STATS 509 HOMEWORK 8

Yuan Yin

Question 2 (a)

(i)

```
library(Ecdat)
data(CRSPday)
crsp=CRSPday[,7]
ar1 = arima(crsp,order=c(1,0,0)); ar1
```

```
##
## Call:
## arima(x = crsp, order = c(1, 0, 0))
##
## Coefficients:
##          ar1  intercept
##       0.0853      7e-04
## s.e.  0.0198      2e-04
##
## sigma^2 estimated as 5.973e-05:  log likelihood = 8706.18,  aic = -17406.37
```

```
ar2 = arima(crsp,order=c(2,0,0)); ar2
```

```
##
## Call:
## arima(x = crsp, order = c(2, 0, 0))
##
## Coefficients:
##          ar1      ar2  intercept
##       0.0865  -0.0141      7e-04
## s.e.  0.0199   0.0199      2e-04
##
## sigma^2 estimated as 5.972e-05:  log likelihood = 8706.43,  aic = -17404.87
```

- I would prefer AR(1) model since we can see that the coefficient of ϕ_2 in AR(2) is -0.014 and the standard error is 0.02, which means that $\phi_2 = 0$ is in the 95% confidence interval, this coefficient is not significant. Since AR(1) is a simpler model, so we prefer AR(1). What's more, AR(1) has smaller AIC, it corresponds with our result.

(ii)

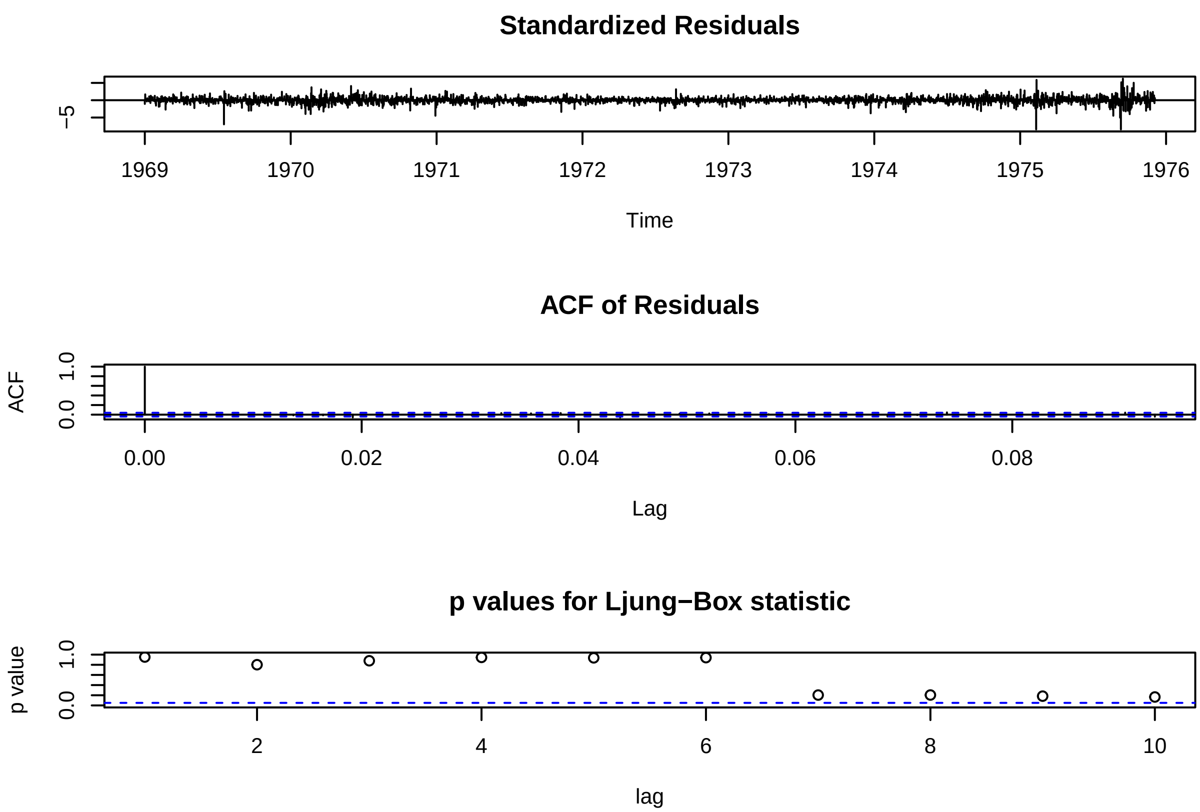
```
left = ar1$coef[1]-1.96*sqrt(ar1$var.coef[1,1])
right = ar1$coef[1]+1.96*sqrt(ar1$var.coef[1,1])
c(left, right)
```

```
##       ar1       ar1
## 0.0464615 0.1241421
```

