```
set.seed(123) # Ustawienie ziarna dla powtarzalności wyników
symulacje <- matrix(0, nrow = n, ncol = h)
for (i in 1:n) {
  # Generowanie losowych rest modelu
  sim_res <- rnorm(h, mean=res1, sd=res2)
  # Symulacja logarytmów cen przy użyciu modelu
  symulacje[i, 1] <- last_price + sim_res[1]
  for (j in 2:h) {
    symulacje[i, j] <- (symulacje[i, j-1] + sim_res[j])
  }
}
```

Hide

Obliczenie 95% przedziałów ufności dla logarytmów cen

Hide

```
log_price_forecast_mean_mc <- colMeans(symulacje)
log_price_forecast_lower <- apply(symulacje, 2, quantile, probs = 0.025)
log_price_forecast_upper <- apply(symulacje, 2, quantile, probs = 0.975)
```

Hide

Konwersja logarytmów cen na rzeczywiste ceny

Hide

```
symulacje_ceny <- exp(symulacje)
```

Hide

Obliczenie 95% przedziałów ufności dla rzeczywistych cen

Hide

```
price_forecast_mean <- colMeans(symulacje_ceny)
price_forecast_lower <- apply(symulacje_ceny, 2, quantile, probs = 0.025)
price_forecast_upper <- apply(symulacje_ceny, 2, quantile, probs = 0.975)
```

Hide

Tworzenie tabeli porównawczej wyników

Hide

```
wyniki_monte_carlo <- data.frame(
  Odczytane_Wartosci_Log = ceny_0ct,
  Prognoza_Log_Ceny = log_price_forecast_mean_mc,
  Przedzial_Log_Ceny_Dolny = log_price_forecast_lower,
  Przedzial_Log_Ceny_Gorny = log_price_forecast_upper,
  Odczytane_Wartosci = exp(ceny_0ct),
  Prognoza_Ceny = price_forecast_mean,
  Przedzial_Ceny_Dolny = price_forecast_lower,
  Przedzial_Ceny_Gorny = price_forecast_upper
)
print(wyniki_monte_carlo)
```

Hide

Odczytane_Wartosci_Log	Prognoza_Log_Ceny	Przedzial_Log_Ceny_Dolny	Przedzial_Log_Ceny_Gorny
273	1.311032	1.333292	1.26747
274	1.311032	1.333502	1.24973
275	1.311032	1.334508	1.23236
276	1.308333	1.335842	1.216895

4 rows | 1-5 of 8 columns

W kontekście wcześniejszych wyników z modelu ARIMA, które analizowano w odniesieniu do stóp zwrotu oraz cen nieogrzewywalnych, wyniki uzyskane z metody Monte Carlo wydają się bardziej stabilne. Porównanie analizy wykazało, że prognozy cen są niewiele różnie od rzeczywistych wartości, co sugeruje, że model skutecznie przewidywał ogólny kierunek cenowy w krótkim okresie.

Metody bootstrapowe

Hide

```
set.seed(123) # Ustawienie ziarna dla powtarzalności wyników
bootstrap <- matrix(0, nrow = n, ncol = h)
for (i in 1:n) {
  # Generowanie losowych rest modelu
  boot_res <- sample(res1, h, replace = TRUE)
  # Symulacja logarytmów cen przy użyciu modelu
  bootstrap[i, 1] <- last_price + boot_res[1]
  for (j in 2:h) {
    bootstrap[i, j] <- (bootstrap[i, j-1] + boot_res[j])
  }
}
```

Hide

Obliczenie średnich prognoz i 95% przedziałów ufności dla logarytmów cen

Hide

```
log_price_forecast_mean_bootstrap <- colMeans(bootstrap)
log_price_forecast_lower <- apply(bootstrap, 2, quantile, probs = 0.025)
log_price_forecast_upper <- apply(bootstrap, 2, quantile, probs = 0.975)
```

Hide

Konwersja logarytmów cen na rzeczywiste ceny

Hide

```
bootstrap_prices <- exp(bootstrap)
```

Hide

Obliczenie średnich prognoz i 95% przedziałów ufności dla rzeczywistych cen

Hide

```
price_forecast_mean <- colMeans(bootstrap_prices)
price_forecast_lower <- apply(bootstrap_prices, 2, quantile, probs = 0.025)
price_forecast_upper <- apply(bootstrap_prices, 2, quantile, probs = 0.975)
```

Hide

Tworzenie tabeli wyników z prognozami i przedziałami ufności

Hide

```
wyniki_bootstrap <- data.frame(
  Odczytane_Wartosci_Log = ceny_0ct,
  Prognoza_Log_Ceny = log_price_forecast_mean_bootstrap,
  Przedzial_Log_Ceny_Dolny = log_price_forecast_lower,
  Przedzial_Log_Ceny_Gorny = log_price_forecast_upper,
  Odczytane_Wartosci = exp(ceny_0ct),
  Prognoza_Ceny = price_forecast_mean,
  Przedzial_Ceny_Dolny = price_forecast_lower,
  Przedzial_Ceny_Gorny = price_forecast_upper
)
print(wyniki_bootstrap)
```

Hide

Odczytane_Wartosci_Log	Prognoza_Log_Ceny	Przedzial_Log_Ceny_Dolny	Przedzial_Log_Ceny_Gorny
273	1.311032	1.333262	1.26747
274	1.311032	1.333575	1.24973
275	1.311032	1.333432	1.238401
276	1.308333	1.334226	1.227396

4 rows | 1-5 of 8 columns

Dla logarytmicznych cen, przedziały ufności są nieco węższe w analizie bootstrapowej w porównaniu do wyników z Monte Carlo, co sugeruje mniejszą niepewność w prognozach. Podobnie, prognozy cen nieogrzewywalnych są zbliżone z rzeczywistymi wartościami, a wąskie przedziały ufności mogą wskazywać na stabilność modelu.

Porównanie obu metod

Hide

```
wyniki_porownawcze <- data.frame(
  Odczytane_Wartosci = ceny_0ct,
  Prognoza_Monte_Carlo = log_price_forecast_mean_mc,
  Prognoza_Bootstrap = log_price_forecast_mean_bootstrap
)
print(wyniki_porownawcze)
```

Hide

Odczytane_Wartosci	Prognoza_Monte_Carlo	Prognoza_Bootstrap
273	1.311032	1.333292
274	1.311032	1.333502
275	1.311032	1.334508
276	1.308333	1.335842

4 rows

Wartości prognozowane zarówno przy metodzie Monte Carlo, jak i bootstrap są bardzo zbliżone do siebie oraz do rzeczywistych odczytanych wartości.

Różnice pomiędzy prognozami z obu metod są minimalne, co może sugerować, że obie podejścia oferują podobny poziom dokładności w prognozowaniu logarytmicznych cen. Wartość prognoz są zbliżone do prognoz GARCH. Różnice między modelami: Prognozy oparte na metodach stochastycznych na ogólną interpretację.

Prognozy z obu metod mieszczą się w wąskich granicach, co sugeruje o stabilności prognoz i sugeruje, że zastosowane modele są odpowiednie do przewidywania przyszłych cen.

Efekty ARCH

Hide

```
spec <- ugarchspec(
  variance.model = list(model = "sgARCH", garchOrder = c(1, 1)),
  mean.model = list(armaOrder = c(1, 2), include.mean = TRUE),
  distribution.model = "norm"
)
model_garch <- ugarchfit(spec = spec, data = data)
print(model_garch)
```

Hide

GARCH Model Fit

Conditional Variance Dynamics

GARCH Model: sgARCH(1,1)

Mean Model: ARFMA(1,6,2)

Distribution: norm

Optimal Parameters

	Estimate	Std. Error	t value	Pr(> t)
mu	1.252308	0.038882	32.207744	0.000000
ar1	1.805804	0.000311	5.8102e+03	0.000000
ar2	-0.668944	0.000321	-5.3294e+03	0.000000
ar3	-0.136679	0.000302	-4.5210e+03	0.000000
ma1	-1.430788	0.000296	-4.7654e+03	0.000000
ma2	0.394894	0.000391	2.0444e+03	0.000000
omega	0.002391	0.000557	6.4544e+02	0.000000
alpha	0.502886	0.154055	3.2650e+00	0.001095
beta	0.005050	0.005719	5.8511e-02	0.928278

Robust Standard Errors

	Estimate	Std. Error	t value	Pr(> t)
mu	1.252308	0.038882	32.207744	0.000000
ar1	1.805804	0.000322	5.6444e+03	0.000000
ar2	-0.668944	0.001286	-520.555103	0.000000
ar3	-0.136679	0.002597	-52.428391	0.000000
ma1	-1.430788	0.032163	-43.863943	0.000000
ma2	0.394894	0.015159	25.710219	0.000000
omega	0.002391	0.002191	0.284449	0.78257
alpha	0.502886	13.082146	0.038485	0.98933
beta	0.005050	2.633861	0.002112	0.99831

LogLikelihood: 571.605

Information Criteria

	Akaike	-4.1358
	Bayes	-4.0375
	Dubina	-4.1389
	Hannan-Quinn	-4.0889

Weighted Ljung-Box Test on Standardized Residuals

	Statistic	p-value
Lag[1]	1.045	1.927e-01
Lag[2]*g(pq)+g(pq-1)[1:4]	12.354	6.429e-13
Lag[4]*g(pq)+g(pq-1)[1:8]	10.871	1.577e-02

D.F.:45

D.F.:45

Weighted Ljung-Box Test on Standardized Squared Residuals

	Statistic	p-value
Lag[1]	1.098	2.946e-01
Lag[2]*g(pq)+g(pq-1)[1:8]	24.545	2.946e-07
Lag[4]*g(pq)+g(pq-1)[1:8]	29.919	4.593e-07

D.F.:162

Weighted ARCH LM Tests

	Statistic	Shape	Scale	p-value
Lag[1]	0.0114	1.563	2.005	0.418
ARCH Lag[5]	1.4212	1.440	1.667	0.413
ARCH Lag[7]	1.5819	2.315	1.543	0.8048

Rhynch stability test

Joint Statistics: 2.0631

Individual Statistics:

mu	0.0923
ar1	0.04595
ar2	0.09359
ar3	0.09228
ma1	0.08374
ma2	0.10464
omega	0.3374
alpha	0.15111
beta	0.13333

Asymptotic Critical Values (10, 5, 1)

Joint Statistics: 1.1

Individual Statistics: 0.35

Sign Bias Test

	t-value	prob	sig
Sign Bias	<0.001	<0.001	<0.001
Negative Sign Bias	1.0317492	0.3031236	
Positive Sign Bias	1.6405649	0.1021713	
Joint Effect	3.7542887	0.2892481	

4 rows

Adjusted Pearson Goodness-of-Fit Test:

	group	statistic	p-value(g-1)
1	20	49.88	0.0001910
2	30	61.07	0.0003223
3	40	64.82	0.0036394
4	50	76.53	0.0071724

Elapsed time: 0.395569

Testy Ljung-Boxa wykazują powolny stopień autokorelacji w resztach, co może sugerować niedopasowanie modelu, ale testy ARCH nie wykazują istotnych oznak autokorelacji. Wyniki te sugerują, że model jest w stanie dobrze opisać dane.

Prognozy modelu ARMA-GARCH

Hide

```
forecast_garch <- ugarchfcast(model_garch, n.ahed = h)
print(forecast_garch)
```

Hide

GARCH Model Forecast

Model: sgARCH

Rhynch: 1.4

Null reject: 0

Out of Samples: 0

0-out forecast [170=10-30-10]:

Series	Sigma
T+1	1.312269
T+2	1.314102
T+3	1.313275
T+4	1.312169

Hide

```
log_garch_mean <- fitted(forecast_garch)
print(log_garch_mean)
```

Hide

1970-09-30

T+1	1.312269
T+2	1.314102
T+3	1.313275
T+4	1.312169

Rezulty z symulacji GARCH

Hide

```
res3 <- residuals(model_garch)
```

Hide

Hide

```
set.seed(123)
mq <- matrix(0, nrow = n, ncol = h)
for (i in 1:n) {
  # Generowanie losowych rest modelu
  sim_res <- rnorm(h, mean=res3, sd=res3)
  # Symulacja logarytmów cen przy użyciu modelu
  mq[i, 1] <- last_price + sim_res[1]
  for (j in 2:h) {
    mq[i, j] <- (mq[i, j-1] + sim_res[j])
  }
}
```

Hide

set.seed(123)

boot <- matrix(0, nrow = n, ncol = h)

for (i in 1:n) {

boot_res <- sample(res3, h, replace = TRUE)

Symulacja logarytmów cen przy użyciu modelu

boot[i, 1] <- last_price + boot_res[1]

for (j in 2:h) {

boot[i, j] <- (boot[i, j-1] + boot_res[j])

}

Hide

```
log_gmc_mean <- colMeans(mq)
gmc_mean <- colMeans(exp(mq))
log_boot_mean <- colMeans(boot)
boot_mean <- colMeans(exp(boot))
```

Hide

Hide

```
ostatnia_porownanie <- data.frame(
  Odczytane_Wartosci = ceny_0ct,
  Prognoza_ARMA = arima_forecast,
  Prognoza_MC = log_price_forecast_mean_mc,
  Prognoza_Bootstrap = log_price_forecast_mean_bootstrap,
  Prognoza_GARCH = unname(log_garch_mean),
  Prognoza_MC_GARCH = log_gmc_mean,
  Prognoza_Bootstrap_GARCH = log_boot_mean
)
print(cotnie(ostatnia_porownanie))
```

Hide

Odczytane_Wartosci	Prognoza_ARMA	Prognoza_MC	Prognoza_Bootstrap	Prognoza_GARCH
273	1.311032	1.312172	1.333292	1.312969
274	1.311032	1.285334	1.333502	1.333375
275	1.311032	1.285068	1.334508	1.333342
276	1.308333	1.286393	1.335842	1.334226

4 rows | 1-5 of 8 columns

Porównanie prognoz: Tabela przedstawia porównanie prognozowanych wartości uzyskanych z różnych modeli: ARIMA, Monte Carlo, Bootstrap oraz GARCH. Prognozy ARIMA są w dużej mierze zbliżone do prognoz GARCH. Różnice między modelami: Prognozy oparte na metodach Monte Carlo i Bootstrap w większości przypadków przewidywają nieco wyższe wartości niż te uzyskane z modeli ARIMA i GARCH.

Przebieżenie testu ARCH

Hide

```
test_arch1 <- ArchTest(res2)
print(test_arch1)
```

Hide

ARCH LM-test; Null hypothesis: no ARCH effects

data: res2

Chi-squared = 5.6088, df = 12, p-value = 0.9345

Hipoteza zerowa testu stwierdza, że nie występuje efektu ARCH, co oznacza, że wariancja reszt jest stała w czasie.

Wyniska p-wartości (0.9345) sugeruje, że nie ma wystarczających dowodów, aby odrzucić hipotezę zerową. Wyniki wskazują, że w analizowanych resztach nie ma istotnych efektów ARCH, co wskazuje na stabilność wariancji reszt.

Hide

```
test_arch2 <- ArchTest(residuals(model_garch))
print(test_arch2)
```

Hide

ARCH LM-test; Null hypothesis: no ARCH effects

data: residuals(model_garch)

Chi-squared = 23.092, df = 12, p-value = 0.01439

Test ARCH sprzeciwia hipotezie zerowej, że nie występuje efektu ARCH w resztach modelu, co oznacza, że wariancja reszt jest stała w czasie. Wyniki te sugerują, że model jest w stanie dobrze opisać dane.

Wartości p-wartości (0.01439) sugeruje, że nie ma wystarczających dowodów, aby odrzucić hipotezę zerową. Wyniki wskazują, że w analizowanych resztach nie ma istotnych efektów ARCH, co wskazuje na stabilność wariancji reszt.

Podstawowy model nie wykazuje efektu ARCH, jednak przeprowadzając go na modelu GARCH uzyskujemy odwrotny wynik.

1. Badanie stacjonarności i logarytmicznych stóp zwrotu

Wstępna analiza stacjonarności została przeprowadzona za pomocą testu ADF (Augmented Dickey-Fuller). Wyniki wykazały stacjonarność stóp zwrotu, co było kluczowe dla przeprowadzenia dalszej analizy czasowej. Chiągając stacjonarność potwierdza, że dane są odpowiednie do modelowania przy użyciu metod ARIMA/ARMA, pozwalając na skuteczniejsze przewidywanie przyszłych zmian cen na podstawie historycznych not