Daria Lasikowska

# Model GARCH(1,1)

## przy zastosowaniu ekonometrii bayesowskiej



## Spis treści



2017-01-28

## Opis zagadnienia

Polska waluta jest szczególnie wrażliwa na zawirowania polityczne i ekonomiczne. Jako waluta kraju wschodzącego mierzy się nie tylko z czynnikami wewnętrznymi wpływającymi na jej kwotowania, ale także globalnymi. Szczególnie widoczne było to chociażby w drugiej połowie 2016 roku, kiedy w czerwcu Brytyjczycy w referendum opowiedzieli się za wystąpieniem z Unii Europejskiej oraz w październiku po zwycięstwie barwnego Donalda Trumpa w wyborach na prezydenta Stanów Zjednoczonych. W obu sytuacjach niepewność na rynku światowym popchnęła zagranicznych inwestorów do ograniczenia zaangażowania w polskiej walucie. Wahliwość kwotowań złotego w stosunku do euro pokazuje podatność polskiej waluty na szoki zewnętrzne. Poniżej postaram się przybliżyć zmienność kursu EUR/PLN za pomocą modelu GARCH(1,1), zarówno w pracy Osiewalskiego (2001) jak i Ardia i Hoogerheide (2010) został on opisany jako najbardziej oszczędny i efektywny.



## Teoretyczne informacje wstępne

Modelowanie danych finansowych wykorzystywane jest do określenia ryzyka związanego z inwestycją, do opisania którego najczęściej zastosowanie znajdują miary zmienności (np. cen akcji, kursów walutowych). Cechą charakterystyczną tych procesów jest występowanie zmiennej wariancji warunkowej. Podczas niepewnej sytuacji gospodarczej występują okresy dużych wahań cen, po których, wraz z chwilową stabilizacją rynku, nastają okresy małej zmienności. Dodatkowo na poziomy cen mają często ogromny wpływ jednorazowe wydarzenia generujące obserwacje nietypowe, co w modelowaniu znajduje odwzorowanie w postaci tzw. „grubych ogonów" rozkładu zmiennej.

Do opisania zmienności wykorzystywane są m.in. parametryczne modele statystyczne, mające odtworzyć możliwie najtrafniej proces generowania danych. Jednym z nich jest model GARCH(p,q) [Generalized Auto-Regressive Conditional Heteroskedasticity model, uogólniony ARCH], w którym obecna wariancja zależy od *p* wcześniejszej realizacji procesu oraz autoregresyjnie od wariancji warunkowych z *q* poprzednich okresów. Poniżej zamieszczono równanie wariancji GARCH(1,1):

$$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

dla t=1,…,T oraz α,β >0, niedotrzymanie tego warunku wiązałoby się z możliwością uzyskania ujemnej wariancji.

Natomiast zmienna opisana jest wzorem:

$$y_t = \varepsilon_t \sigma_t$$

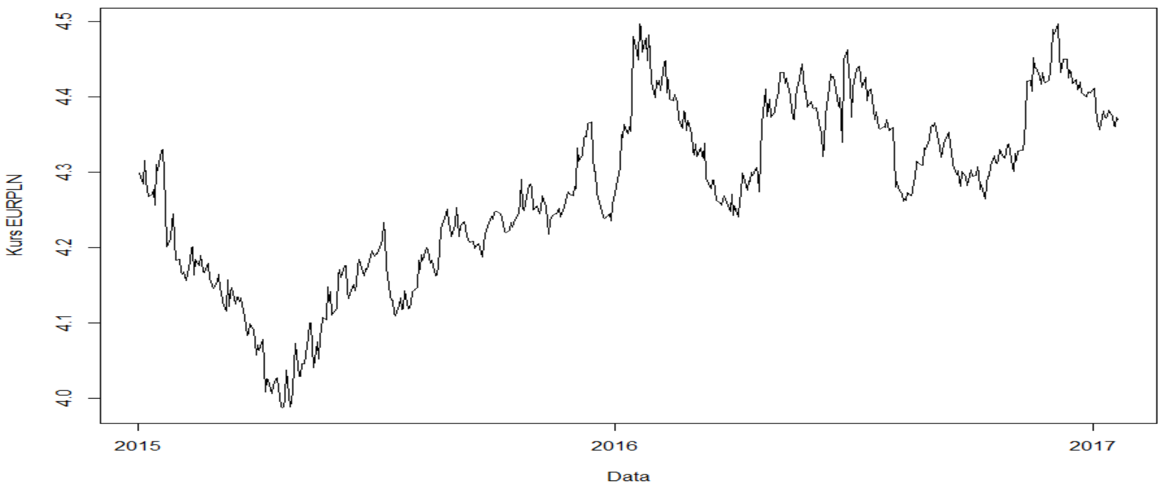
gdzie $\varepsilon_t$ są i.i.d., dla t=1,…,T. Powszechnie przyjmuje się, że składniki losowe mają rozkład normalny, jednak ze względu na wspomniane wcześniej „grube ogony” w analizie szeregów czasowych danych finansowych najczęściej stosuje się rozkład t-Studenta.