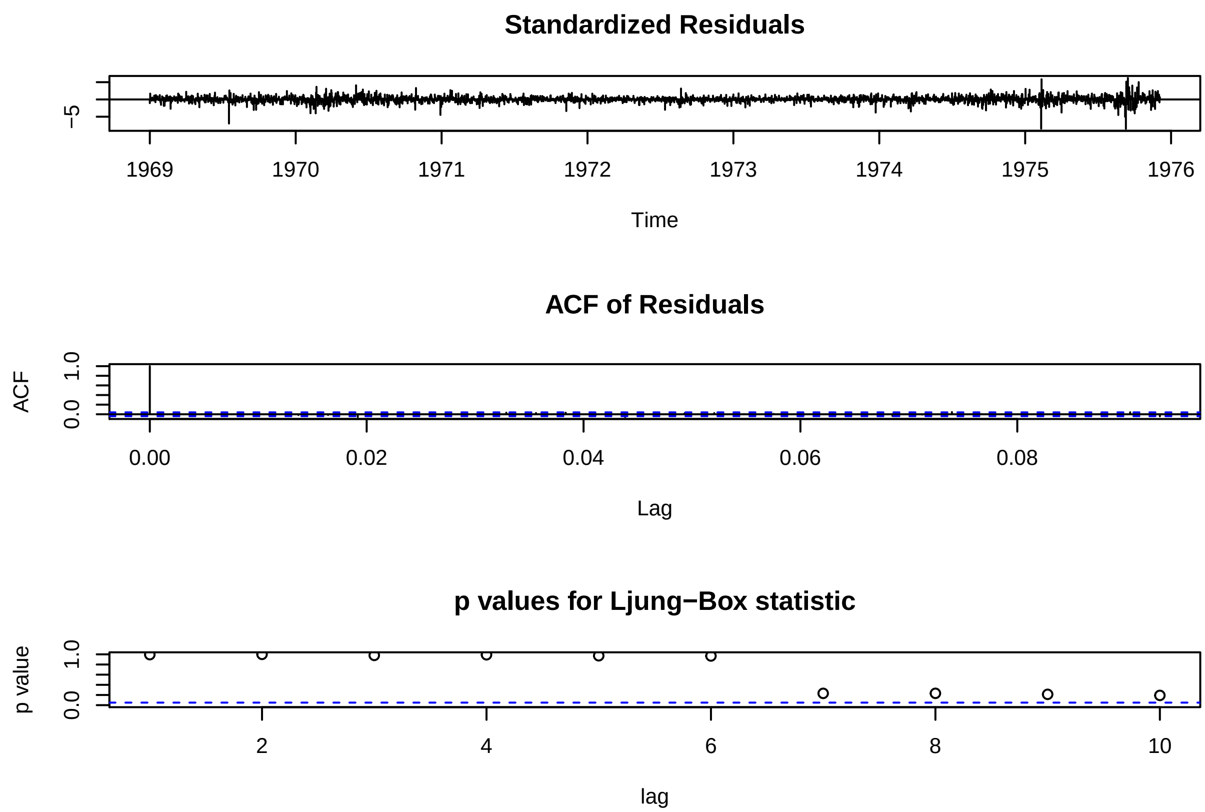
- Seeing that the 95% interval of ϕ in AR(1) is $0.085 \pm 1.96 * SE = [0.0465, 0.1241]$.

(iii)

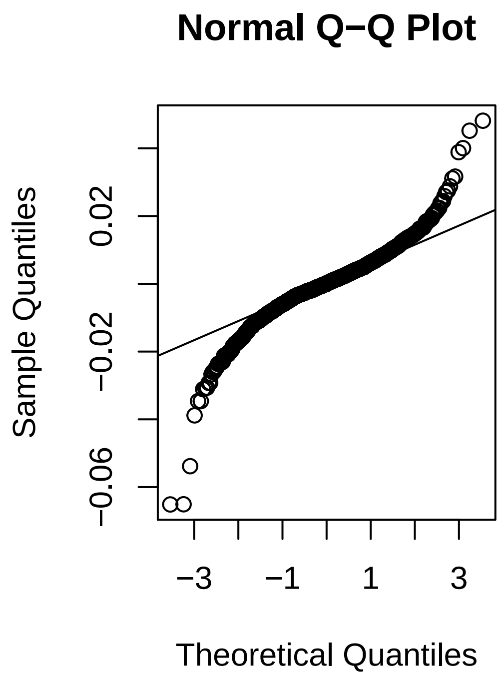
```r
# Diagnostics plots from arima
tsdiag(ar1)
```

