Case of stock-market by using GARCH

以 光大证券(SH601788)股票的对数收益率为例。

光大证券[601788]A股实时行情 - 百度股市通

有 400个观测值。读入数据：

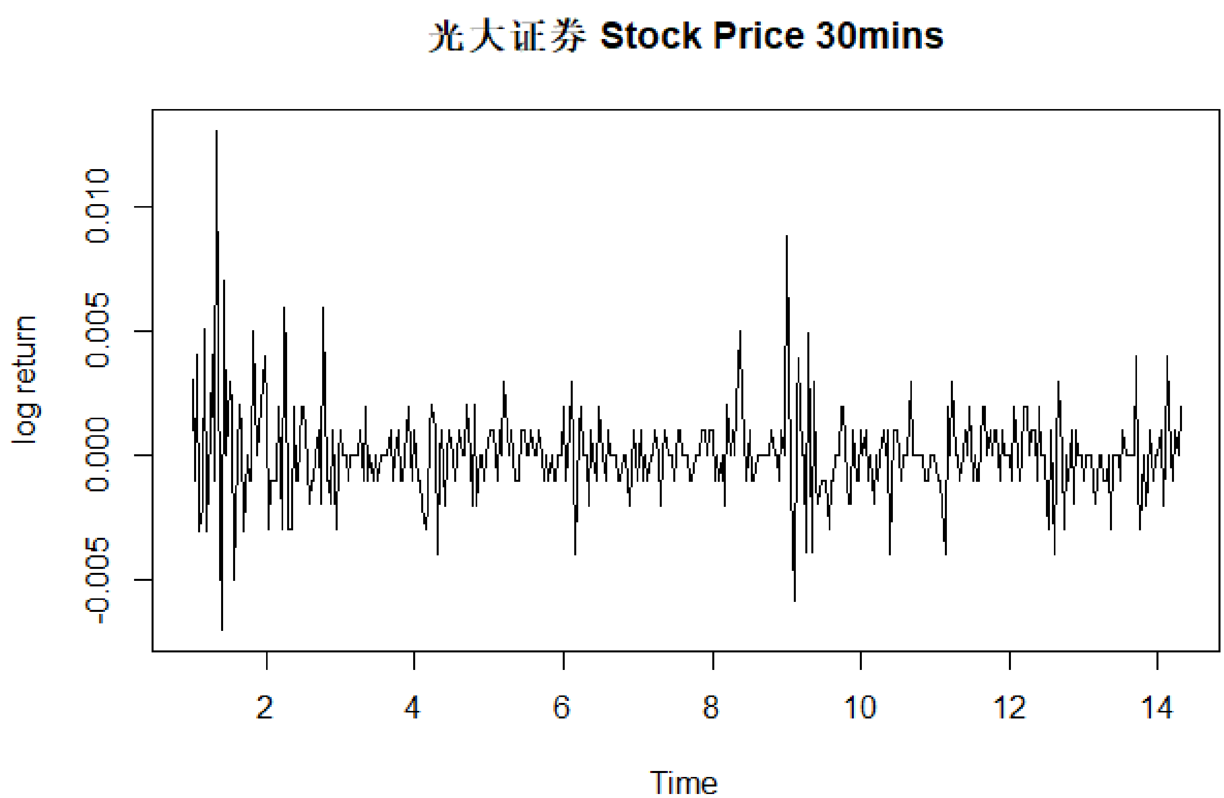
```
www<- read.table("D:/教学资料/研究生/时间序列/Chapter6/601788.txt")
print(www)
stock <- www[8]
```

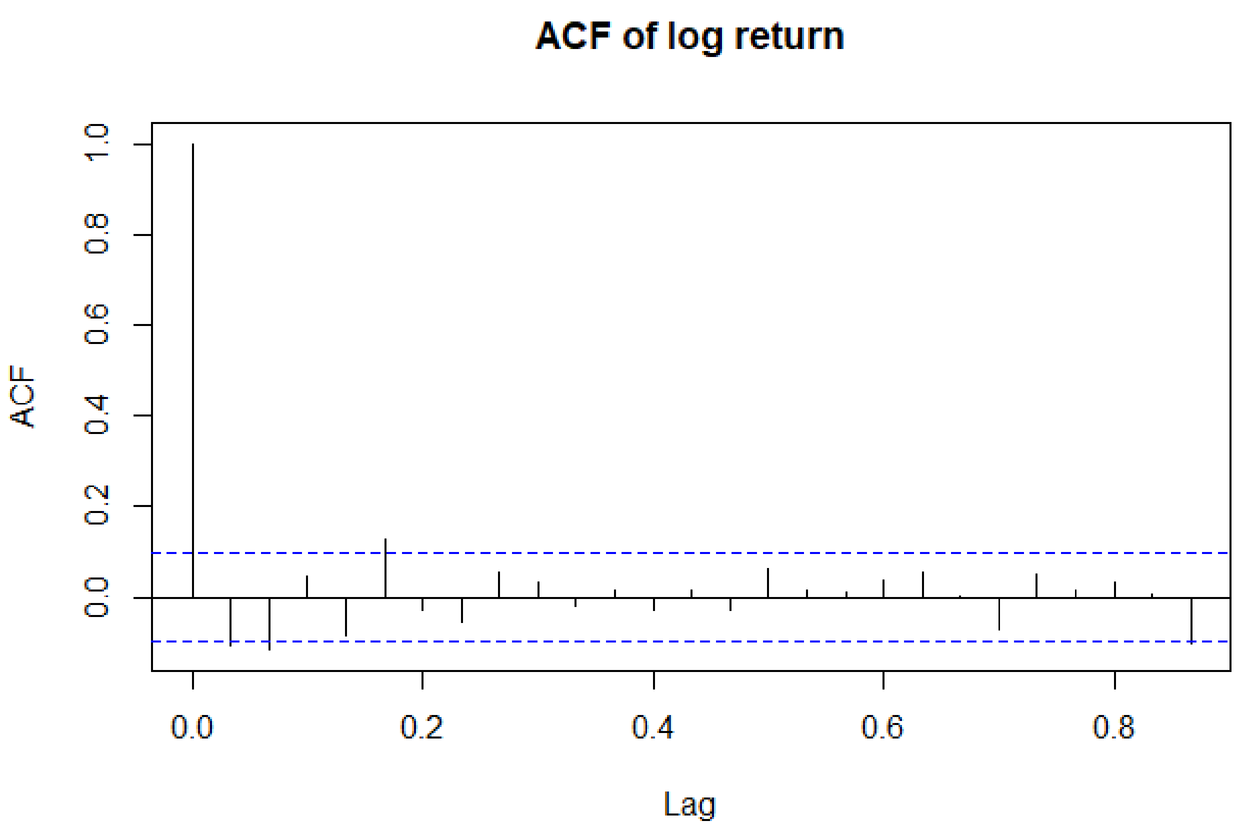
	时间	收盘价	开盘价	最高价	最低价	成交额（元）	成交量（手）	对数收益率
1	20130801093100	9.84	9.81	9.9	9.81	649650	591	0.003053437
2	20130801093200	9.83	9.84	9.87	9.82	634400	576	-0.001016777
3	20130801093300	9.87	9.83	9.87	9.82	1182750	1077	0.004060919
4	20130801093400	9.84	9.84	9.84	9.83	605256	551	-0.003044142
5	20130801093500	9.82	9.82	9.84	9.82	392178	357	-0.002034589
6	20130801093600	9.87	9.82	9.87	9.82	413637	376	0.005078731
7	20130801093700	9.84	9.82	9.87	9.82	147590	134	-0.003044142
8	20130801093800	9.85	9.84	9.85	9.84	487357	443	0.001015744
9	20130801093900	9.89	9.84	9.89	9.84	872757	791	0.00405269
10	20130801094000	9.88	9.86	9.9	9.86	144723	131	-0.001011634

1. Data Analysis

First of all we turn the data into time series with 30 mins. We note this series as $\left{ r_t \right}$.

```r
#转化为时间序列
ts.stock <- ts(stock, frequency=30)
#作对数收益率的时间序列图:
plot(ts.stock, ylab="log return", main="光大证券 Stock Price Monthly Log Return")
```





We can see from ACF figure of $\left{ r_t \right}$, with only a few lags slightly out of boundary, so we can consider it as a white noise.（这里的白噪声特指零均值、不相关的弱平稳时间序列）。

Then we apply ***Ljung Box white noise test***:

```
Box.test(ts.stock, lag=30, type="Ljung")
```

```
> #L jung白噪声检验
> Box.test(ts.stock, lag=30, type="Ljung")

        Box-Ljung test

data:  ts.stock
X-squared = 41.24, df = 30, p-value = 0.083
```

Since p-value is 0.083, it passed the white noise test at 0.05 level.

Then we test whether the mean of $\left\{ r_t \right\}$ is 0:

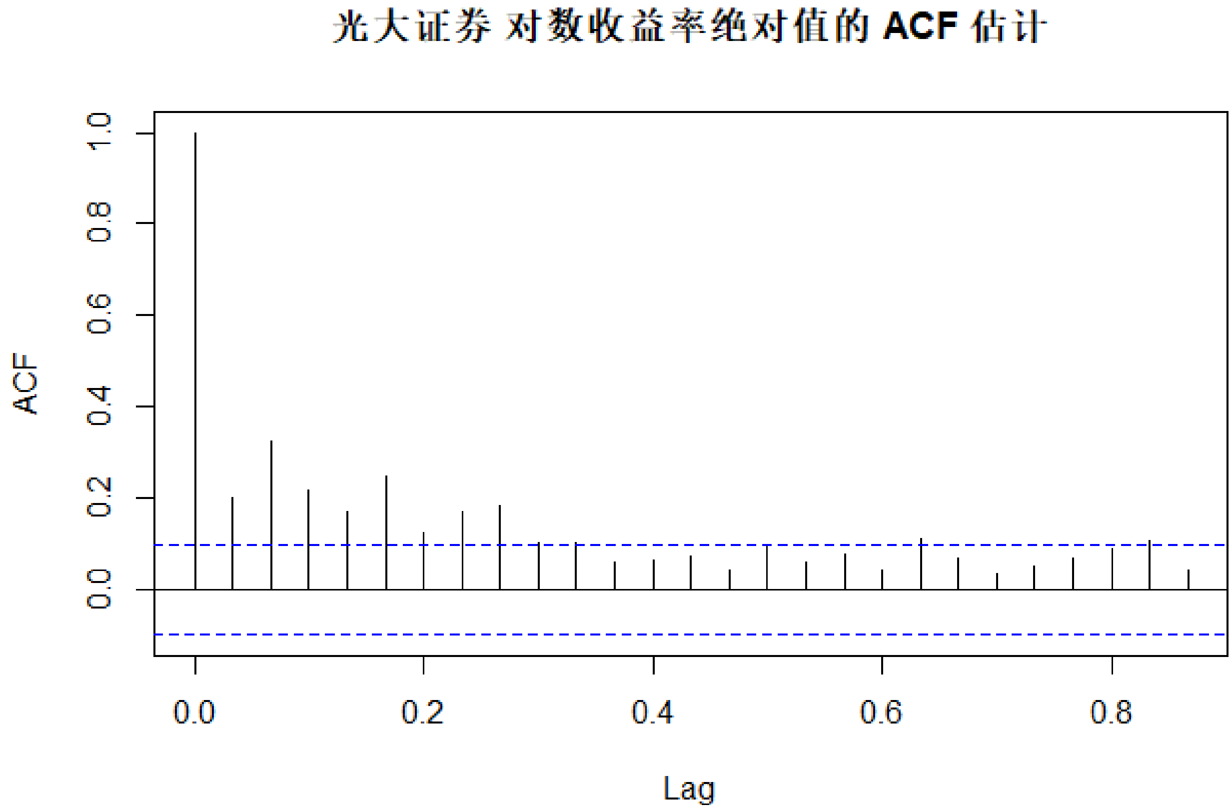
```
t.test(c(ts.stock))
```

```
> #检验 {r_t} 序列的均值是否等于零:
> t.test(c(ts.stock))

        One Sample t-test

data:  c(ts.stock)
t = 0.62201, df = 399, p-value = 0.5343
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.0001251804  0.0002410546
sample estimates:
   mean of x
5.793709e-05
```

We can see the mean approach to 0.

Then we consider about the $\left\{|r_t|\right\}$, so make an ACF figure of it:

```
acf(abs(ts.stock), main="光大证券 对数收益率绝对值的 ACF 估计")
```



We can see lots of values are out of bound.

Apply the **Ljung-Box test**:

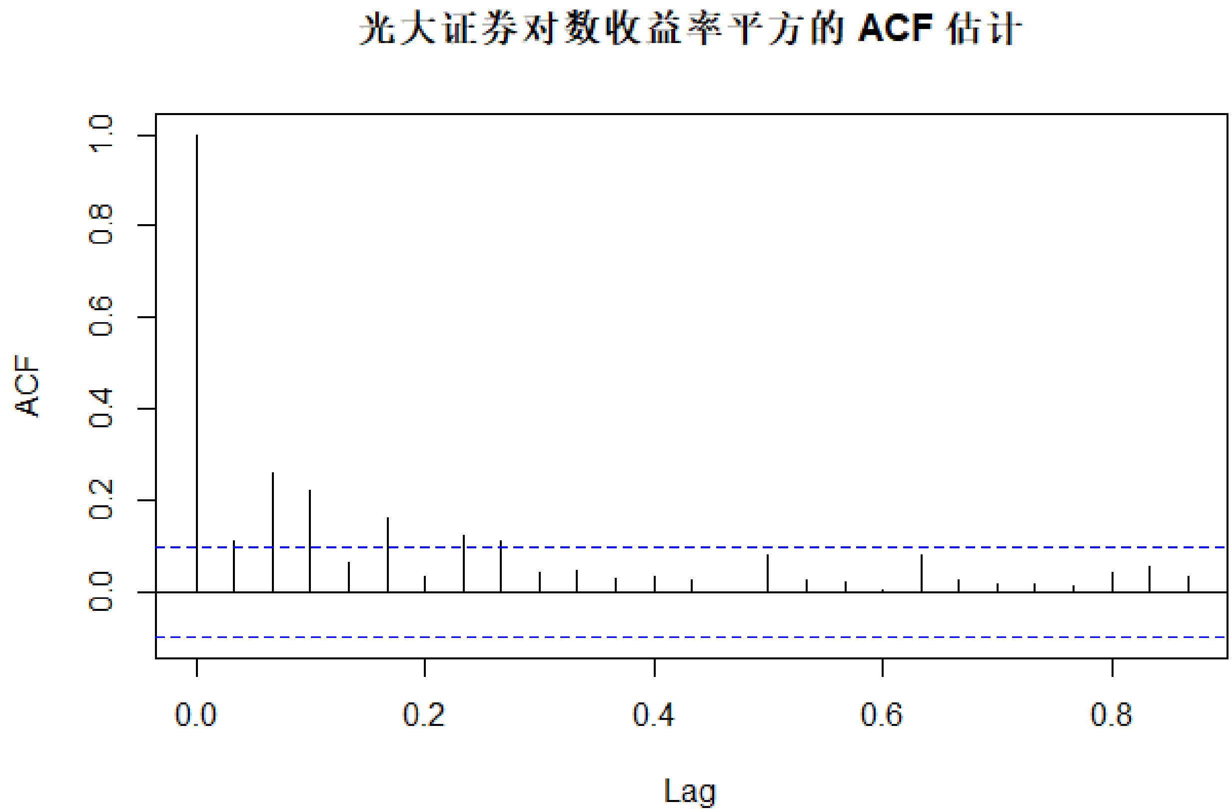
```
Box.test(abs(ts.stock), lag=12, type="Ljung")
```

```
> Box.test(abs(ts.stock), lag=30, type="Ljung")

        Box-Ljung test

data:  abs(ts.stock)
X-squared = 207.9, df = 30, p-value < 2.2e-16
```

The result significantly rejected the white noise null hypothesis. So we can say, $\left{ |r_t|\right}$ performs: stable, no correlation but not independent with terms nearby.

Furthermore we print the ACF and apply the Ljung-box test to $\left{ r_t^2\right}$:

```
acf(ts.stock^2, main="光大证券对数收益率平方的 ACF 估计")
Box.test(ts.stock^2, lag=30, type="Ljung")
```




```
> Box.test(ts.stock^2, lag=30, type="Ljung")

        Box-Ljung test

data:  ts.stock^2
X-squared = 99.825, df = 30, p-value = 1.978e-09
```

The ACF estimation and white-noise test shows $\left{r_t\right}$ have the auto-correlation,not independent.

2. Construct model

First of all we need to test whether the data fits Garth model.

```
#ArchTest检验（函数代码此处省略
ArchTest(ts.stock - mean(ts.stock), m=30)
```

```
xmat27       1.214e-01  3.095e-02   3.921 0.000107 ***
xmat28       2.857e-02  3.099e-02   0.922 0.357215
xmat29      -2.376e-02  3.025e-02  -0.785 0.432818
xmat30      -5.240e-02  3.021e-02  -1.734 0.083749 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.825e-06 on 339 degrees of freedom
Multiple R-squared:  0.1637,    Adjusted R-squared:  0.08972
F-statistic: 2.212 on 30 and 339 DF,  p-value: 0.0003862
```

It passed the test, so it's proper way to use Garch.

We use normal distribution to construct GARCH model.

```
#采用正态条件分布建立 GARCH(1,1) 模型:
library(fGarch, quietly = TRUE)
mod1 <- garchFit(~ 1 + garch(1,1), data=ts.stock, trace=FALSE)
summary(mod1)
```

```
> summary(mod1)

Title:
 GARCH Modelling

Call:
 garchFit(formula = ~1 + garch(1, 1), data = ts.stock, trace = FALSE)

Mean and Variance Equation:
 data ~ 1 + garch(1, 1)
<environment: 0x000001b057b37c58>
 [data = ts.stock]

Conditional Distribution:
 norm

Coefficient(s):
        mu       omega      alpha1       beta1
2.9739e-05  3.7088e-07  1.6130e-01  7.2689e-01

Std. Errors:
 based on Hessian

Error Analysis:
        Estimate  Std. Error  t value Pr(>|t|)
mu     2.974e-05   7.645e-05    0.389   0.6973
omega  3.709e-07   1.602e-07    2.315   0.0206 *
alpha1 1.613e-01   7.451e-02    2.165   0.0304 *
beta1  7.269e-01   1.048e-01    6.938 3.98e-12 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Log Likelihood:
 1992.165    normalized:  4.980412

Description:
 Sun Nov 13 22:06:23 2022 by user: 10306
```

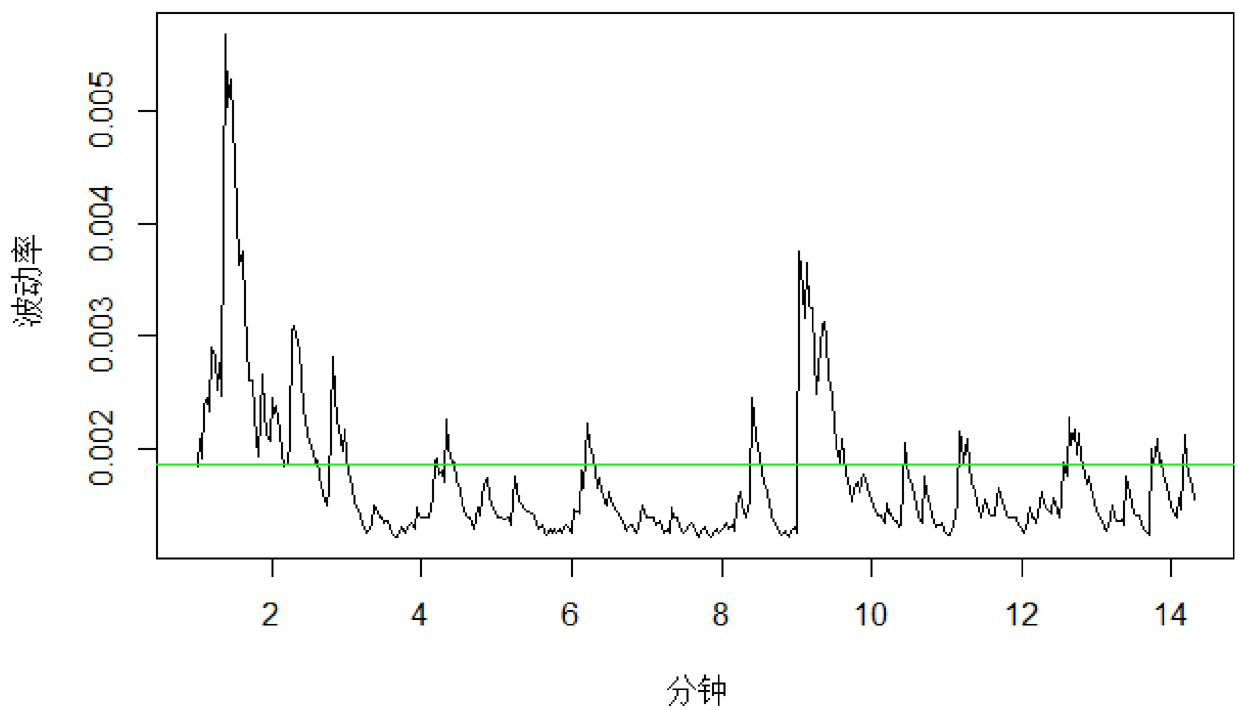

```
Description:
 Sun Nov 13 22:06:23 2022 by user: 10306
```


```
Standardised Residuals Tests:
                                Statistic p-Value
 Jarque-Bera Test   R    Chi^2  1096.244  0
 Shapiro-Wilk Test  R    W      0.8974312 9.409665e-16
 Ljung-Box Test     R    Q(10)  3.834157  0.9545185
 Ljung-Box Test     R    Q(15)  6.140595  0.977288
 Ljung-Box Test     R    Q(20)  8.71917   0.9859375
 Ljung-Box Test     R^2  Q(10)  1.722616  0.998058
 Ljung-Box Test     R^2  Q(15)  2.481819  0.9998786
 Ljung-Box Test     R^2  Q(20)  8.210088  0.9903841
 LM Arch Test       R    TR^2   1.586213  0.999824
```

```
Information Criterion Statistics:
      AIC       BIC       SIC      HQIC 
-9.940824 -9.900909 -9.941021 -9.925017 
```

Then we try to fit the volatility figure.

```
# 拟合的波动率图形：
vola <- volatility(mod1)
plot(ts(vola, start=start(ts.stock), frequency=frequency(ts.stock)),
     xlab=" 分钟", ylab=" 波动率")
abline(h=sd(ts.stock), col="green")
```



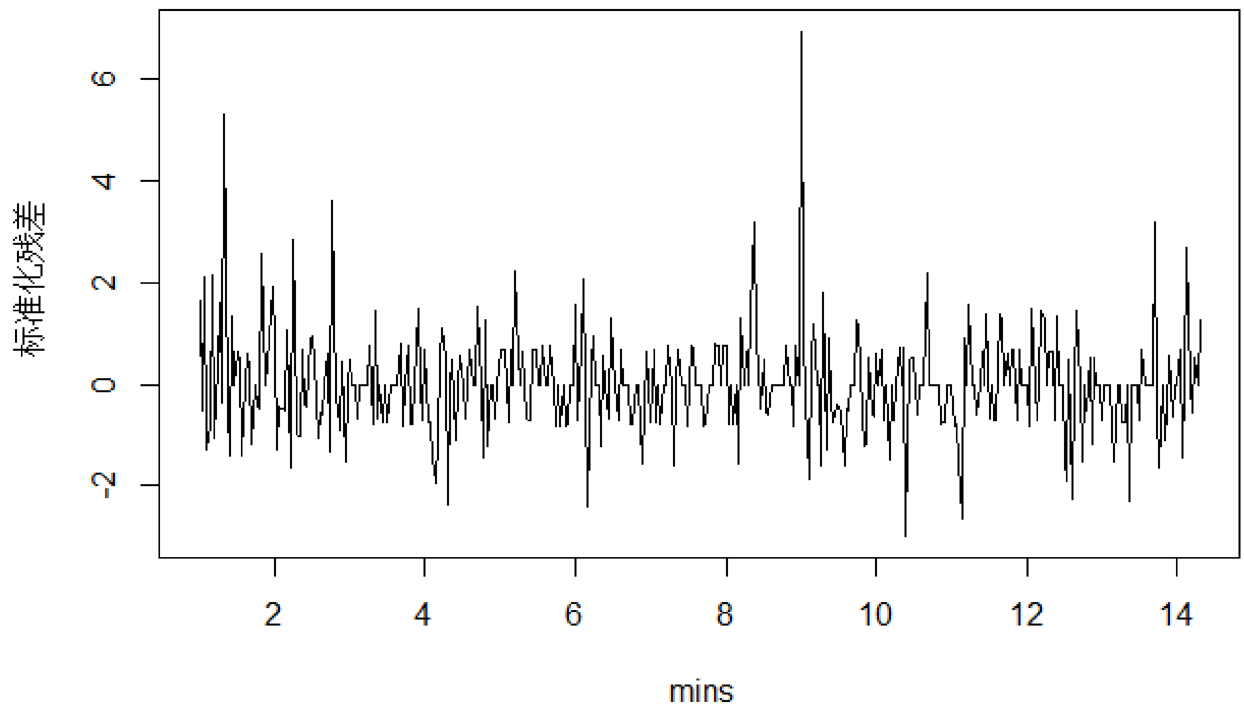
The green line is the standard error of the sample.

We can see from the figure: In the open time of morning and afternoon, the volatility reach a high level.

3. Test of the model

First we have an overview of standardized residual:

