STAT 443: Lab 7

Flora Zhang 52135365

7 Mar, 2022

```r
TempPG <- read_csv("TempPG.csv")
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   Year = col_double(),
##   Jan = col_double(),
##   Feb = col_double(),
##   Mar = col_double(),
##   Apr = col_double(),
##   May = col_double(),
##   Jun = col_double(),
##   Jul = col_double(),
##   Aug = col_double(),
##   Sep = col_double(),
##   Oct = col_double(),
##   Nov = col_double(),
##   Dec = col_double(),
##   Annual = col_double(),
##   Winter = col_double(),
##   Spring = col_double(),
##   Summer = col_double(),
##   Autumn = col_double()
## )
```

```r
TempPG
```

```
## # A tibble: 90 x 18
##     Year   Jan   Feb   Mar   Apr   May   Jun   Jul   Aug   Sep   Oct   Nov   Dec
##    <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
##  1  1919  -9.8 -17.4 -13.2 -2      0.9   4     6.8   5.7   2.5  -3.1 -13.3 -12.7
##  2  1920 -18.7 -15.3  -9.4  -6.3  -3.1   3.8   8.5   5.8   2.4   0.8  -3.2 -10
##  3  1921 -15.9 -11.3  -7.7  -2.8   1     5     5.7   5.8   2.3   0.9  -9.9 -15.6
##  4  1922 -15.8 -25.2 -11.4  -3.3   0.7   2.8   5.7   4.4   2.3  -1.7  -4.1 -15.8
##  5  1923 -16   -11.7  -7    -3.7   1.7   6.4   6.8   7.8   1    -3.1  -3.2 -10.4
##  6  1924 -13.2  -8.7  -8.4  -2.8   0.8   4     6.8   5.1   2.5   0.7  -8.2 -20.9
##  7  1925 -15.8 -16    -8.7  -3.2   0.8   4.5   6.5   5.4   0.3  -4.5  -4.3  -3.8
##  8  1926  -7.6  -4.7  -4.4  -1.3   0.7   3.6   5.2   4.2  -2.6  -1.9  -9   -13.3
##  9  1927 -17.1 -16.5  -7.8  -4.2  -0.4   5.1   6.1   4.4   0.6  -3   -15.6 -24.3
## 10  1928 -13.2 -10.8  -7.3  -2.6   2.7   7.2   8.1   6.1   1    -0.8  -3.2 -10.3
## # ... with 80 more rows, and 5 more variables: Annual <dbl>, Winter <dbl>,
## #   Spring <dbl>, Summer <dbl>, Autumn <dbl>
```

1. Fit the AR(1) model to the annual minimum temperatures using the arima()command, and write down your fitted model.

```r
annual = ts(data=TempPG$Annual)
x = arima(annual, order=c(1,0,0) ,include.mean=T)
x
```

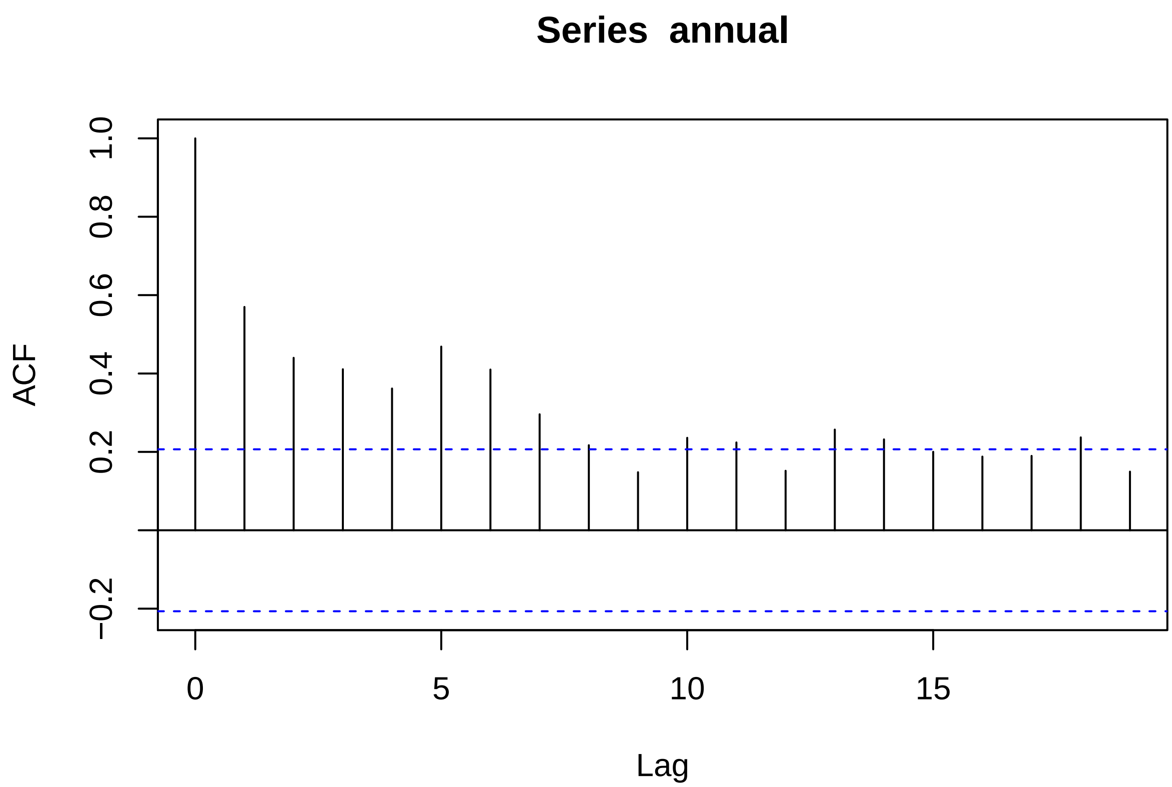
```
##
## Call:
## arima(x = annual, order = c(1, 0, 0), include.mean = T)
##
## Coefficients:
##          ar1  intercept
##       0.5843    -1.9591
## s.e.  0.0864     0.2810
##
## sigma^2 estimated as 1.265:  log likelihood = -138.49,  aic = 282.99
```

The parameter is: ar1=0.5843 with a s.e. of 0.0864, and intercept = -1.9591 with s.e. of 0.2810 The model is: y_t = -1.9591 + 0.5843y_{t-1} + e_t

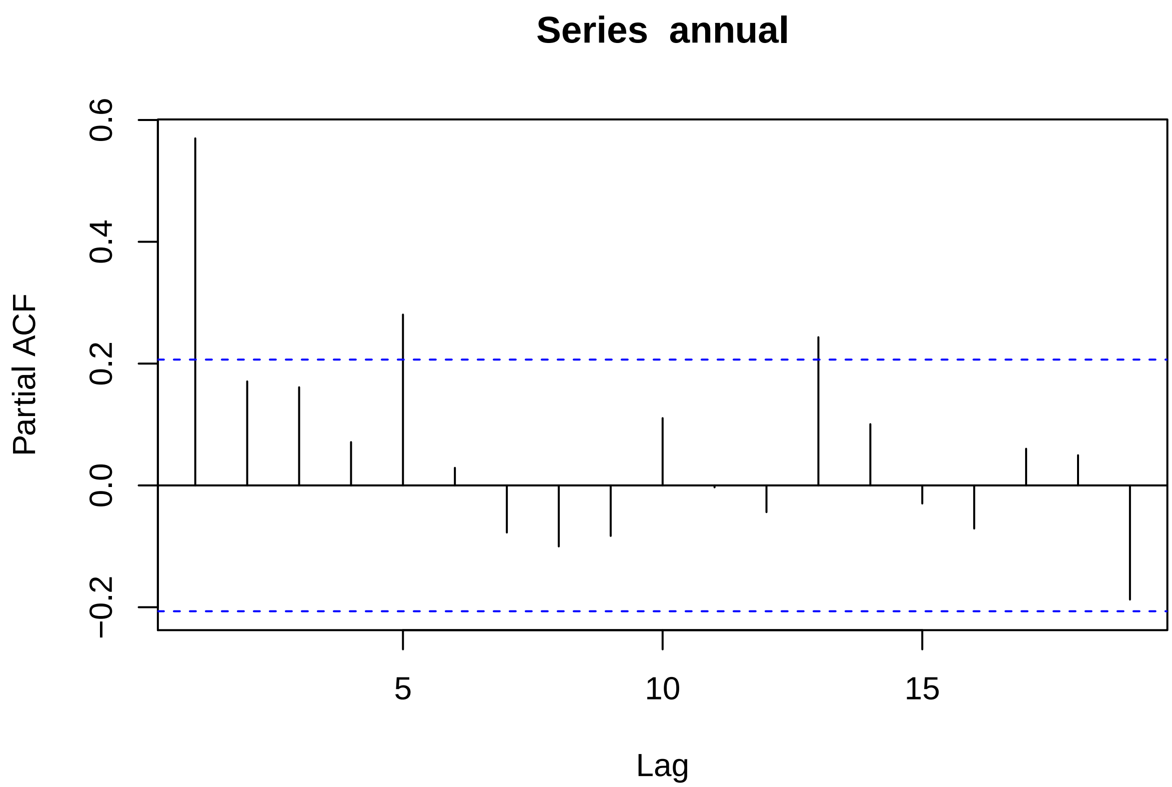
2. Look again at the acf of the annual minimum temperatures. In what way does the acf not behave as you would expect for the fitted AR(1) model?

```r
#generate simulations for actual AR(1) model


acf(annual)
```



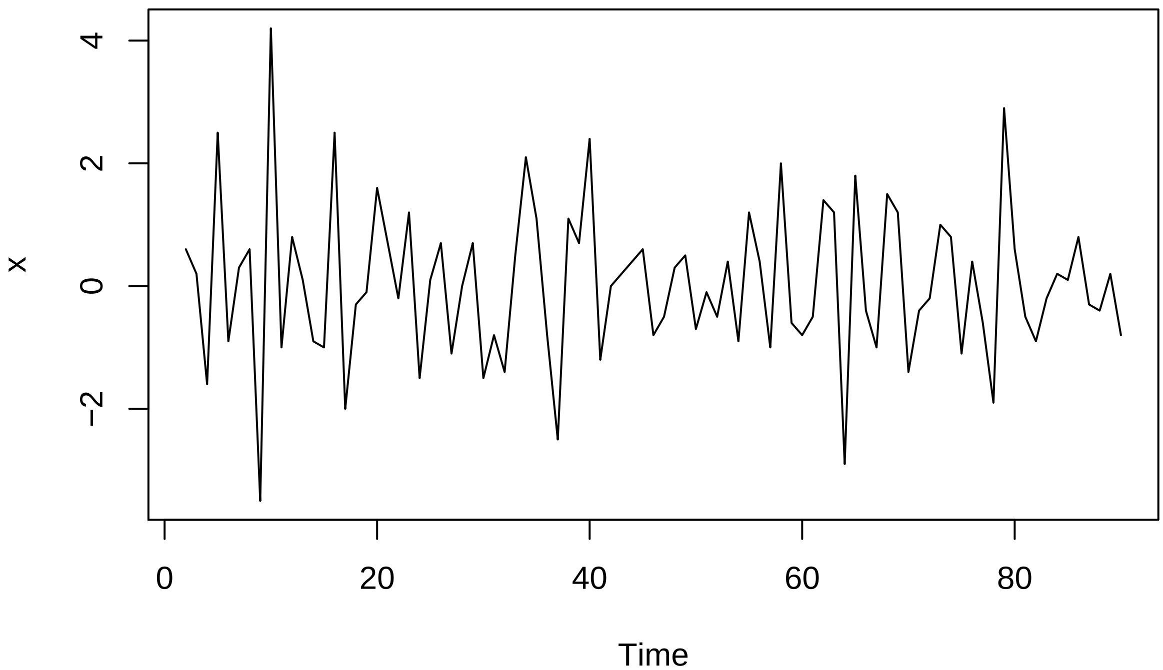

```
pacf(annual)
```



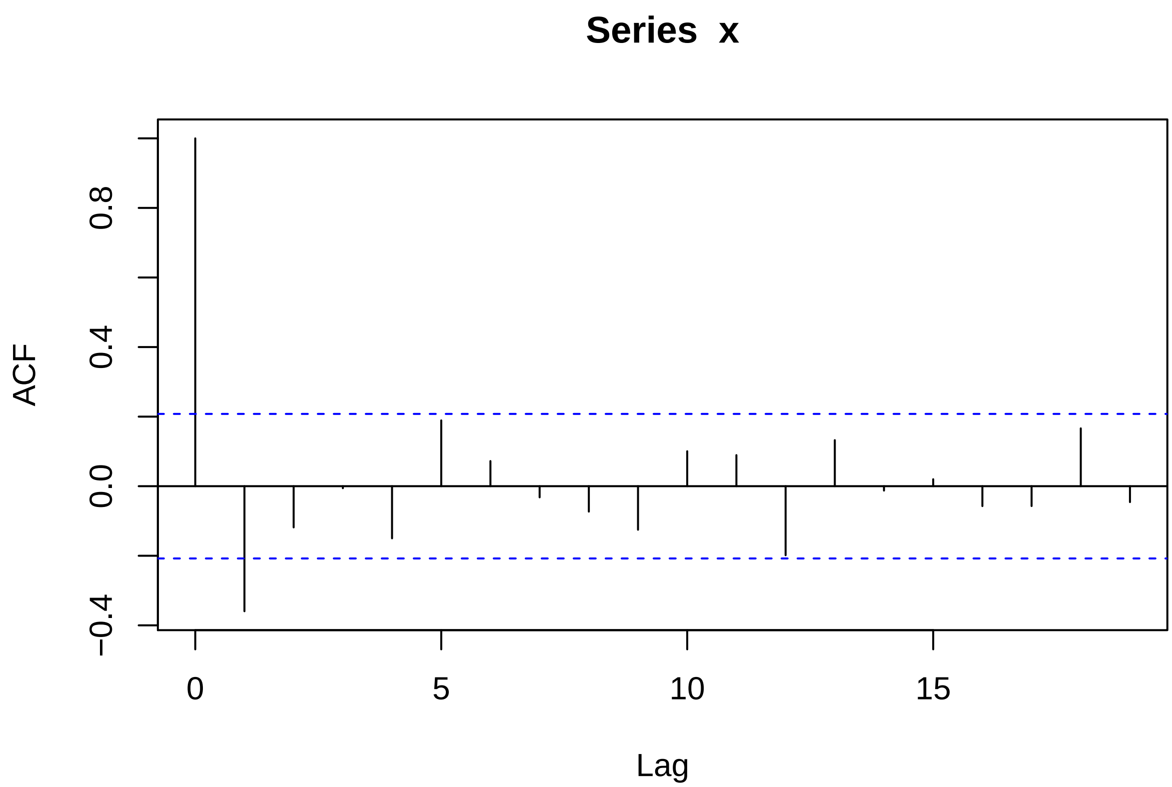
We would expect the acf for an AR(1) model to decrease exponentially to 0. In this acf plot, we can see that the graph does not directly decrease exponentially. For example, in period 5, the value increases. In addition, the ACF seems to be decreasing very slowly, and values remain well above the significance range (dotted blue lines), indicating that the time series is not stationary. Additionally, the pacf plot does not cut off after lag 1, which is inconsistent with the behaviour of AR(1).

3. Plot the series of first differences of minimum annual temperatures. Plot the acf and pacf of the differences. What model would you suggest for the differences?

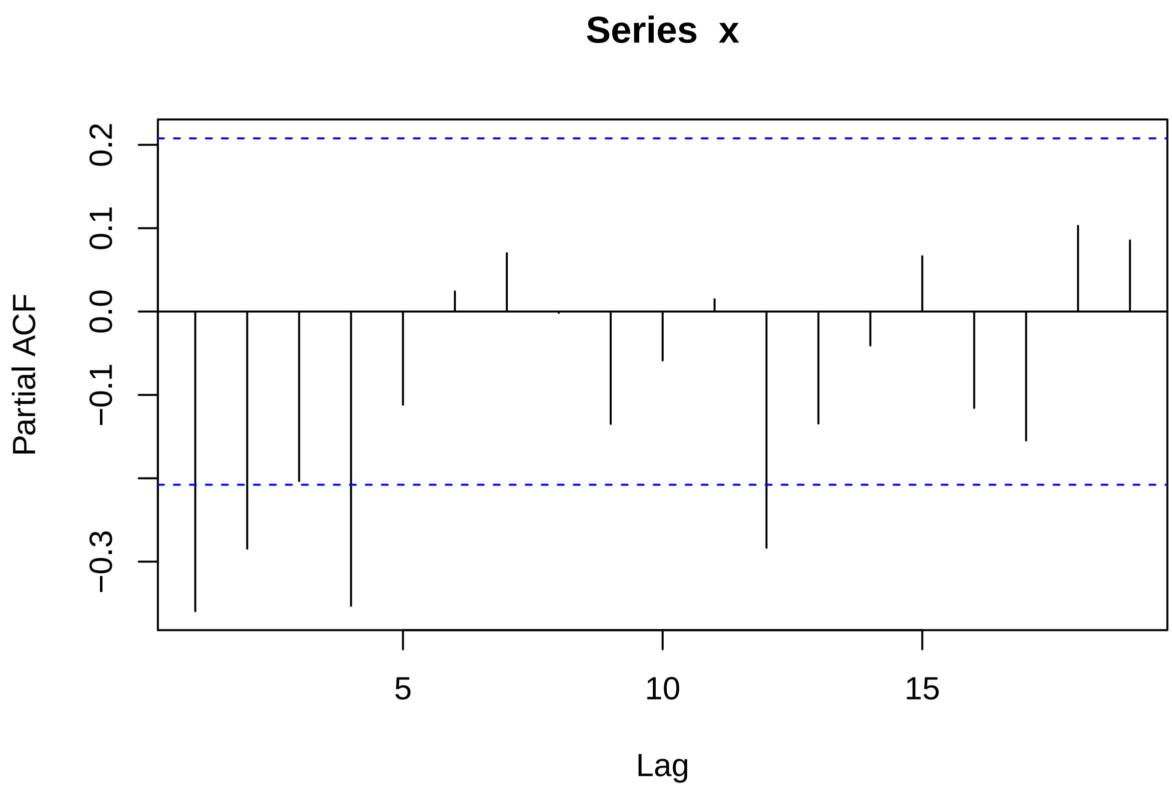
```
x = diff(annual, lag=1)
plot(x)
```




```
acf(x)
```




```
pacf(x)
```



From the newly differenced graph, we can see that the ACF cuts off after lag 1. