# STATS 509 Winter 2022, Solution to Homework 8

## 1

### (a)

$$X = |Y| = \begin{cases} 0, & \text{with prob. } \frac{1}{6} \\ 1, & \text{with prob. } \frac{1}{6} \\ 2, & \text{with prob. } \frac{2}{3} \end{cases}$$

For $X \in \{0, 1\}$, it is straightforward that

$$\begin{aligned} \mathbb{E}[Y \mid X = 0] &= 0 \cdot \mathbb{P}(Y = 0 \mid X = 0) = 0 \\ \mathbb{E}[Y \mid X = 1] &= 1 \cdot \mathbb{P}(Y = 1 \mid X = 1) = 1 \end{aligned}$$

For $X = 2$, we have

$$\begin{aligned} \mathbb{E}[Y \mid X = 2] &= -2 \cdot \mathbb{P}(Y = -2 \mid X = 2) + 2 \cdot \mathbb{P}(Y = 2 \mid X = 2) \\ &= -2 \cdot \frac{1/6}{4/6} + 2 \cdot \frac{3/6}{4/6} = 1 \end{aligned}$$

The best predictor of $Y$ based on $X$ is:

$$\hat{Y} = \mathbb{E}[Y \mid X = x] = \begin{cases} 0, & x = 0 \\ 1, & x = 1 \\ 1, & x = 2 \end{cases}$$

Thus

$$\mathrm{MSPE} = \mathbb{E}\left[(\hat{Y} - Y)^2\right] = \frac{1}{6}\left[(1 - (-2))^2 + (1 - 1)^2 + (0 - 0)^2\right] + \frac{1}{2}(1 - 2)^2 = 2.$$

### (b)

Fit

$$\hat{Y} = \mu_Y + \Sigma_{YX}\Sigma_X^{-1}\left(X - \mu_X\right).$$

Then

$$\begin{gathered} \mu_Y = \frac{1}{6}(-2 + 1 + 0) + \frac{1}{2}(2) = \frac{5}{6}, \quad \mu_X = 0 \cdot \frac{1}{6} + 1 \cdot \frac{1}{6} + 2 \cdot \frac{2}{3} = \frac{3}{2} \\ \sigma_Y^2 = \mathbb{E}\left[Y^2\right] - \mathbb{E}[Y]^2 = \frac{17}{6} - \left(\frac{5}{6}\right)^2 = \frac{77}{36}, \quad \sigma_X^2 = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = \frac{17}{6} - \left(\frac{3}{2}\right)^2 = \frac{7}{12} \\ \Sigma_{XY} = \Sigma_{YX} = \mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \frac{1}{6}(-4 + 1 + 0) + \frac{1}{2}(4) - \frac{5}{6} \cdot \frac{3}{2} = \frac{1}{4} \end{gathered}$$

Therefore the best linear linear predictor is,

$$\hat{Y} = \frac{5}{6} + \frac{1}{4}\left(\frac{7}{12}\right)^{-1}\left(X - \frac{3}{2}\right) = \frac{3}{7}X + \frac{4}{21} \quad \text{and}$$

$$\text{MSPE} = \sigma_Y^2 - \Sigma_{YX}\Sigma_X^{-1}\Sigma_{XY} = \frac{77}{36} - \left(\frac{1}{4}\right)^2\left(\frac{12}{7}\right) = \frac{128}{63}$$

## (c)

Best linear predictor assumes that $Y$ and $X$ has a linear relationship while the best predictor does not. The MSPE in (b) is larger than than of (a) because the class of predictors considered in (a) include linear and nonlinear, and is the best possible linear/nonlinear function in terms of minimizing MSPE. By only considering the best linear predictor in (b), it is clear that MSPE from (a) will always be less than that from (b).

