STAT 443: Lab 7

Wenxuan Zan (61336194)

05 March, 2023

Question 1

```
data <- read.csv("lab7data.csv",header = TRUE)
annual_ts <- ts(data$Annual, start = c(1919), end = c(2008))
ar1model <- arima(annual_ts, order = c(1,0,0),include.mean = TRUE)
ar1model
```

```
##
## Call:
## arima(x = annual_ts, order = c(1, 0, 0), include.mean = TRUE)
##
## Coefficients:
##          ar1  intercept
##       0.5843    -1.9591
## s.e.  0.0864     0.2810
##
## sigma^2 estimated as 1.265:  log likelihood = -138.49,  aic = 282.99
```

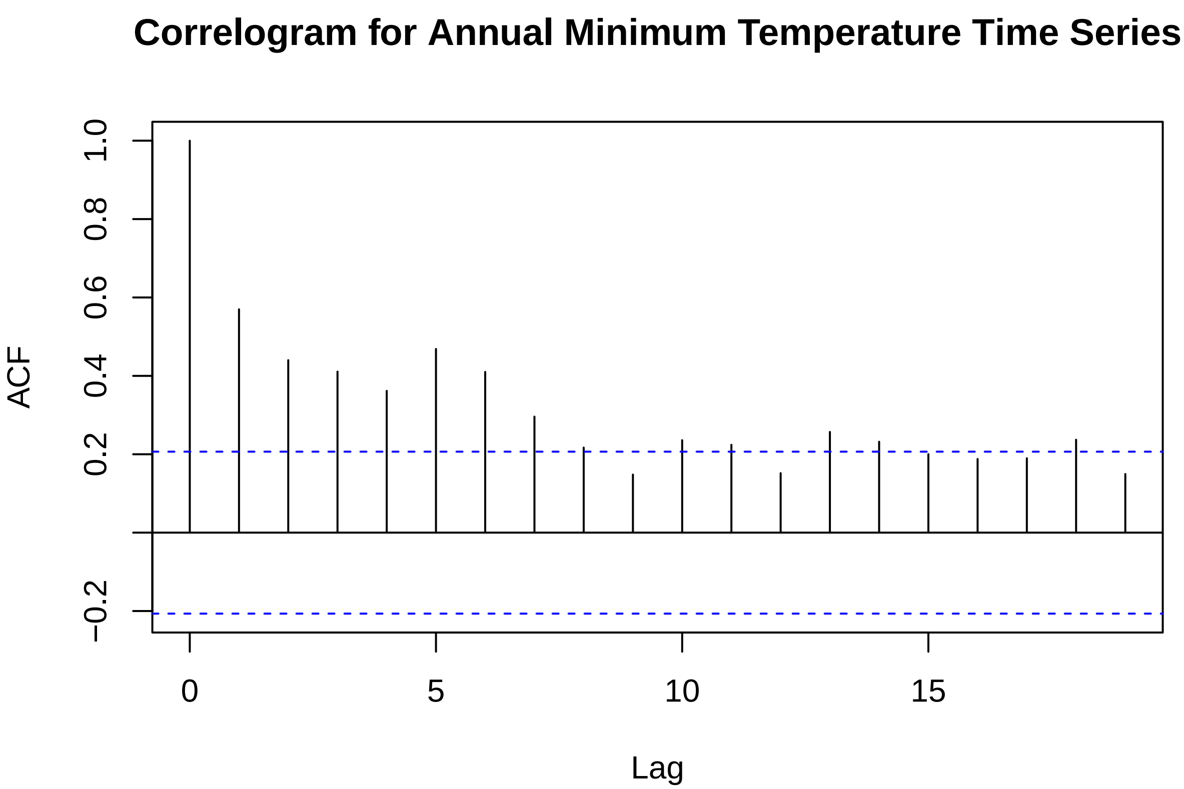
The fitted model is

$$X_t - \hat{\mu} = 0.5843(X_{t-1} - \hat{\mu}) + Z_t$$

where $\hat{\mu} = -1.9591$ and $Z_t \sim WN(0, 1.265)$

Question 2

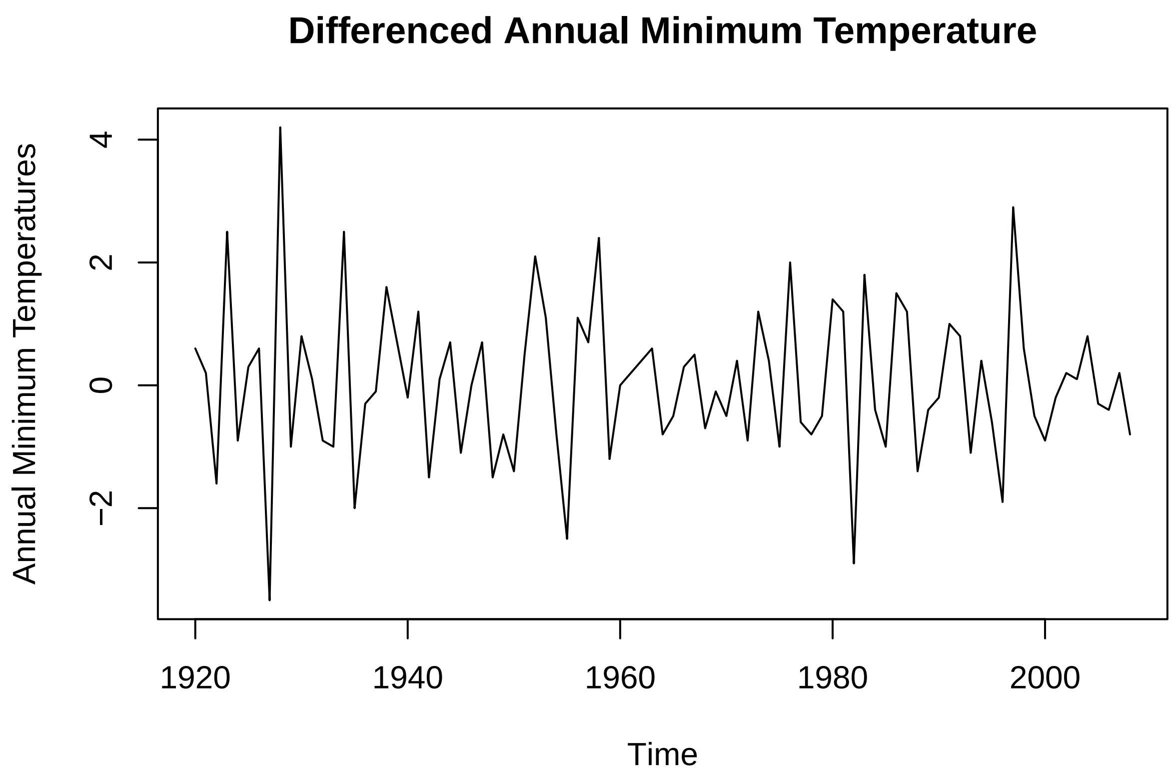
```
acf(annual_ts,
    main = "Correlogram for Annual Minimum Temperature Time Series")
```



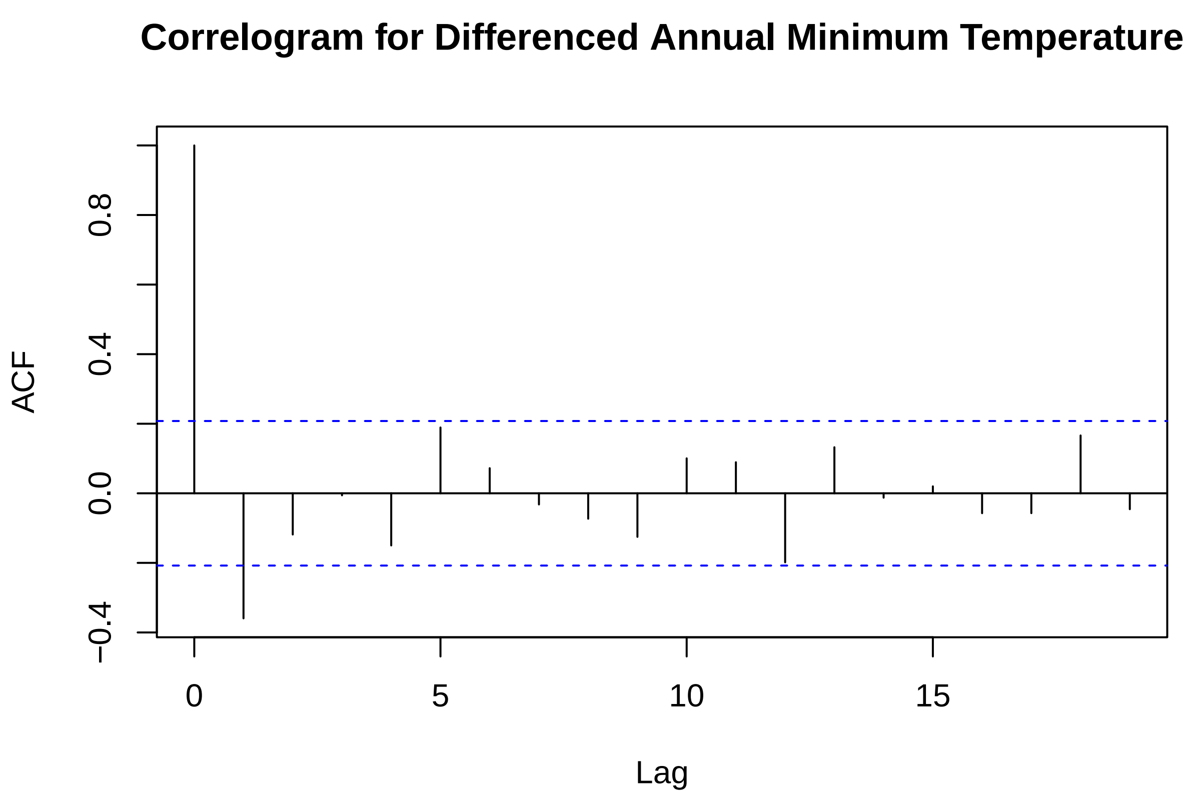
We would expect acf values for an AR(1) process to decay exponentially, but the current acf values for the temperature data first show a exponential decay until lag 5 and has acf values exceeds the significant cut-off even at large lags.