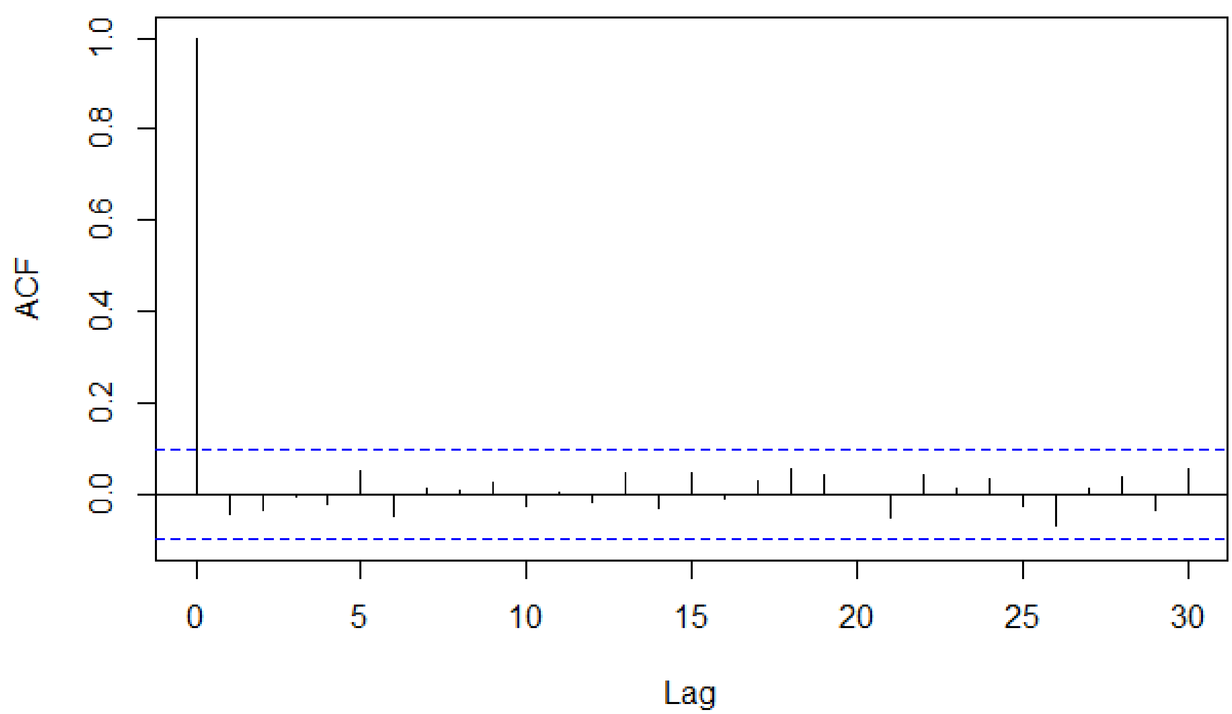
```
#标准化残差的时间序列图：
resi <- residuals(mod1, standardize=TRUE)
plot(ts(resi, start=start(ts.stock), frequency=frequency(ts.stock)),
     xlab=" mins", ylab=" 标准化残差")
```

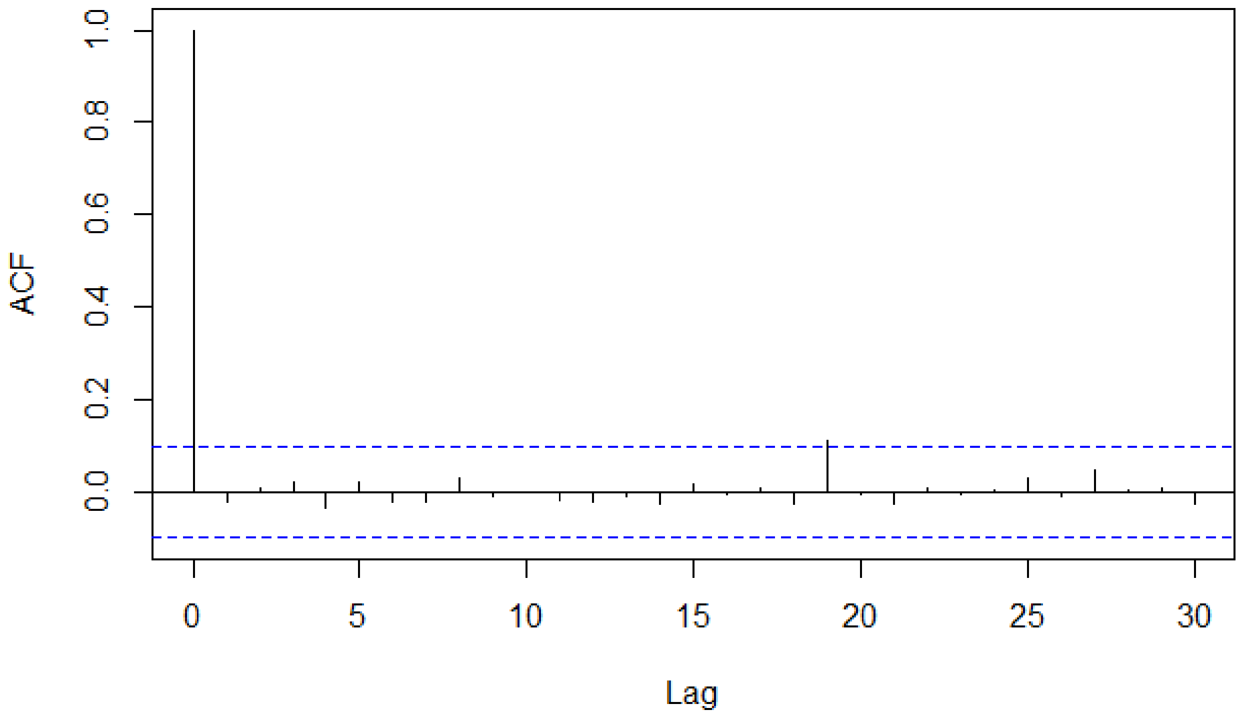


We can see except some outliers, it satisfies the idd (independent identically distributed).

Next we test the ACF of $\left{ a_t \right}$ and $\left{ a_t^2 \right}$:

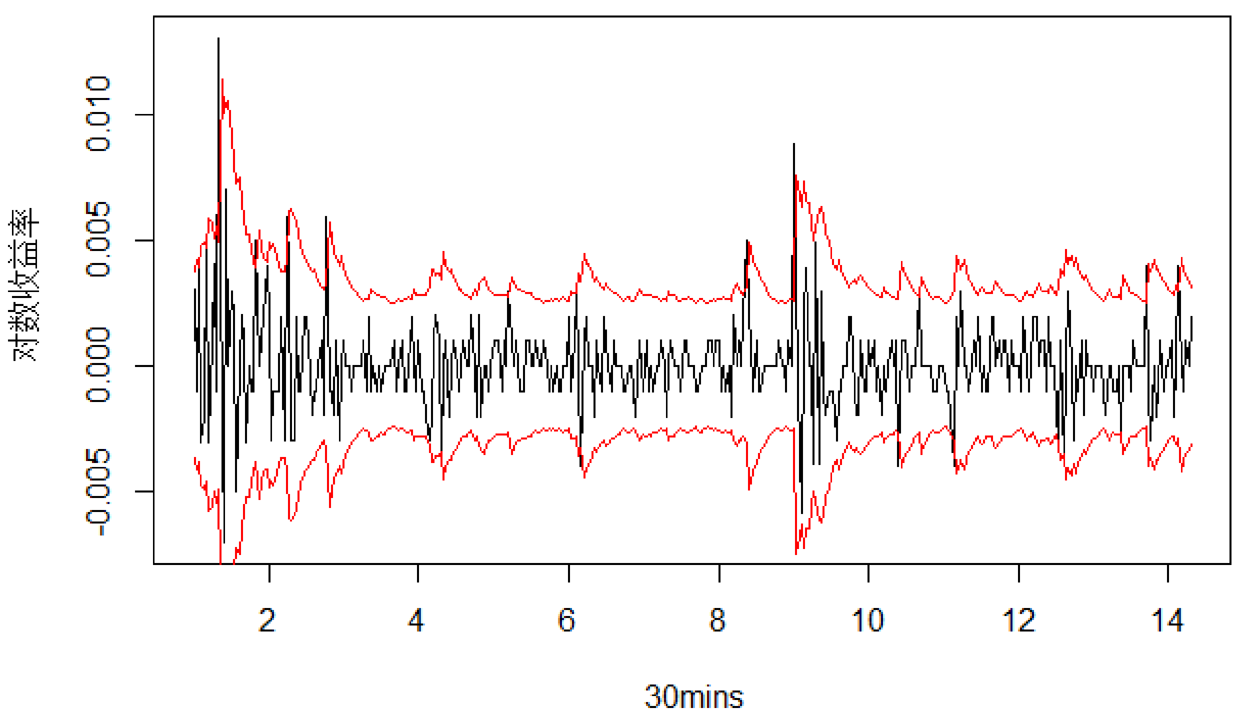
```
resi <- residuals(mod1, standardize=TRUE)
acf(resi, lag.max=30, main="")
resi <- residuals(mod1, standardize=TRUE)
acf(resi^2, lag.max=30, main="")
```





Only have a single point slightly exceeds the boundary.

Finally, we use $ \hat \mu_t + 2\hat\sigma_t$ as 95% CI, connect each CI points as lines in t-axis direction, we get a figure:



It can be seen that the values of logarithmic rate of return is almost in the prediction interval.

Or we can draw a conclusion that we get a proper model.