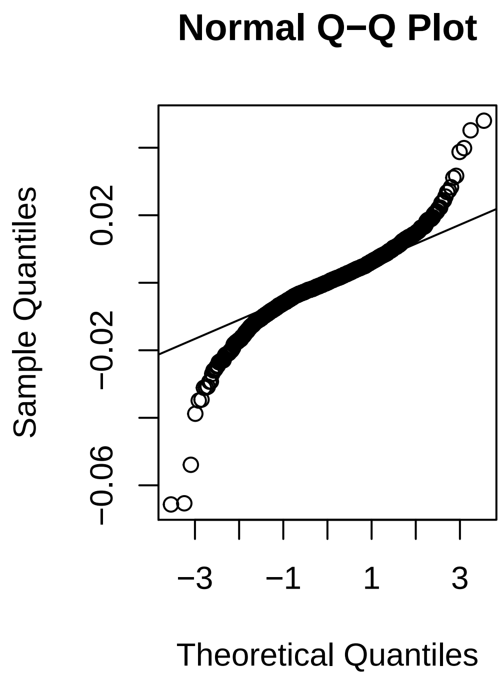



```r
tsdiag(ar2)
```




```
## QQ plots of residuals
par(mfrow = c(1,2))
qqnorm(ar1$resid); qqline(ar1$resid)
qqnorm(ar2$resid); qqline(ar1$resid)
```





Seeing that both models fail to reject that residuals are white noise, and both QQ-plots show that residuals have heavy tails. This indicates that we can't tell differences between two models. For simplifying our models, we should choose AR(1). This doesn't change our answer

Question 2 (b)

```
library(forecast)
auto.arima(crsp, start.p=0, max.p = 4, max.q = 4, seasonal = FALSE, ic = "aic")
```

```
## Series: crsp
## ARIMA(0,0,1) with non-zero mean
##
## Coefficients:
##          ma1   mean
##       0.0869  7e-04
## s.e.  0.0199  2e-04
##
## sigma^2 estimated as 5.977e-05:  log likelihood=8706.36
## AIC=-17406.73   AICc=-17406.72   BIC=-17389.22
```

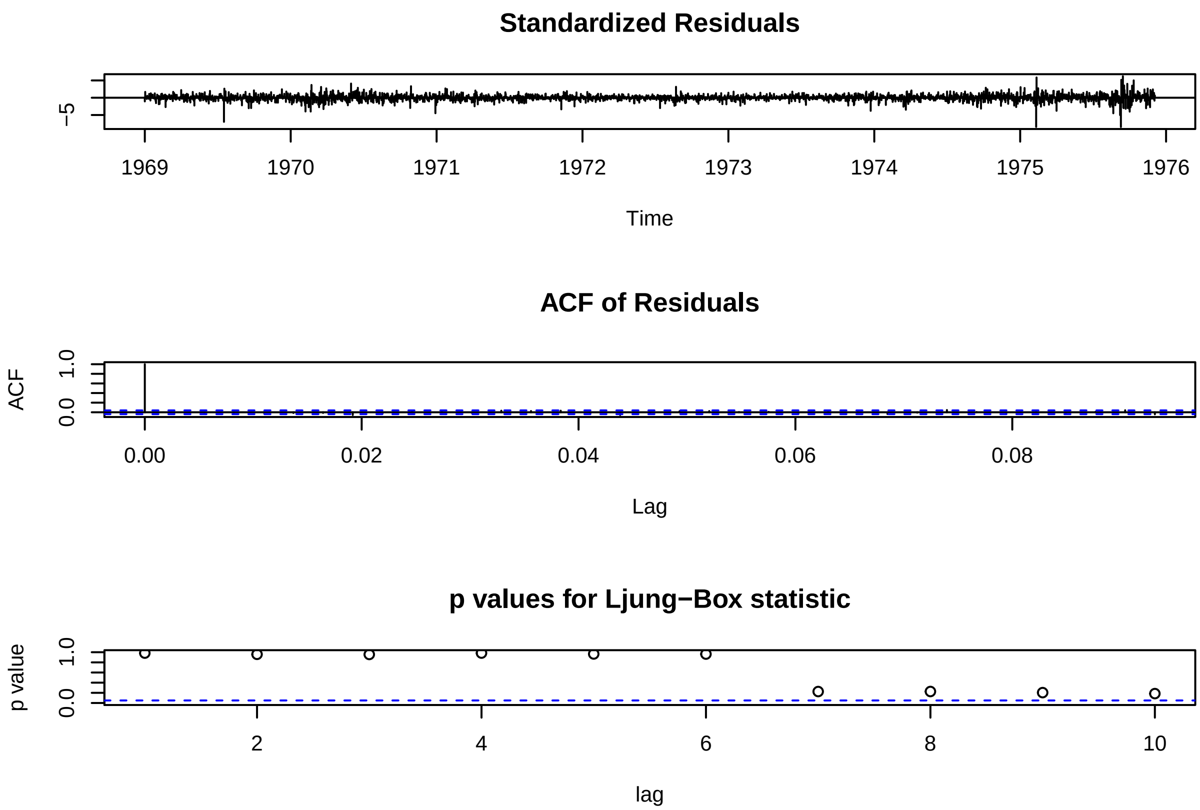
Seeing that MA(1) has the smallest AIC, let's make diagnostics on MA(1).

```
ma1 = arima(crsp,order=c(0,0,1)); ma1
```

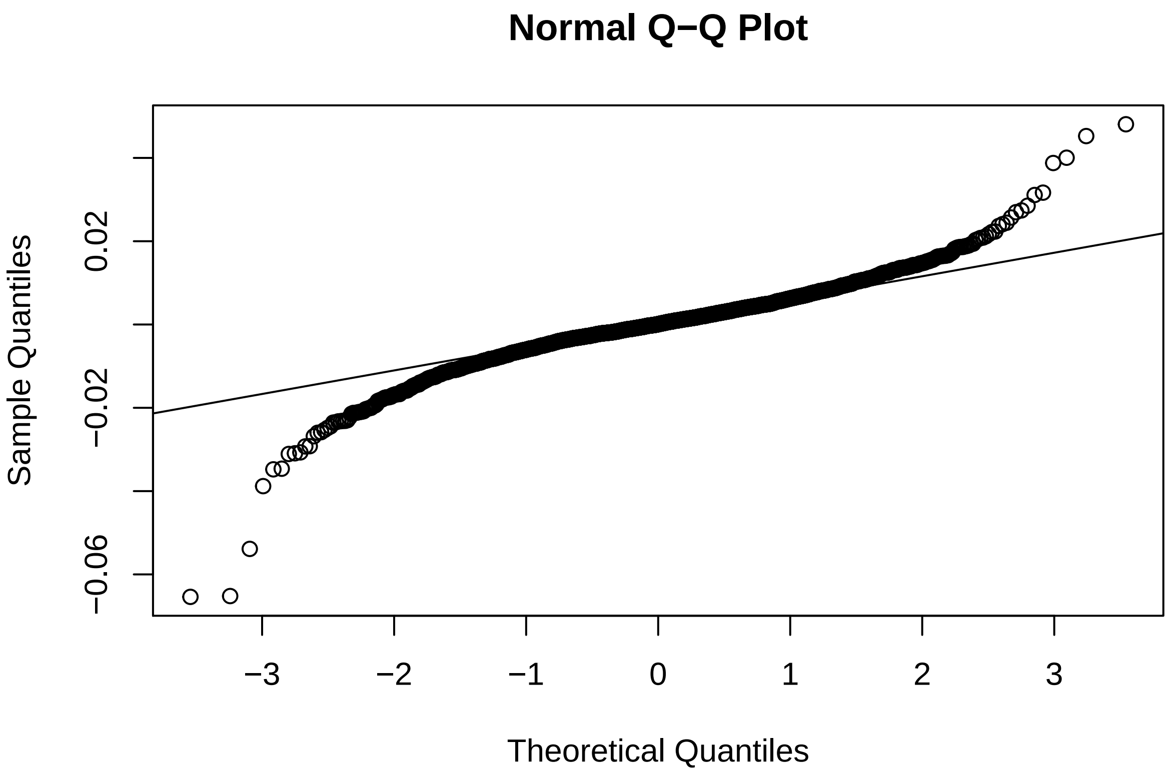
```
##
## Call:
## arima(x = crsp, order = c(0, 0, 1))
##
## Coefficients:
```

```
##          ma1  intercept
##       0.0869      7e-04
## s.e.  0.0199      2e-04
##
## sigma^2 estimated as 5.972e-05:  log likelihood = 8706.36,  aic = -17406.73
```

```r
tsdiag(ma1)
```




```r
qqnorm(ma1$resid); qqline(ma1$resid)
```



- Finding that with similar results with AR(1) and AR(2), we still fail to reject the residuals to be white noise, but from QQ-plot we can see the distribution is not normal since its heavy tails.
- Since MA(1) is the best AMRA model to fit the data, but there is not much difference between MA(1) and AR(1), maybe AMRA model is not a good model to fit the data, or maybe the residuals are nor normal distribution white noise but other type of white noise.