## Modelowanie

Niestety bayesowskie podejście w analizie danych finansowych jest dość oszczędnie reprezentowane w postaci dostępnych pakietów programu R. W wyliczeniach skorzystam z pakietu bayesGARCH, którego zasadniczą wadą jest brak możliwości określenia rzędu innego niż p=1 i q=1, co uniemożliwia porównanie innych specyfikacji oraz ograniczenia wprowadzania rozkładów a priori. Zaletą jest prostota zastosowania bazująca na algorytmie Metropolisa - Hastingsa.

### Szereg czasowy

Kurs EUR/PLN obejmuje okres od 2 stycznia 2015r. do 20 stycznia 2017r. co daje T=532.

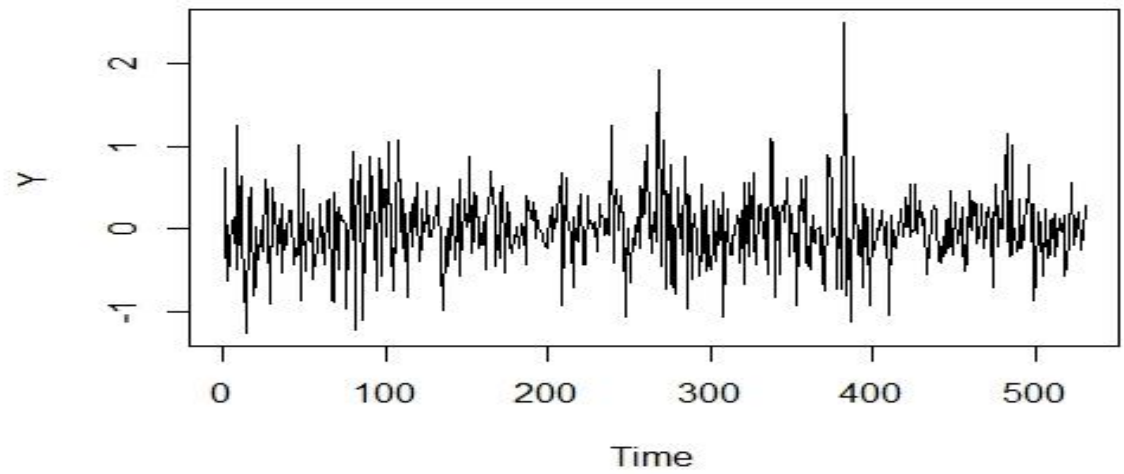



**Wykres 1 Notowanie EURPLN w analizowanym okresie, dane: stooq.pl**

Wstępnie powyższy szereg czasowy trudno określić jako stacjonarny, co więcej od drugiego kwartału charakteryzuje się trendem rosnącym. W związku z powyższym dalej podstawą do modelowania będzie szereg dziennych logarytmicznych stóp zwrotu wyrażony w procentach:

