高余慈 B08204005

177/180

Subject: Time Series Midterm Exam   No.:   Date: 2023 / 05 / 02

1. Assume $(1-0.5B)X_t = (1+0.6B)(1-0.3B)a_t$, $\{a_t\} \sim WN(0,1)$

30

(a) Express X_t as an $MA(\infty)$ process

(b) Find the Autocovariance generating function of X_t based on (a)

(c) Find recursive formula for the autocorrelation function of X_t , $k \geq 3$

(a) $X_t = \left(\frac{1}{1-0.5B}\right)(1+0.6B)(1-0.3B)a_t$

$= (1+0.5B + 0.25B^2 + \cdots)(1+0.6B)(1-0.3B)a_t$

$= (1+0.5B + 0.25B^2 + \cdots)(1+0.3B-0.18B^2)a_t$

$= a_t[(1+0.3B-0.18B^2) + (0.5B+0.15B^2-0.09B^3) + (0.25B^2+0.015B^3-0.045B^4) + (0.125B^3+0.0475B^4-0.0225B^5)+\cdots]$

$= a_t(1+0.8B + 0.22B^2 + 0.11B^3 + \cdots)$

$= a_t + 0.8a_{t-1} + \sum_{k=2}^{\infty} 0.22 \times \left(\frac{1}{2}\right)^{k-2} a_{t-k}$ ※

(b) $r(B) = \psi(B)\psi(B^{-1})\sigma_a^2$

$= \left(\sum_{i=0}^{\infty}(0.5B)^i\right)(1+0.3B-0.18B^2)\left(\sum_{i=0}^{\infty}(0.5B)^{-i}\right)(1+0.3B^{-1}-0.18B^{-2})$

$= \left(\sum_{k=-\infty}^{\infty}\left(\frac{1}{2}\right)^{|k|}\left(\frac{4}{3}\right)B^k\right)(-0.18B^{-2}+0.246B^{-1}+1.1224+0.246B-0.18B^2)$

$= -0.18\left(\sum_{k=-\infty}^{\infty}\left(\frac{1}{2}\right)^{|k|}\left(\frac{4}{3}\right)B^{k-2}\right) + 0.246\left(\sum_{k=-\infty}^{\infty}\left(\frac{1}{2}\right)^{|k|}\left(\frac{4}{3}\right)B^{k-1}\right)$

$-0.18\left(\sum_{k=-\infty}^{\infty}\left(\frac{1}{2}\right)^{|k|}\left(\frac{4}{3}\right)B^{k+2}\right) + 0.246\left(\sum_{k=-\infty}^{\infty}\left(\frac{1}{2}\right)^{|k|}\left(\frac{4}{3}\right)B^{k+1}\right)$

$+1.1224\left(\sum_{k=-\infty}^{\infty}\left(\frac{1}{2}\right)^{|k|}\left(\frac{4}{3}\right)B^{k}\right)$ ※

$\Rightarrow r_k = \frac{4}{3}\left(-0.18\left(\frac{1}{2}\right)^{|k+2|} + 0.246\left(\frac{1}{2}\right)^{|k+1|} + 1.1224\left(\frac{1}{2}\right)^{|k|} + 0.246\left(\frac{1}{2}\right)^{|k-1|} - 0.18\left(\frac{1}{2}\right)^{|k-2|}\right)$ ※

(c) $\rho_k = 0.5\rho_{k-1}$, for $k \geq 3$ ※

2. Assume that the quarterly log return r_t of an asset follows the model:

30

$(1-0.4B)(1-B)(1-B^4)r_t = (1-0.2B^4)a_t$, $\{a_t\} \sim WN(0,1)$

(a) $W_t = (1-B)(1-B^4)r_t$ with regular and seasonal differencing , $Var(W_t)$?

(b) Suppose $r_{799}=0.9$, $r_{798}=0.5$, $r_{797}=0.7$, $r_{796}=0.6$, $r_{795}=0.4$

$r_{794}=0.3$, $a_{799}=0.1$, $a_{798}=0.3$, $a_{797}=-0.4$, $a_{796}=-0.5$

$a_{795}=0.2$, Find 1-step ahead forecast of r_{800} at origin $t=799$

(c) Error ?

95% prediction interval of r_{800} ?

No.
Date : :

(a) $\Rightarrow (1-0.4B)W_t = (1-0.2B^4)a_t$

$\Rightarrow W_t = 0.4W_{t-1} + a_t - 0.2a_{t-4}$

$\gamma_0 = Var(W_t) = Var(0.4W_{t-1} + a_t - 0.2a_{t-4})$

$= 0.16\gamma_0 + \sigma_a^2 + 0.04\sigma_a^2 - 2\cdot 0.2\, Cov(a_t, a_{t-4})$ 0

$+ 2\cdot 0.4\, Cov(W_{t-1}, a_t) - 2\cdot 0.4\cdot 0.2\, Cov(W_{t-1}, a_{t-4})$ 0

$W_{t-1} = (1+0.4B + (0.4)^2B^2 + (0.4)^3B^3 + \cdots)(1-0.2B^4)a_{t-1}$

$= a_{t-1} + 0.4a_{t-2} + (0.4)^2a_{t-3} + (0.4)^3a_{t-4} + \cdots$

$\Rightarrow Cov(W_{t-1}, a_{t-4}) = Cov(a_{t-1} + 0.4a_{t-2} + (0.4)^2a_{t-3} + (0.4)^3a_{t-4} + \cdots, a_{t-4})$

$= 0.4^3\, Cov(a_{t-4}, a_{t-4})$

$= 0.4^3\, Var(a_{t-4}) = 0.4^3\sigma_a^2$

$\Rightarrow \gamma_0 = 0.16\gamma_0 + 1 + 0.04 - 0.16\cdot 0.4^3$

$\Rightarrow 0.84\gamma_0 = 1 + 0.04 - 0.01024 = 1.02976$

$\Rightarrow \gamma_0 = 1.225905$ #

(b) $(1-1.4B + 0.4B^2 - B^4 + 1.4B^5 - 0.4B^6)r_t = (1-0.2B^4)a_t$

$r_t - 1.4r_{t-1} + 0.4r_{t-2} - r_{t-4} + 1.4r_{t-5} - 0.4r_{t-6} = a_t - 0.2a_{t-4}$

$\Rightarrow r_{800} = 1.4r_{799} - 0.4r_{798} + r_{796} - 1.4r_{795} + 0.4r_{794} + a_{800} - 0.2a_{796}$