# 2

## (a)

Let $Z_t = Y_t - \mu$, then $Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \epsilon_t$, for $\forall k > 0$ :

$$\begin{aligned}\gamma_Z(k) &= \operatorname{Cov}(Z_t, Z_{t-k}) \\ &= \operatorname{Cov}(\phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \epsilon_t, Z_{t-k}) \\ &= \phi_1 \operatorname{Cov}(Z_{t-1}, Z_{t-k}) + \phi_2 \operatorname{Cov}(Z_{t-2}, Z_{t-k}) + \operatorname{Cov}(\epsilon_t, Z_{t-k})\end{aligned}$$

But note that

$$\operatorname{Cov}(\epsilon_t, Z_{t-k}) = \operatorname{Cov}(\epsilon_t, Y_{t-k} - \mu) = \operatorname{Cov}(\epsilon_t, Y_{t-k}) = 0$$

which implies

$$\gamma_\varepsilon(k) = \phi_1 \gamma_\varepsilon(k-1) + \phi_2 \gamma_\varepsilon(k-2)$$

Since

$$\gamma_Z(k) = \operatorname{Cov}(Z_t, Z_{t-k}) = \operatorname{Cov}(Y_t - \mu, Y_{t-k} - \mu) = \operatorname{Cov}(Y_t, Y_{t-k}) = \gamma_Y(k)$$

this implies that

$$\gamma_Y(k) = \phi_1 \gamma_Y(k-1) + \phi_2 \gamma_Y(k-2)$$

which implies

$$\frac{\gamma_Y(k)}{\gamma_Y(0)} = \phi_1 \frac{\gamma_Y(k-1)}{\gamma_Y(0)} + \phi_2 \frac{\gamma_Y(k-2)}{\gamma_Y(0)}$$

and hence

$$\rho_Y(k) = \phi_1 \rho_Y(k-1) + \phi_2 \rho_Y(k-2)$$

## (b)

When $k = 1$ : $\quad \rho_Y(1) = \phi_1 \rho_Y(0) + \phi_2 \rho_Y(-1)$
When $k = 2$ : $\quad \rho_Y(2) = \phi_1 \rho_Y(1) + \phi_2 \rho_Y(0)$
Since then:

$$\rho_Y(-1) = \operatorname{Cor}(Y_t, Y_{t+1}) = \operatorname{Cor}\left(Y_{t+1}, Y_{(t+1)-1}\right) = \rho_Y(1)$$

$$\begin{cases} \rho_Y(1) = \phi_1 \rho_Y(0) + \phi_2 \rho_Y(1) \\ \rho_Y(2) = \phi_1 \rho_Y(1) + \phi_2 \rho_Y(0) \end{cases}$$

$$\Rightarrow \quad \begin{pmatrix} \rho_Y(1) \\ \rho_Y(2) \end{pmatrix} = \begin{pmatrix} \rho_Y(0) & \rho_Y(1) \\ \rho_Y(1) & \rho_Y(0) \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}$$

$$\rho_Y(0) = \frac{\gamma_Y(0)}{\gamma_Y(0)} = 1$$

$$\Rightarrow \quad \begin{pmatrix} \rho_Y(1) \\ \rho_Y(2) \end{pmatrix} = \begin{pmatrix} 1 & \rho_Y(1) \\ \rho_Y(1) & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}.$$

### (c)

Since $\rho(1) = 0.3, \rho(2) = 0.2$

$$\begin{pmatrix} 0.3 \\ 0.2 \end{pmatrix} = \begin{pmatrix} 1 & 0.3 \\ 0.3 & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}$$

$$\Rightarrow \quad \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} 1 & 0.3 \\ 0.3 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0.3 \\ 0.2 \end{pmatrix} = \begin{pmatrix} 24/91 \\ 11/91 \end{pmatrix}$$

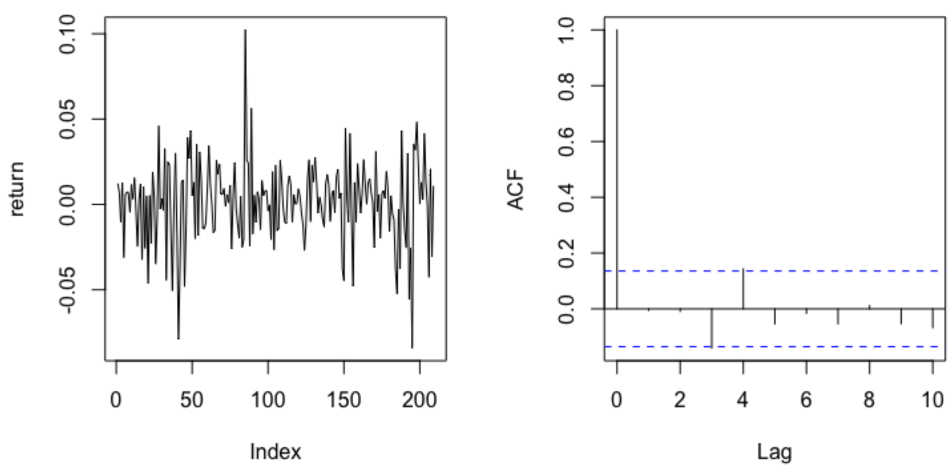
We have

$$\rho(3) = \phi_1 \rho(2) + \phi_2 \rho(1) = \frac{24}{91} \times \frac{1}{5} + \frac{11}{91} \times \frac{3}{10} = \frac{81}{910}.$$

## 3

### (a)

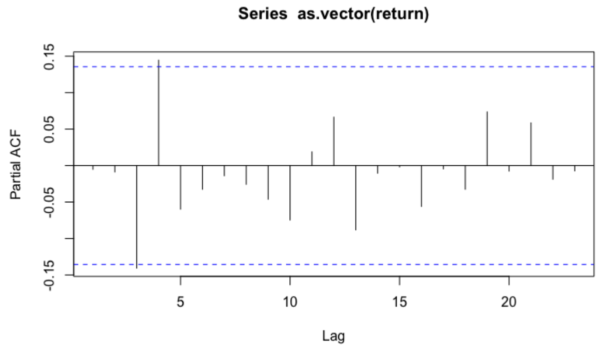
```
> par(mfrow=c(1,2))
> data = read.csv('RUT_03_2015-03_2019.csv')
> return = diff(data$Adj.Close)/data$Adj.Close[-length(data$Adj.Close)]
> plot(return, type = 'l')
> acf(as.vector(return), lag=10, main = '')
```



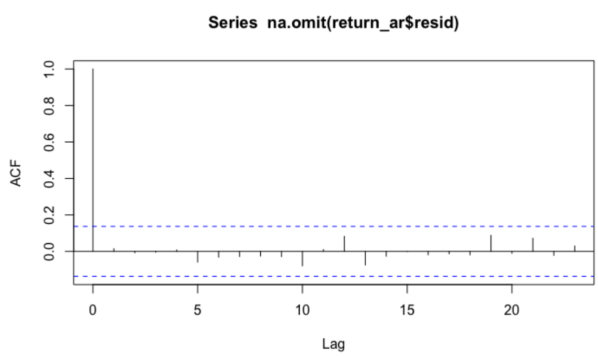

```
> Box.test(return, lag=10, type='Ljung')

Box-Ljung test

data:  return
X-squared = 11.605, df = 10, p-value = 0.3124
```

### (b)

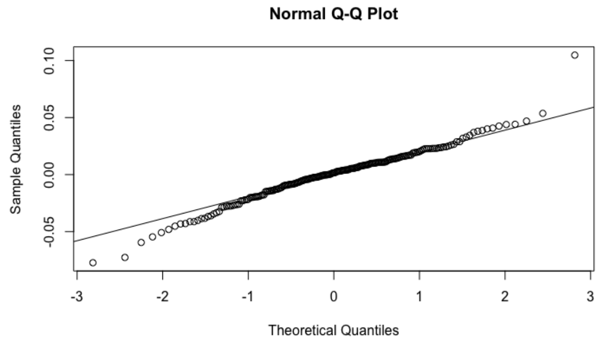
```
> pacf(as.vector(return))
```




```
> return_ar = ar(return, aic = TRUE, method = c("mle"))
> return_ar

Call:
ar(x = return, aic = TRUE, method = c("mle"))

Coefficients:
      1        2        3        4
 0.0144  -0.0050  -0.1422   0.1489

Order selected 4  sigma^2 estimated as  0.0005315

> acf(na.omit(return_ar$resid))
```

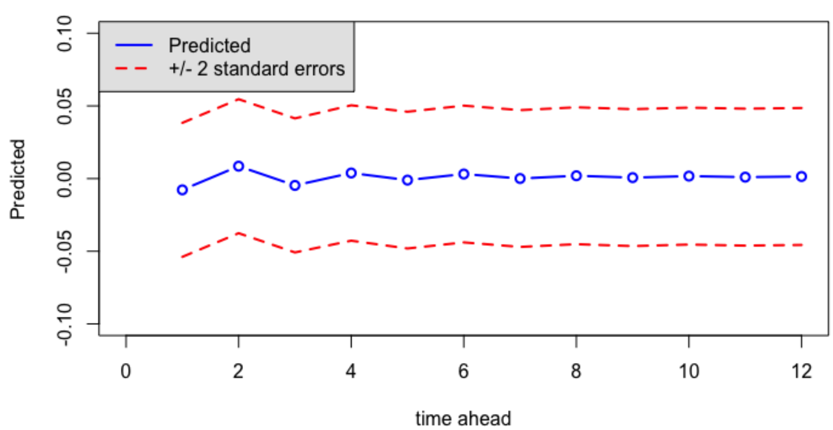



```
> qqnorm(return_ar$resid)
> qqline(return_ar$resid)
```




### (c)

```
> acf(na.omit(return_ar$resid))
> qqnorm(return_ar$resid)
> qqline(return_ar$resid)
> # c
> forecasts = predict(return_ar,newdata=return[206:209],n.ahead = 12, se.fit = TRUE)
> plot(c(1:12),forecasts$pred,type='b',lty=1,xlab='time ahead',
+       ylab='Predicted',xlim=c(0,12),ylim=c(-0.1,0.1),lwd=2,col='blue')
> lines(c(1:12),forecasts$pred+2*forecasts$se,lty=2,lwd=2,col='red')
> lines(c(1:12),forecasts$pred-2*forecasts$se,lty=2,lwd=2,col='red')
> legend('topleft', c("Predicted","+/- 2 standard errors"), lty=c(1,2),
+         lwd=c(2,2),col=c("blue","red"), bg="gray90")
```

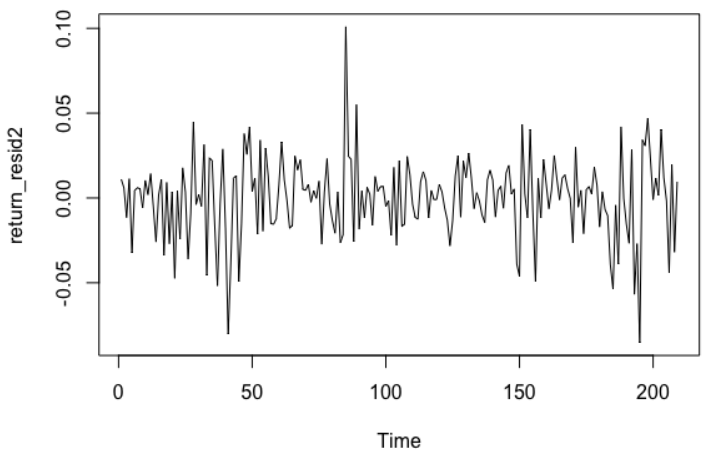



```
> show(cbind(forecasts$pred+2*forecasts$se, forecasts$pred-2*forecasts$se))
Time Series:
Start = 5
End = 16
```

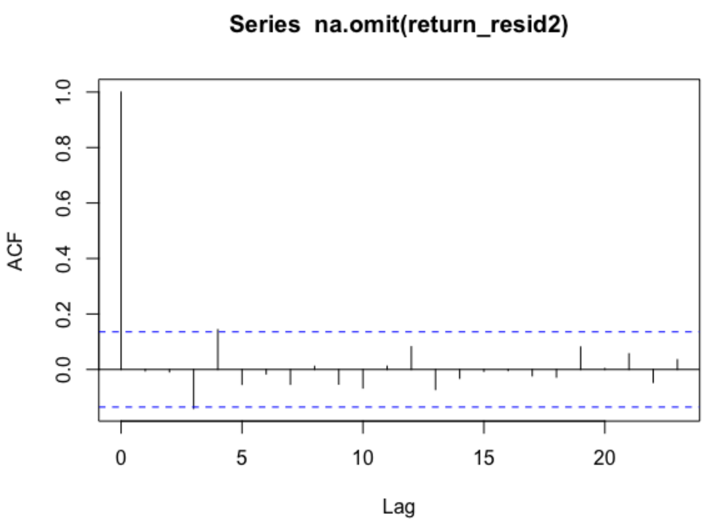
```
Frequency = 1
   forecasts$pred + 2 * forecasts$se forecasts$pred - 2 * forecasts$se
 5                        0.03834824                       -0.05386721
 6                        0.05464058                       -0.03758449
 7                        0.04140758                       -0.05081854
 8                        0.05042102                       -0.04273434
 9                        0.04600227                       -0.04810566
10                        0.05019697                       -0.04391239
11                        0.04710367                       -0.04702181
12                        0.04904609                       -0.04515732
13                        0.04777280                       -0.04644880
14                        0.04881040                       -0.04541129
15                        0.04810093                       -0.04612410
16                        0.04855309                       -0.04567552
```

**(d)**

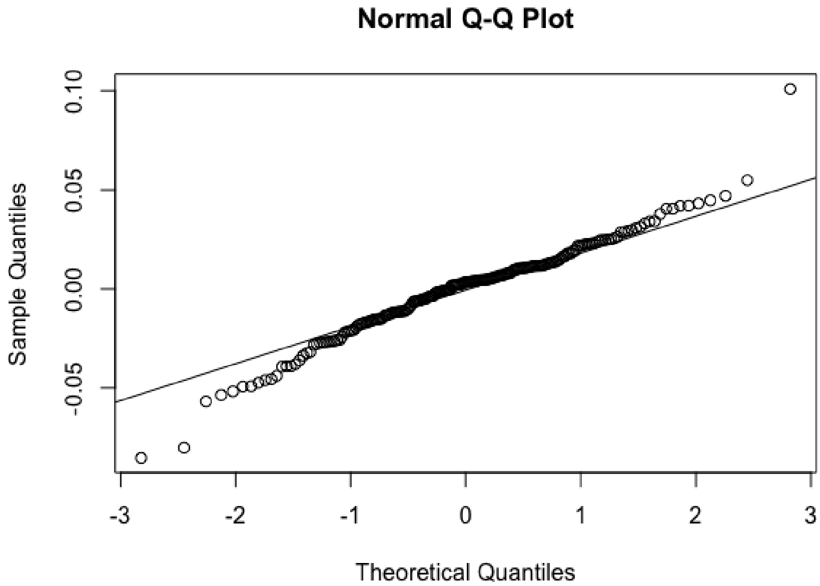
```
> library(forecast)
> auto.arima(return, ic=c("aic"))
Series: return
ARIMA(0,0,0) with zero mean

sigma^2 estimated as 0.0005564:  log likelihood=486.57
AIC=-971.15   AICc=-971.13   BIC=-967.8
> return_arma = arima(return, order = c(0,0,0), seasonal = list(order = c(0,0,0), period = NA)
+                      xreg = NULL, include.mean = TRUE, transform.pars = TRUE, fixed = NULL, in
+                      method = c("CSS-ML","ML","CSS"))
> # check residuals:
> return_resid2 = return_arma$resid
> plot(return_resid2, type = "l")
```




```
> acf(na.omit(return_resid2))
```




```
> qqnorm(return_resid2)
> qqline(return_resid2)
```

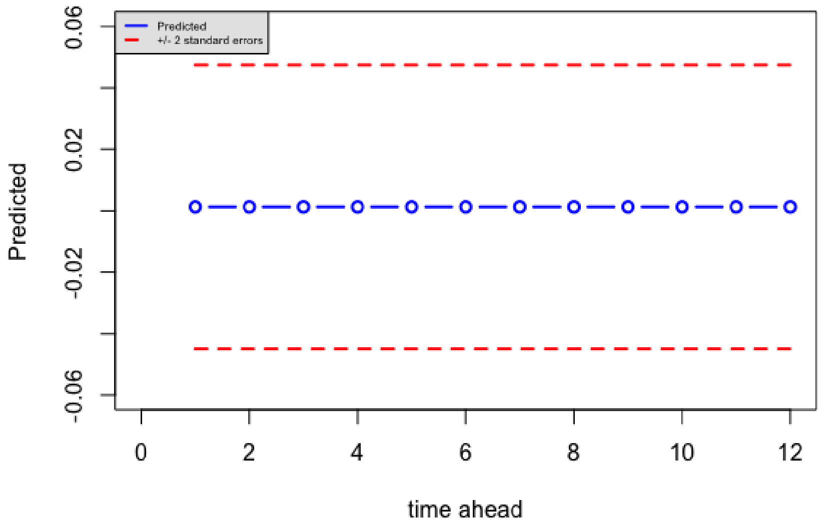



```
> show(c(mean(na.omit(return_resid2)), var(na.omit(return_resid2))))
[1] -3.627259e-15  5.574523e-04
```

From the ACF plot, on can see that the residuals are uncorrleated with lags. In the Q-Q plot, residuals seem to have heavy tails, which indicates that the noises might not exactly come from normal white noise. Mean and var of residuals are -3.626687e-15 and 5.574523e-04, Both of which are quite small.

### e

```
> xnew2 = return[204:209]
> forecasts2 <- predict(return_arma, newdata = xnew2, n.ahead = 12, se.fit = TRUE)
> plot(c(1:12), forecasts2$pred, type='b', lty=1, xlab='time ahead', ylab='Predicted',
   xlim=c(0,12), ylim=c(-0.06,0.06), lwd=2, col='blue')
> lines(c(1:12),forecasts2$pred+1.96*forecasts2$se,lty=2,lwd=2,col='red')
> lines(c(1:12),forecasts2$pred-1.96*forecasts2$se,lty=2,lwd=2,col='red')
> legend('topleft', c("Predicted","+/- 2 standard errors"), lty=c(1,2),
   lwd=c(2,2), col=c("blue","red"), bg="gray90", cex = 0.5)
```




```
> confidence_interval_lower_bound = forecasts2$pred - 1.96*forecasts2$se
> confidence_interval_upper_bound = forecasts2$pred + 1.96*forecasts2$se
> confidence_interval = cbind(confidence_interval_lower_bound,
                     confidence_interval_upper_bound)
> show(confidence_interval)
Time Series:
Start = 210
End = 221
Frequency = 1
    confidence_interval_lower_bound confidence_interval_upper_bound
210                      -0.0449072                        0.047424
211                      -0.0449072                        0.047424
212                      -0.0449072                        0.047424
213                      -0.0449072                        0.047424
214                      -0.0449072                        0.047424
215                      -0.0449072                        0.047424
216                      -0.0449072                        0.047424
217                      -0.0449072                        0.047424
218                      -0.0449072                        0.047424
219                      -0.0449072                        0.047424
220                      -0.0449072                        0.047424
221                      -0.0449072                        0.047424
```