$$y_t = 100\ln\left(\frac{x_t}{x_{t-1}}\right).$$

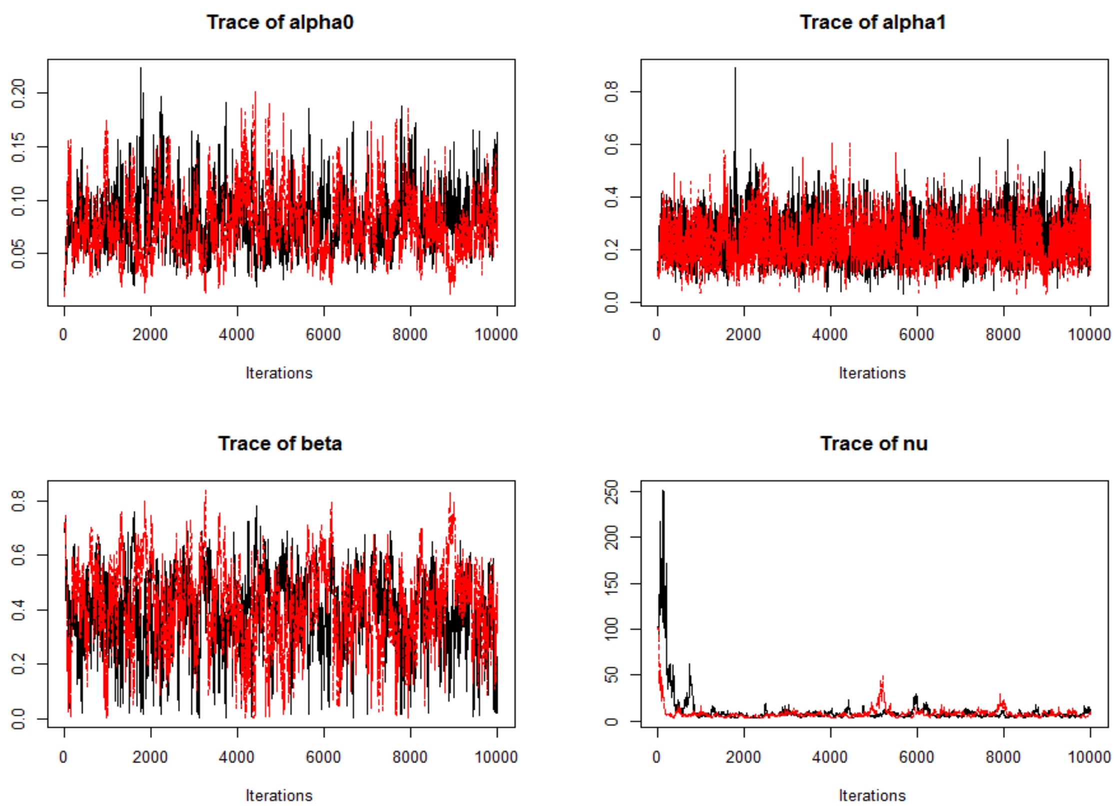
Po przekształceniu szereg w większym stopniu przypomina „biały szum”.




Wykres 2 Szereg logarytmicznych stóp zwrotu kursu EURPLN, styczeń 2015-styczeń 2017

## Zbieżność łańcuchów MCMC

Na tak przekształconych danych obliczam GARCH(1,1) z rozkładem t-Studenta. Ponieważ przed modelowaniem nie posiadam wiedzy odnośnie rozkładu a priori, pozostawiam wartości startowe zapisane dla tej funkcji przez autorów pakietu. Metodą MCMC generuję dwa łańcuchy o długości 10000 obserwacji.




Wykres 3 Zbieżności łańcuchów otrzymanych metodą MCMC

Test Gelmana – Rubina wskazuje na zbieżność łańcuchów, oszacowania w większości są poniżej wartości 1.1 oraz wszystkie są mniejsze niż odpowiadające im wartości z górnej granicy przedziału ufności.

```
Potential scale reduction factors:

       Point est. Upper C.I.
alpha0       1.02       1.10
alpha1       1.01       1.05
beta         1.02       1.10
nu           1.11       1.32

Multivariate psrf

1.05
```

Algorytm MCMC uzyskał bardzo wysokie odsetki przyjętych wartości 87% dla oszacowań α oraz 94% dla β, świadcząc, że przyjęte rozkłady są rekurencyjne. Liczba stopni swobody czyli nu nie zależy od swoich poprzednich realizacji, jest losowana przy każdej iteracji algorytmu Metropolisa-Hastingsa.

```
  alpha0    alpha1      beta        nu 
0.8713371 0.8713371 0.9424942 1.0000000
```

Stopień skorelowania wartości w łańcuchach ze swoimi poprzedniczkami z okresu opóźnionego o jeden jest bardzo wysoki, aby uniknąć tej autokorelacji z łańcuchów zostaną wybrane co drugie wartości. Dodatkowo odrzuconych zostanie 5000 pierwszych losowań jako burn in, aby uniknąć obciążenia związanego ze wstępnymi założeniami co do losowania.

```
         alpha0     alpha1      beta        nu
Lag 0  1.0000000 1.00000000 1.0000000 1.0000000
Lag 1  0.9298854 0.67967724 0.9549540 0.9867633
Lag 5  0.7770793 0.30435875 0.8292899 0.9356798
```