$\Rightarrow \hat{r}_{799}(1) = 1.4\cdot 0.9 - 0.4\cdot 0.5 + 0.6 - 1.4\cdot 0.4 + 0.4\cdot 0.3 - 0.2\cdot(-0.5)$

$= 1.32$ #

(c) $\hat{e}_{799}(1) = r_{800} - \hat{r}_{799}(1) = a_{800} \Rightarrow se(\hat{e}_{799}(1)) = se(a_{800}) = 1$ #

95 % P.I. : $1.32 \pm 1.96\cdot 1 \Rightarrow [-0.64, 3.28]$ #

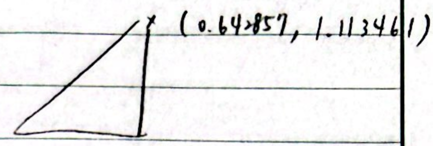
Chryo culture

3

20

$C(B) = 405 - 639B + 497B^2 - 196B^3 = 0$

$B = 0.642857 \pm 1.113461i$ and 1.25 (By R)



$\Rightarrow \quad \alpha = \sqrt{0.642857^2 + 1.113461^2} = 1.285714$

$\theta = \tan^{-1}\left(\frac{d}{c}\right) = \tan^{-1}\left(\frac{1.113461}{0.642857}\right) = 1.047198$

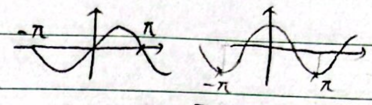
and $p = \frac{2\pi}{\theta} = 6$ #

Yes, the average period of the cycle is 6 #

10

$E(Z_t) = E(U)\,E(\sin(9t+\theta)) + E(\cos(9t+\theta))$ (By U, θ indep)

$= 0 + E(\cos(9t+\theta))$



$= \int_{-\pi}^{\pi} \frac{1}{2\pi} \cos(9t+\theta)\, d\theta$

$= \frac{1}{2\pi}\left[\int_{-\pi}^{\pi} \cos(\theta)\cos(9\theta)\,d\theta - \int_{-\pi}^{\pi} \sin(\theta)\sin(9\theta)\,d\theta\right]$ $(9t-\pi)$

$= \frac{1}{2\pi}(0-0) = 0$ #

$Cov(Z_t Z_{t+k}) = E(Z_t Z_{t+k}) - \underbrace{E(Z_t)E(Z_{t+k})}_{0}$

求 $E(Z_t Z_{t+k}) = E\big((U\sin(9t+\theta) + \cos(9t+\theta))(U\sin(9(t+k)+\theta) + \cos(9(t+k)+\theta))\big)$

$= E[U^2 \sin(9t+\theta)\sin(9(t+k)+\theta)] +$ $(\theta, U$ indep.$)$

$E[U\sin(9t+\theta)\cos(9(t+k)+\theta)] +$ $\rightsquigarrow 0$ $(\theta, U$ indep.$)$, $E(U)=0$

$E[U\sin(9(t+k)+\theta)\cos(9t+\theta)] +$ $\rightsquigarrow 0$ $(\theta, U$ indep.$)$ $E(U)=0$

$E[\cos(9(t+k)+\theta)\cos(9t+\theta)]$ $(\theta, U$ indep$)$

$\because Var(U) = E(U^2) - \underbrace{E(U)}_{0} = 1 \quad \therefore E(U^2) = 1$

$\cos(a+b) = \cos a\cos b - \sin a \sin b$

$+) \quad \cos(a-b) = \cos a\cos b + \sin a \sin b$

$\frac{1}{2}(\cos(a+b) + \cos(a-b)) = \cos a \cos b$

$-\frac{1}{2}(\cos(a+b) - \cos(a-b)) = \sin a \sin b$

$= E[\underbrace{\sin(9t+\theta)}_{b}\underbrace{\sin(9(t+k)+\theta)}_{a}] +$

$E(\underbrace{\cos(9(t+k)+\theta)}_{a}\underbrace{\cos(9t+\theta)}_{b})$

$\theta + 9t + 9$
$\theta \quad 9t$
$9(2t+$

$= E\left[\frac{1}{2}(\cos(9(2t+k)+2\theta) + \cos(9k))\right]$

$E\left[-\frac{1}{2}(\cos(9(2t+k)+2\theta) - \cos(9k))\right]$

$= \frac{1}{2}E(\cos(9k)) + \frac{1}{2}E(\cos(9k)) = \cos(9k)$ ※ 跟 t 沒關係

故 $\{Z_t\}$ is covariance stationary #

Time Series midterm exam

B082040005 高念慈

2023-05-14

40

1.

- 2022/1/3 ~ 2022/12/30
- 初始 100
- aapl 0.3
- intc 0.5
- msft 0.2
- 2023/1/3 to 2023/3/31
- aapl 0.2
- intc 0.2
- msft 0.6

```
data_1 = read.csv("C:/Users/user/Desktop/time_series/HW/2023_midexam/data_1.csv",header=T)
head(data_1)
```

```
##         date     AAPL     INTC     MSFT
## 1 2022-01-03 180.6839 50.56260 330.8138
## 2 2022-01-04 178.3907 50.49609 325.1414
## 3 2022-01-05 173.6455 51.18977 312.6599
## 4 2022-01-06 170.7468 51.32280 310.1893
## 5 2022-01-07 170.9156 50.78115 310.3474
## 6 2022-01-10 170.9354 52.46309 310.5747
```

```
AAPL_return = dailyReturn(ts(data_1[,2]))
INTC_return = dailyReturn(ts(data_1[,3]))
MSFT_return = dailyReturn(ts(data_1[,4]))

return_data = data.frame(AAPL_return, INTC_return, MSFT_return)

rownames(return_data) = as.Date(data_1[,1], format = "%Y-%m-%d")
colnames(return_data) = c('AAPL', 'INTC', 'MSFT')
```

a. What is the 4-period log-return of the portfolio from 2022-04-25 to 2022-04-29 ?

```
portfolio_data = data.frame(data_1[1:251,1])
portfolio_data[, 2] = data_1[1:251,2]/data_1[1,2]*30
portfolio_data[, 3] = data_1[1:251,3]/data_1[1,3]*50
portfolio_data[, 4] = data_1[1:251,4]/data_1[1,4]*20
portfolio_data[, 5] = portfolio_data[,2] + portfolio_data[,3] + portfolio_data[4]

colnames(portfolio_data) = c('date', 'AAPL', 'INTC', 'MSFT', 'value')
log(portfolio_data[82,5]) - log(portfolio_data[78,5])
```

```
## [1] -0.05043428
```

```
## -0.05043428
```

```
a_data = return_data[79:82,]

log((prod(a_data[,1]+1)-1)*0.3 + (prod(a_data[,2]+1)-1)*0.5 + (prod (a_data[,3]+1)-1)*0.2 +1)
```

```
## [1] -0.05001056
```

```
## -0.05001056
```

b. What is the average daily simple return of the portfolio from 2022-04-01 to 2022-04-29 ?

```
b_data = return_data[63:82,]

(prod (b_data[,1]*0.3 + b_data[,2]*0.5 + b_data[,3]*0.2 + 1))^(1/length(b_data[,1]))-1
```

```
## [1] -0.005738161
```

```
## -0.005738161
```

```
(portfolio_data[82,5]/portfolio_data[63,5])^(1/length (b_data[,1]))-1
```

```
## [1] -0.005048979
```

```
## -0.005048979
```

c. What is the cumulative simple return of the portfolio from 2022/1/3 to 2022/12/30 ?

```
c_data = return_data[2:251,]
prod((c_data[,1])*0.3 + (c_data[,2])*0.5 + (c_data[,3])*0.2 + 1) - 1
```

```
## [1] -0.3829065
```

```
## -0.3829065
```

```
portfolio_data[251,5]/portfolio_data[1,5]-1
```

```
## [1] -0.38176
```

```
## -0.38176
```

d. What is the cumulative log return of the portfolio from 2022/1/3 to 2023/3/31 ?

```
portfolio_data_2 = data.frame(data_1[252:313,1])

portfolio_data_2[, 2] = data_1[252:313,2]/data_1[252,2]*portfolio_data[251,5]*0.2
portfolio_data_2[, 3] = data_1[252:313,3]/data_1[252,3]*portfolio_data[251,5] *0.2
portfolio_data_2[, 4] = data_1[252:313,4]/data_1[252,4] *portfolio_data[251,5]*0.6

portfolio_data_2[, 5] = portfolio_data_2[,2] + portfolio_data_2[,3] + portfolio_data_2[, 4]

log((portfolio_data[251,5]/portfolio_data[1,5])*(portfolio_data_2[62, 5]/portfolio_data_2[1,
5]))
```

```
## [1] -0.2695324
```

```
## -0.2695324
```

```
d_data = return_data[252:313,]

c_return = prod((c_data[,1])*0.3 + (c_data[,2])*0.5 + (c_data[,3])*0.2 + 1)* prod((d_data[,
1])*0.2 + (d_data[,2])*0.2 + (d_data[,3])*0.6 +1)

log(c_return)
```

```
## [1] -0.2744475
```

```
## -0.2744475
```

2. 47

```
data_2 = read.csv("C:/Users/user/Desktop/time_series/HW/2023_midexam/data_2.csv",header=T)
head(data_2)
```

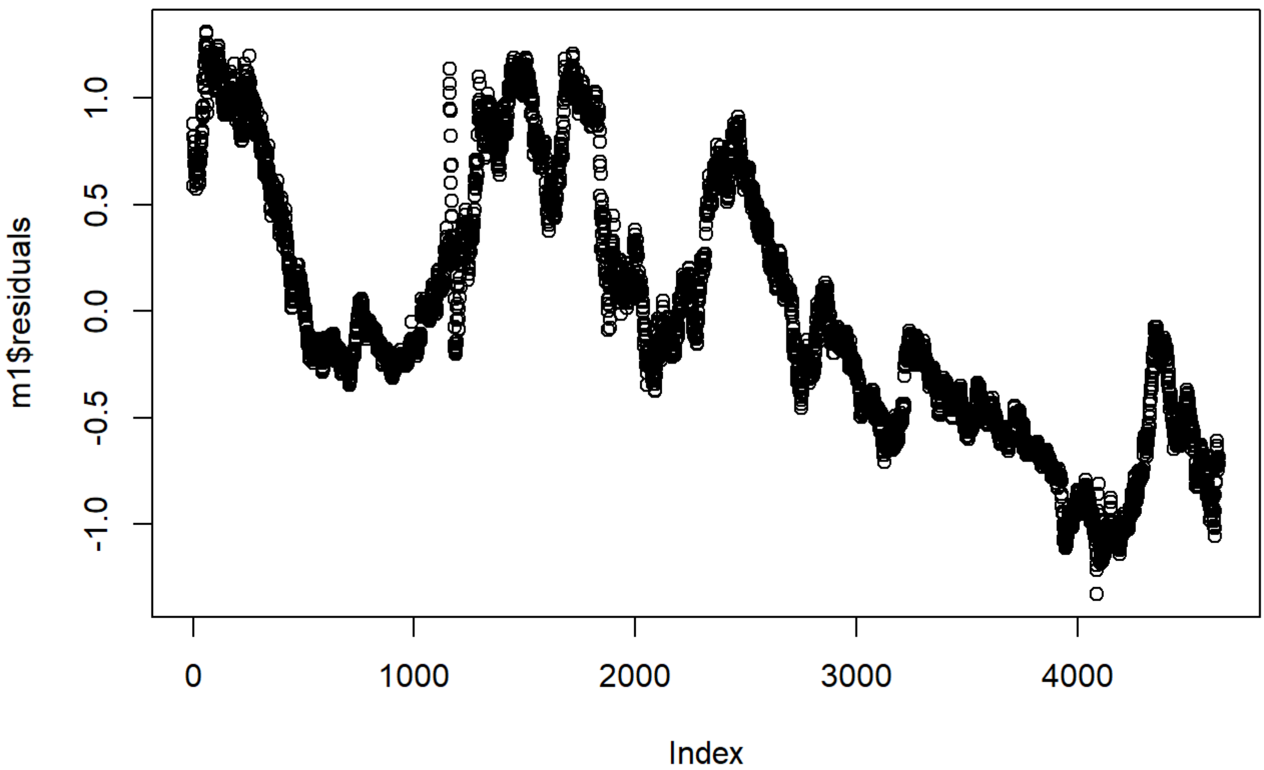
```
##        date ten_years_rate three_years_rate
## 1  2003/5/6          3.788            1.888
## 2  2003/5/7          3.678            1.816
## 3  2003/5/8          3.454            1.829
## 4  2003/5/9          3.682            1.819
## 5 2003/5/12          3.639            1.797
## 6 2003/5/13          3.608            1.796
```

a. Build a regression model using ten_years_rate as the dependent variable and three_years_rate as independent variable. Perform the goodness of fit test on the residuals of the fitted model.

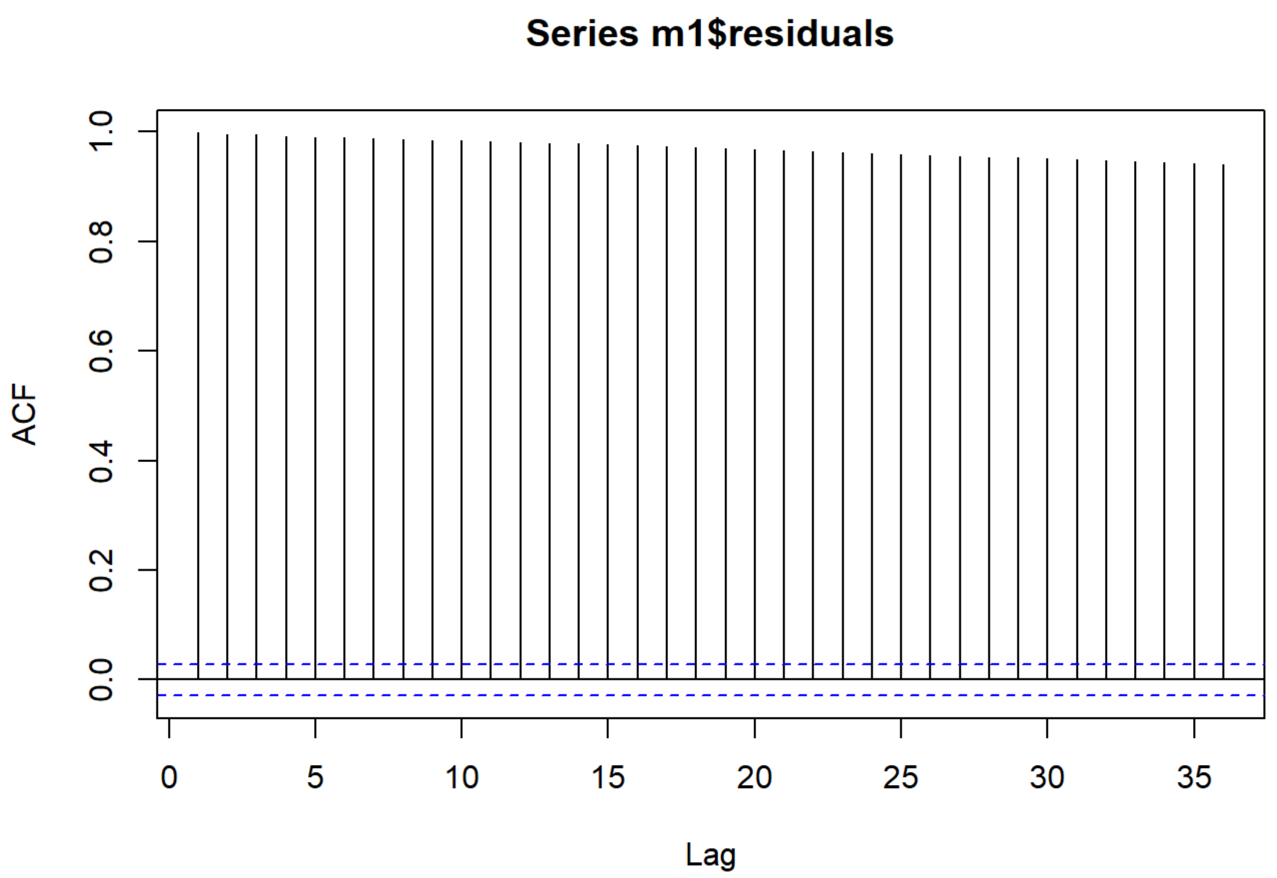
```
m1 = lm(ten_years_rate~three_years_rate,data=data_2)
summary(m1)
```

```
##
## Call:
## lm(formula = ten_years_rate ~ three_years_rate, data = data_2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3282 -0.4565 -0.1299  0.4910  1.3092
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)       1.56804    0.01442   108.7   <2e-16 ***
## three_years_rate  0.71047    0.00650   109.3   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6053 on 4638 degrees of freedom
## Multiple R-squared:  0.7203, Adjusted R-squared:  0.7203
## F-statistic: 1.195e+04 on 1 and 4638 DF,  p-value: < 2.2e-16
```

```
plot(m1$residuals)
```



```
acf(m1$residuals)
```




```
Box.test(m1$residuals,lag=12)
```

```
##
##  Box-Pierce test
##
## data:  m1$residuals
## X-squared = 54405, df = 12, p-value < 2.2e-16
```

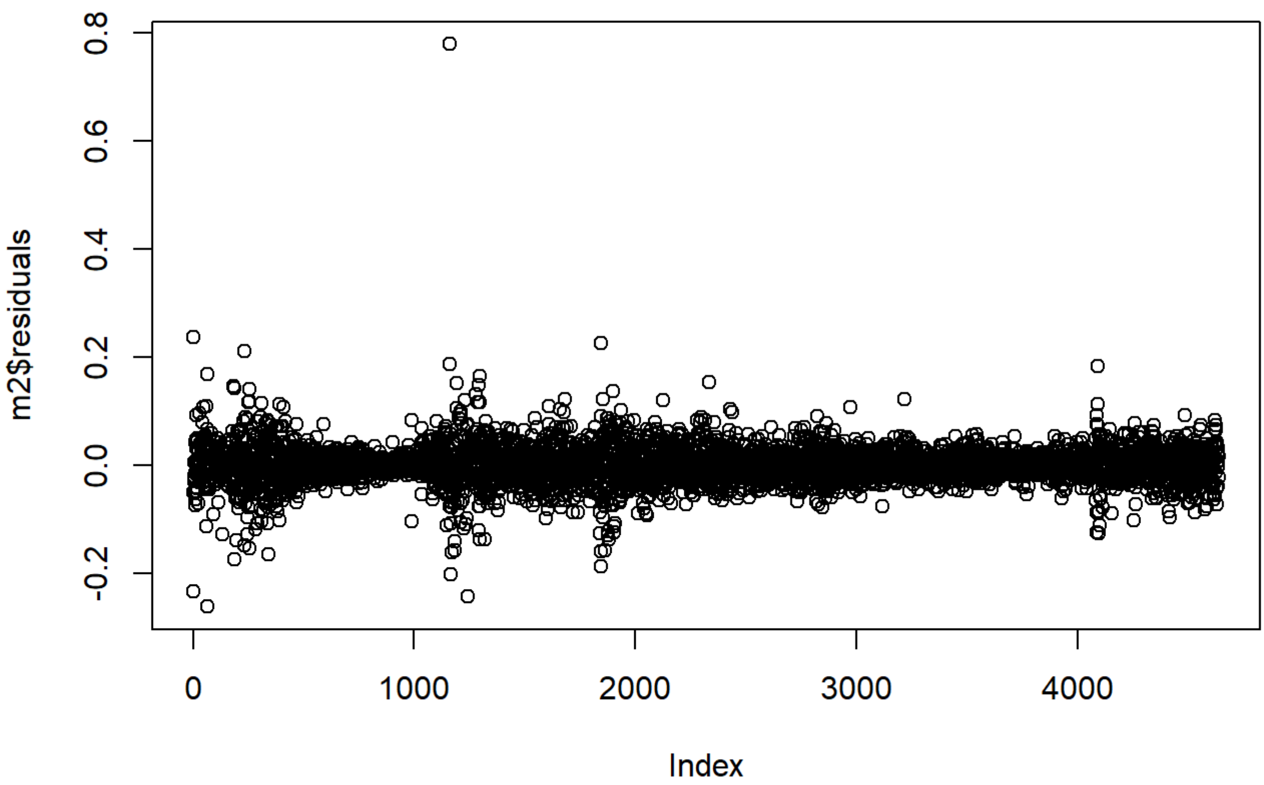
- p value<0.05，拒絕H0，序列相關，該模型不夠，(不好)

b. Build a regression model using the first difference of ten_years_rate as the dependent variable and the first difference of three_years_rate as independent variable. Perform the goodness of fit test on the residuals of the fitted regression model.

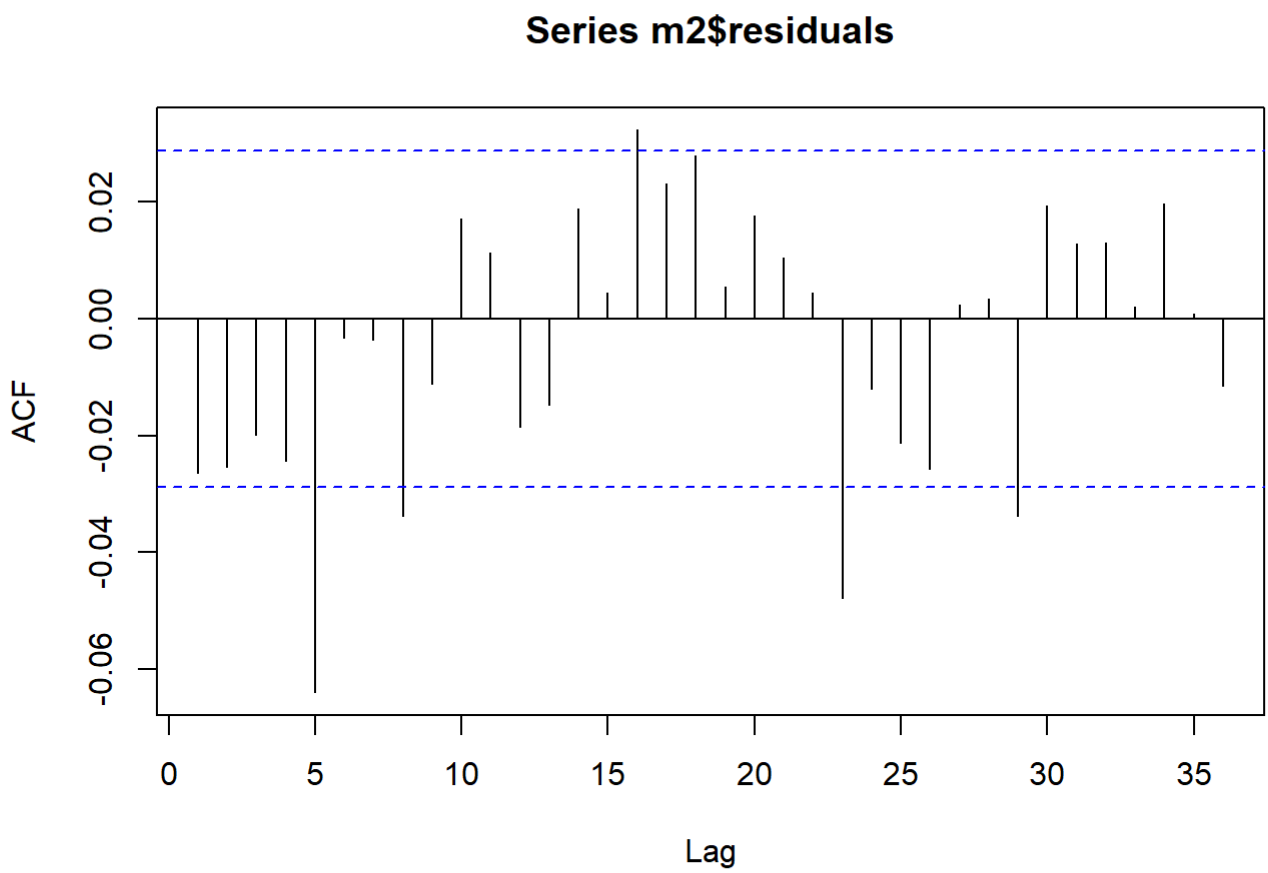
```
c10 = diff(data_2$ten_years_rate)
c3 = diff(data_2$three_years_rate)

m2 = lm(c10~c3,data=data_2)
summary(m2)
```

```
##
## Call:
## lm(formula = c10 ~ c3, data = data_2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.26182 -0.01646 -0.00046  0.01591  0.77922
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.0003607  0.0005245  -0.688    0.492
## c3           0.8179692  0.0101568  80.534   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03572 on 4637 degrees of freedom
## Multiple R-squared:  0.5831, Adjusted R-squared:  0.583
## F-statistic:  6486 on 1 and 4637 DF,  p-value: < 2.2e-16
```

```
plot(m2$residuals)
```




```
acf(m2$residuals)
```




```
Box.test(m2$residuals,lag=12)
```

```
##
##  Box-Pierce test
##
## data:  m2$residuals
## X-squared = 39.231, df = 12, p-value = 9.638e-05
```

- p value<0.05，拒絕H0，序列相關，該模型不夠，(不好)

c. Build a regression with time series error model using the first difference of ten_years_rate as the dependent variable and the first difference of three_years_rate as independent variable. Perform the goodness of fit test on the residuals of the fitted regression model

- 由ACF圖考慮error為MA(5)

```
estmodel1 = arima(c10,order = c(0,0,5),xreg = c3,include.mean = F)
estmodel1
```

```
##
## Call:
## arima(x = c10, order = c(0, 0, 5), xreg = c3, include.mean = F)
##
## Coefficients:
##           ma1      ma2      ma3      ma4      ma5    xreg
##       -0.0310  -0.0297  -0.0275  -0.0298  -0.0657  0.8179
## s.e.   0.0147   0.0147   0.0148   0.0153   0.0145  0.0101
##
## sigma^2 estimated as 0.001266:  log likelihood = 8892.29,  aic = -17772.57
```

```
rbind(estmodel1$coef-2*sqrt(diag(estmodel1$var.coef)),estmodel1$coef+2*sqrt(diag(estmodel1$var.coef)))
```

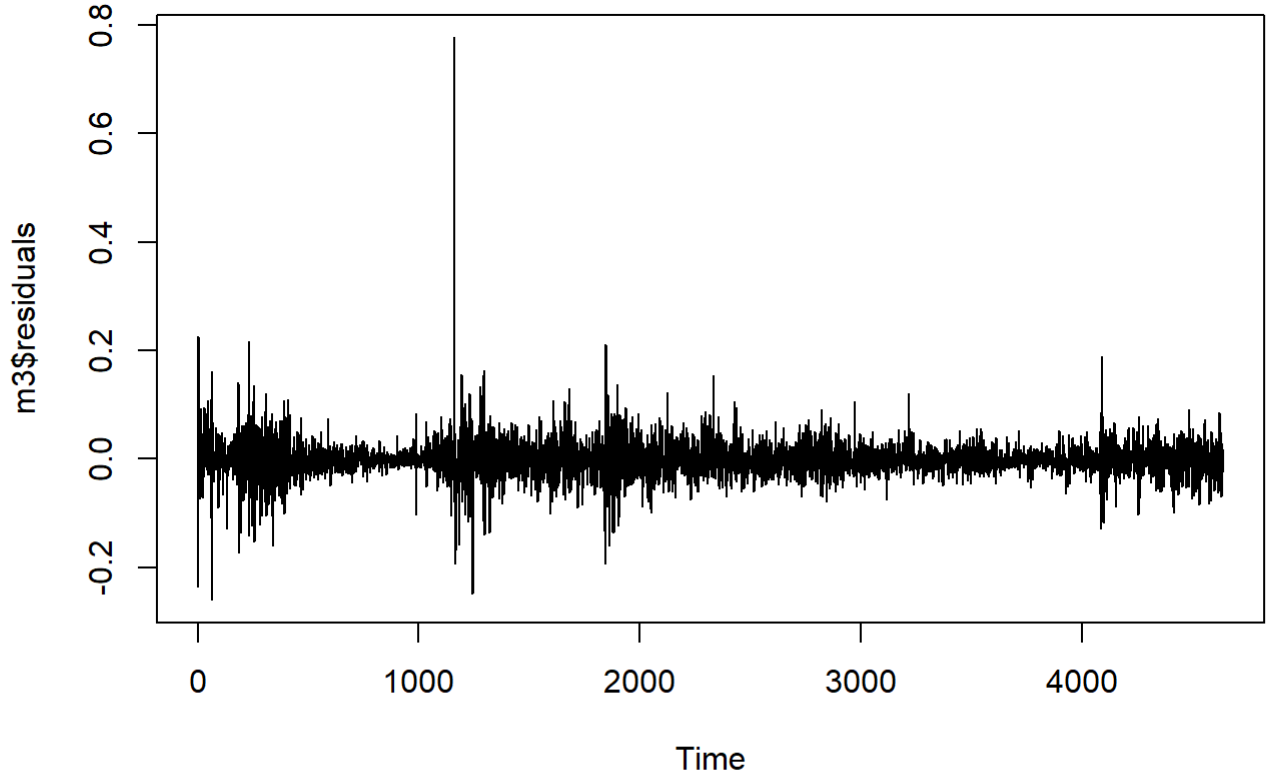
```
##               ma1           ma2          ma3          ma4         ma5
## [1,] -0.060353800 -0.0592141864 -0.057129771 -0.0604373803 -0.09465683
## [2,] -0.001670301 -0.0002464556  0.002179023  0.0009153795 -0.03671049
##           xreg
## [1,] 0.7977481
## [2,] 0.8381361
```

- 移掉 ma3 ma4

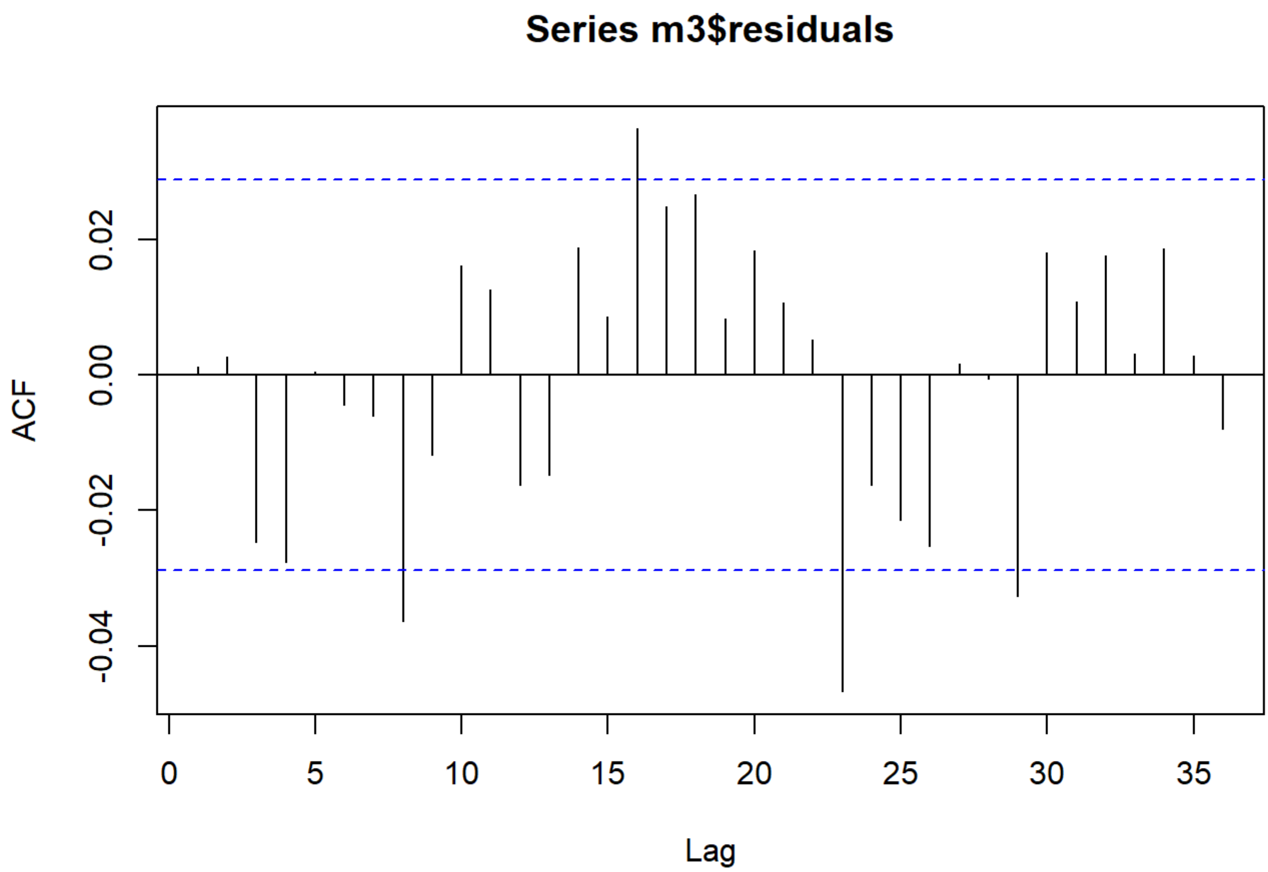
```
m3 = arima(c10,
           order = c(0,0,5),
           fixed = c(NA,NA,0,0,NA,NA),
           xreg = c3,
           include.mean = F)
m3
```

```
##
## Call:
## arima(x = c10, order = c(0, 0, 5), xreg = c3, include.mean = F, fixed = c(NA,
##     NA, 0, 0, NA, NA))
##
## Coefficients:
##           ma1      ma2  ma3  ma4      ma5    xreg
##       -0.0314  -0.0317    0    0  -0.0656  0.8164
## s.e.   0.0147   0.0152    0    0   0.0145  0.0101
##
## sigma^2 estimated as 0.001268:  log likelihood = 8888.61,  aic = -17769.21
```

```
plot(m3$residuals)
```




```
acf(m3$residuals)
```



```
Box.test(m3$residuals,lag=12, fitdf=5-2)
```

-3

```
##
##  Box-Pierce test
##
## data:  m3$residuals
## X-squared = 16.659, df = 9, p-value = 0.05434
```

- p value>0.05，不拒絕H0，序列不相關，該模型似乎足夠

d. Based on the results of (a), (b) and (c), which model would you suggest to use ? Justify your answer

- In conclusion, I suggest to use model (c). Although the R-squared of (c) is smaller than (a)，(跟(b)差不多), the acf residual of model (c) has no serial correlation

```
rsq = (sum(c10^2)-sum(m3$residuals^2))/sum(c10^2)   # R square
rsq #  0.5861534
```

```
## [1] 0.585496
```

e. Set $X_t = ten\_years\_rate_t - three\_years\_rate_t$.
Build a time series model for X_t and check it goodness of fit test on the residuals.

```
x_t = data_2$ten_years_rate - data_2$three_years_rate
adfTest(x_t, lags = 12,type="c")
```

```
##
## Title:
##  Augmented Dickey-Fuller Test
##
## Test Results:
##   PARAMETER:
##     Lag Order: 12
##   STATISTIC:
##     Dickey-Fuller: -1.1392
##   P VALUE:
##     0.635
##
## Description:
##  Sun May 14 03:53:02 2023 by user: user
```

```
adfTest(diff(x_t), lags = 12,type="c")
```

```
##
## Title:
##  Augmented Dickey-Fuller Test
##
## Test Results:
##   PARAMETER:
##     Lag Order: 12
##   STATISTIC:
##     Dickey-Fuller: -20.8531
##   P VALUE:
##     0.01
##
## Description:
##  Sun May 14 03:53:02 2023 by user: user
```

```
eacf(diff(x_t))
```

```
## AR/MA
##   0 1 2 3 4 5 6 7 8 9 10 11 12 13
## 0 o o o o x x o o x o o o  o  o  o
## 1 x o o o x o o o o o o o  o  o  o
## 2 x x o o x o o o o o o o  o  o  o
## 3 x x x o x o o o o o o o  o  o  o
## 4 x x x x x o o o o o o o  o  o  o
## 5 x x x x x o o o o o o o  o  o  o
## 6 x o x x x o o o o o o o  o  o  o
## 7 x x x o x o x o o o o o  o  o  o
```

- p value>0.05，不拒絕H0，ϕ1等於一，(有單根，要做差分)
- 考慮ARIMA(1,1,1)

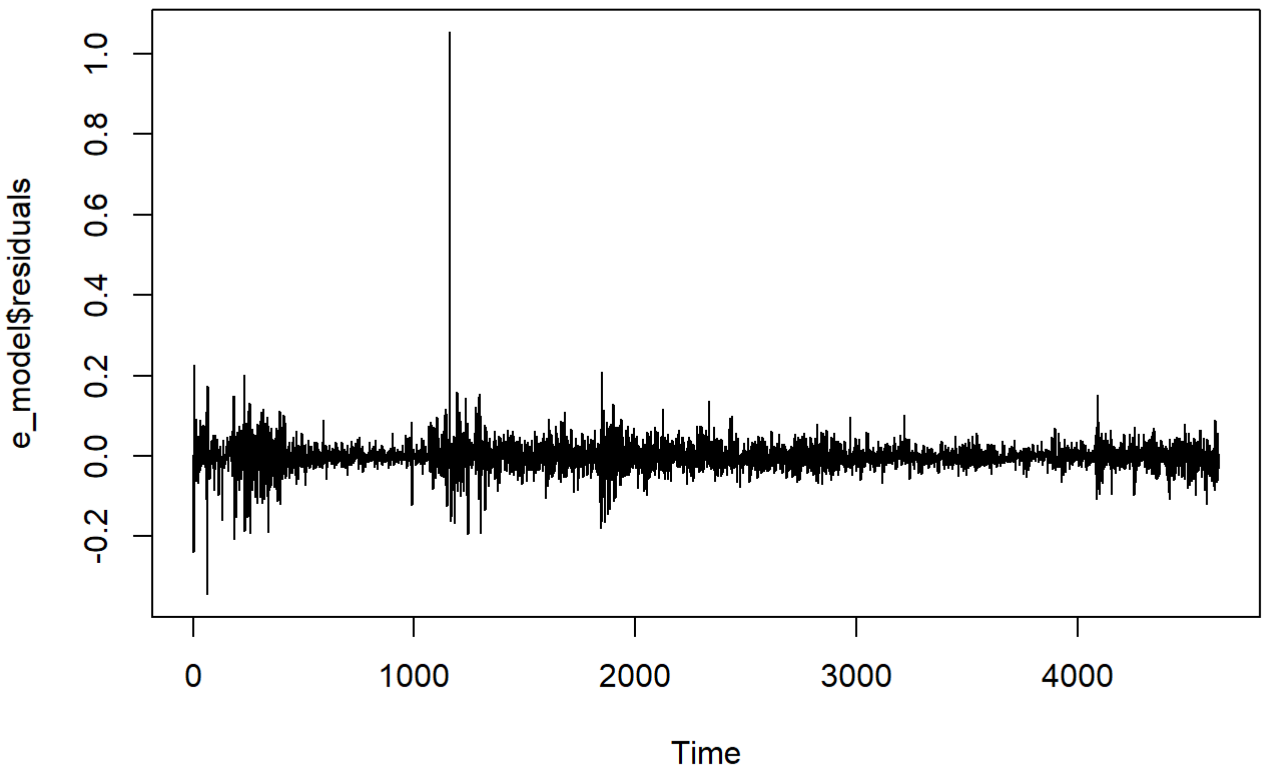
```
e_model = arima(x_t, order = c(1,1,1), method="ML")
e_model
```

```
##
## Call:
## arima(x = x_t, order = c(1, 1, 1), method = "ML")
##
## Coefficients:
##          ar1      ma1
##       0.7900  -0.8308
## s.e.  0.0476   0.0429
##
## sigma^2 estimated as 0.001358:  log likelihood = 8729.88,  aic = -17455.76
```

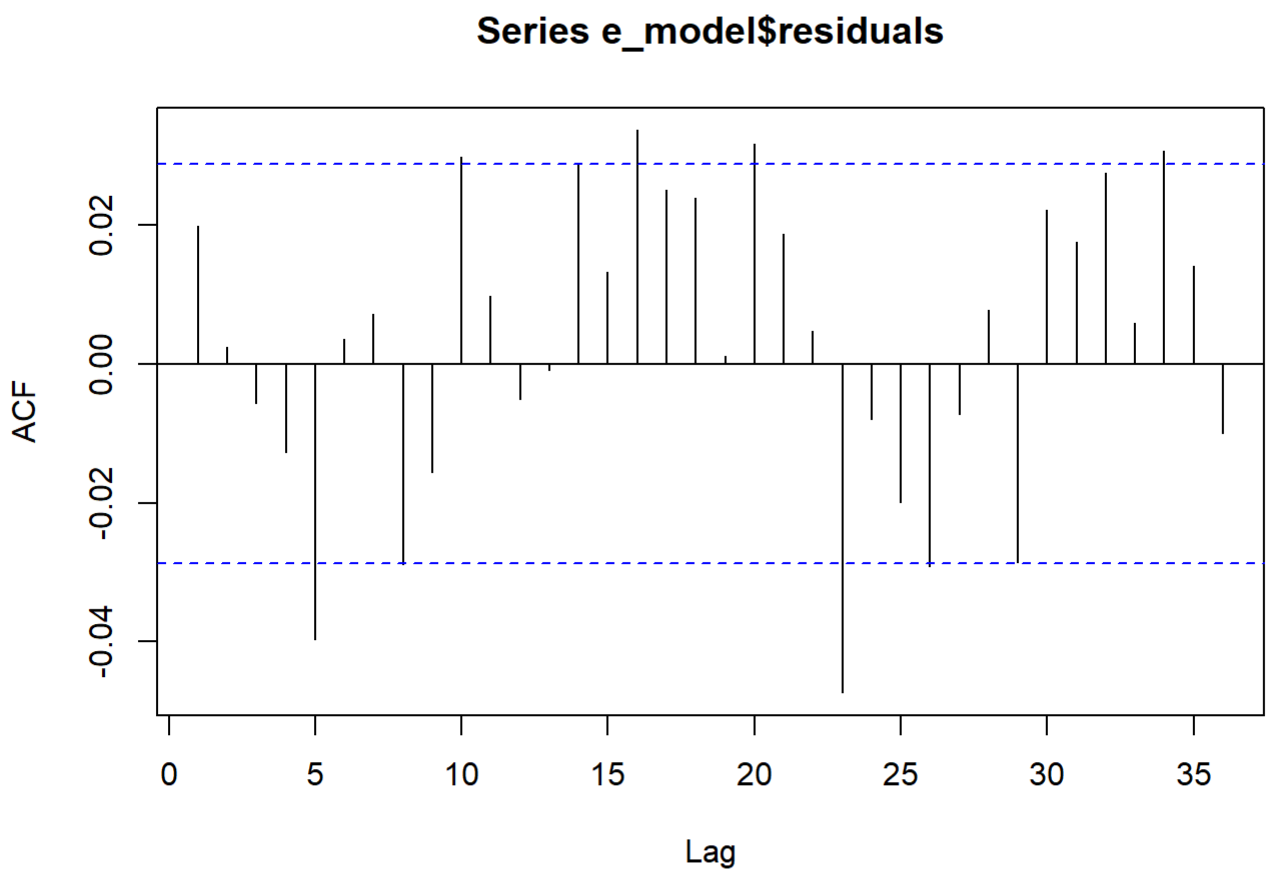
```
rbind(e_model$coef-2*sqrt(diag(e_model$var.coef)),
      e_model$coef+2*sqrt(diag(e_model$var.coef)))
```

```
##              ar1        ma1
## [1,] 0.6949216 -0.9167132
## [2,] 0.8851273 -0.7449603
```

```
plot(e_model$residuals)
```




```
acf(e_model$residuals)
```




```
Box.test(e_model$residuals, lag=7, type='Ljung-Box', fitdf=2)
```

```
##
##  Box-Ljung test
##
## data:  e_model$residuals
## X-squared = 10.396, df = 5, p-value = 0.06477
```

- p value>0.05，不拒絕H0，序列不相關，該模型似乎足夠