Question 3

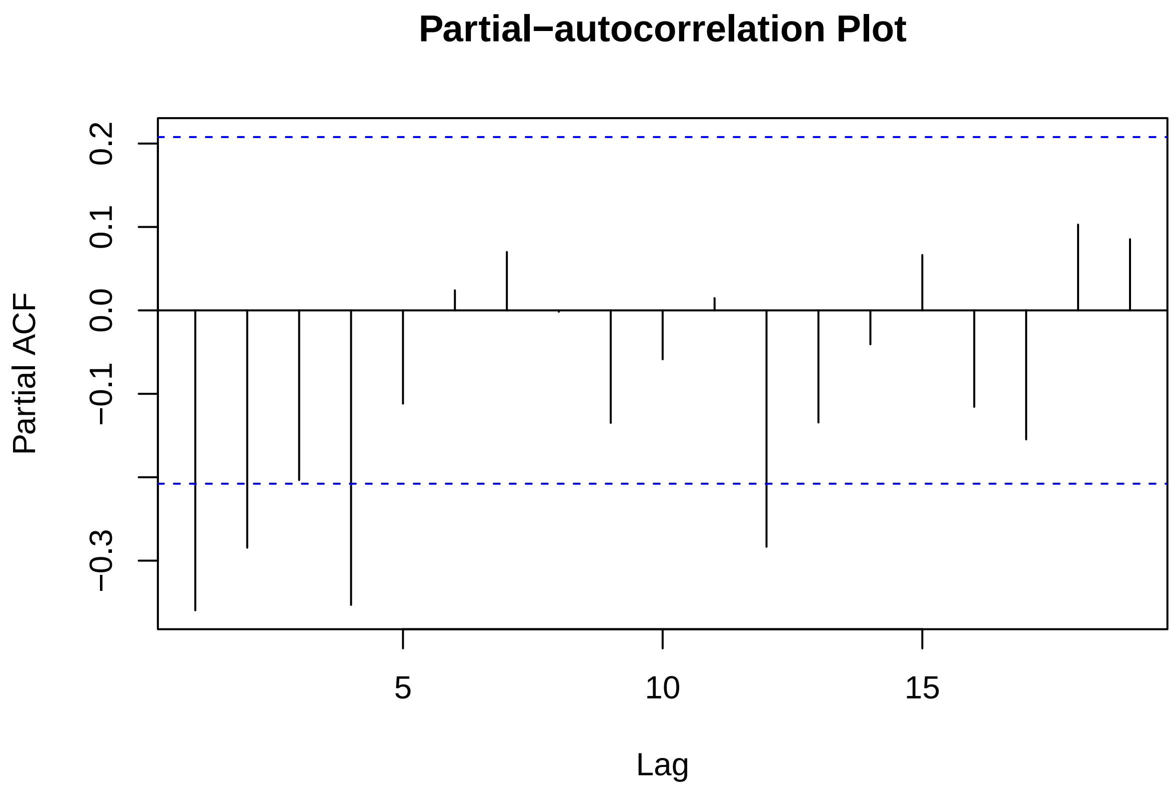
```
diffed_data <- diff(annual_ts, lag = 1, difference = 1)
plot(diffed_data,
     ylab = "Annual Minimum Temperatures",
     main = "Differenced Annual Minimum Temperature")
```




```r
acf(diffed_data,
    main = "Correlogram for Differenced Annual Minimum Temperature")
```



```r
pacf(diffed_data,
     main = "Partial-autocorrelation Plot")
```



The series of lag 1 difference now appears to be a MA(1) process as the acf plot shows a clear cut-off at lag 1.

Question 4

```r
arimamodel <- arima(annual_ts, order = c(0,1,1))
arimamodel
```

```
##
## Call:
## arima(x = annual_ts, order = c(0, 1, 1))
##
## Coefficients:
##           ma1
##       -0.7504
## s.e.   0.0892
##
## sigma^2 estimated as 1.143:  log likelihood = -132.65,  aic = 269.29
```

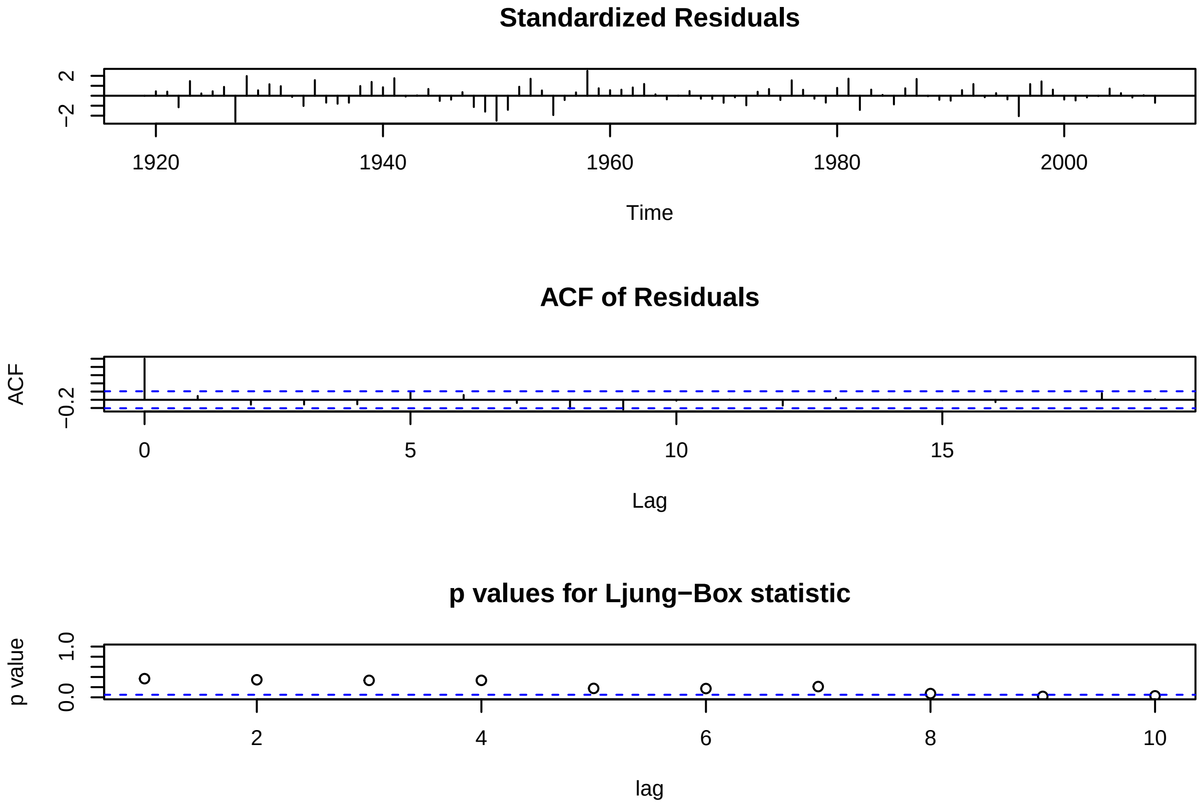
The fitted model is

$$W_t = -0.7504Z_{t-1} + Z_t$$

where $W_t = X_t - X_{t-1}$ and $Z_t \sim WN(0, 1.143)$

Question 5

```
tsdiag(arimamodel)
```



Most of the standardized residuals fall between the ± 2 range with a few exceeding this range. All acf values fall between the $\pm 2/\sqrt{n}$ range except for lag 0 which is expected. The Ljung-Box statistics are not significant for lag smaller than 9, and are significant for lag 9 and 10. Overall, the model fit is reasonable.

Question 6

```
q6_result <- tibble(Model = c("AR(1)", "ARIMA(0,1,1) "),
                    AIC = c(round(ar1model$aic,2),round(arimamodel$aic,2)))
kable(q6_result, caption = "AIC for Competing Models")
```

Table 1: AIC for Competing Models

Model	AIC
AR(1)	282.99
ARIMA(0,1,1)	269.29

Looking at the above result, the ARIMA(0,1,1) has a smaller AIC compared to AR(1) model, therefore the ARIMA(0,1,1) model is prefered.