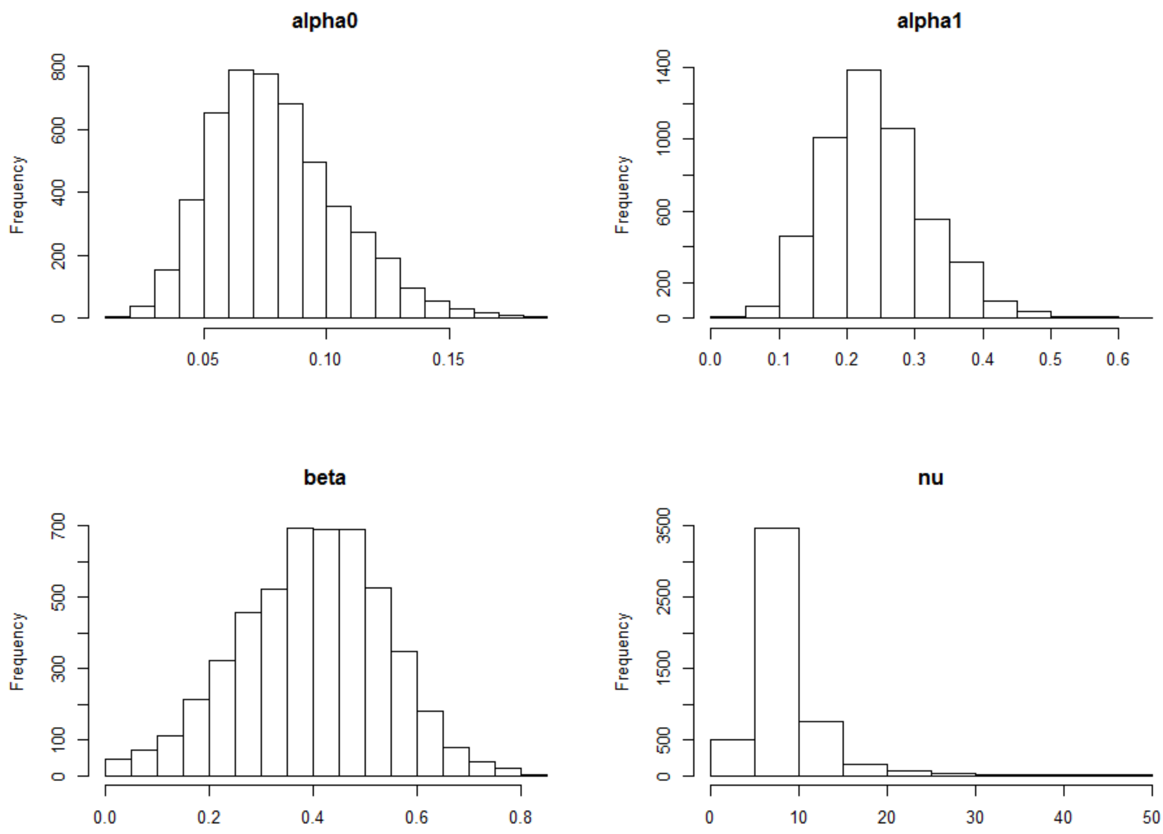
Poniżej zamieszczone wyniki oszacowań wskazują na większe znaczenie opóźnionej wariancji warunkowej niż przeszłości procesu w tworzeniu bieżącej wariancji warunkowej. Szczególną uwagę zwracają oszacowania nu tj. stopni swobody, gdyż potwierdzają występowanie „grubych ogonów", a co za tym idzie wykorzystanie rozkładu t-Studenta. Dla wartości nu>30 adekwatne byłoby zastosowanie warunkowego rozkładu normalnego, jednak w opisywanym przypadku 97,5% oszacowań stopni swobody przyjęło wartości mniejsze niż 20.

```
1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

          Mean      SD  Naive SE Time-series SE
alpha0 0.07946 0.02647 0.0003743       0.002049
alpha1 0.24167 0.07606 0.0010757       0.003577
beta   0.39784 0.13992 0.0019788       0.011579
nu     8.37018 4.30772 0.0609204       0.515319
```

```
2. Quantiles for each variable:

          2.5%     25%     50%     75%   97.5%
alpha0 0.03573 0.06035 0.07609 0.09578  0.1390
alpha1 0.11253 0.18860 0.23506 0.28781  0.4064
beta   0.10289 0.30377 0.40526 0.49498  0.6595
nu     4.34254 5.87243 7.30445 9.49101 19.8755
```

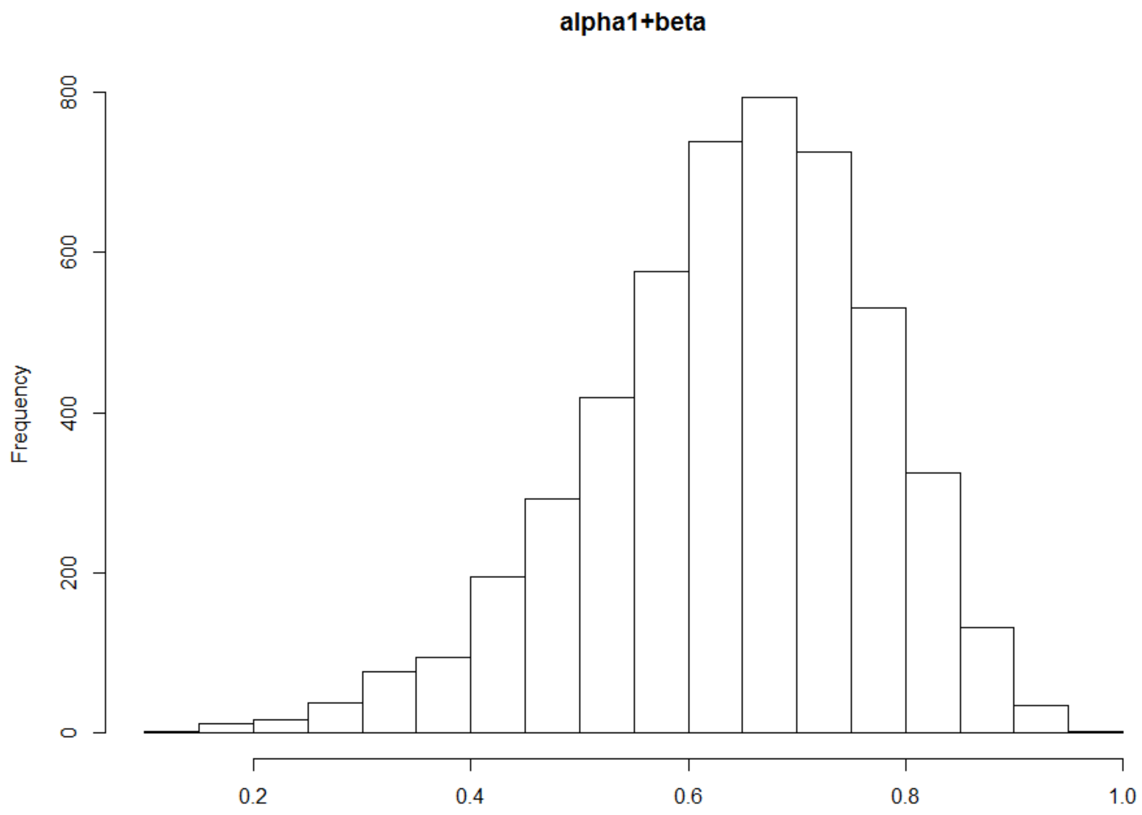
### Rozkłady warunkowe oszacowanych parametrów

Poniższe histogramy charakteryzują się wyraźną skośnością, zgodnie z pracą Ardia i Hoogerheide (2010) posiadane stosunkowo dużej próby – w przypadku tego modelu 532 obserwacji – nie uzasadnia użycia rozkładu asymptotycznie normalnego jako rozkładu oszacowań parametrów.




Wykres 4 Histogramy szacowanych parametrów

Za pomocą histogramu można także sprawdzić stacjonarność kowariancyjną procesu. Jeśli suma $\alpha 1$ i $\beta$ jest bliska 1 poprzednie realizacje i wariancje będą w większym stopniu budować wariancję warunkową dla przyszłych okresów.




Wykres 5 Histogram sumy oszacowań $\alpha_1$ i β

## Bibliografia

Ardia D., Hoogerheide L.F. "Bayesian Estimation of the GARCH(1,1)Model with Student-t Innovations", The R Journal Vol. 2/2, 2010

https://journal.r-project.org/archive/2010-2/RJournal_2010-2_Ardia+Hoogerheide.pdf

Osiewalski J., „Ekonometria bayesowska w zastosowaniach", Wydawnictwo AE w Krakowie, 2001.rozdział 3.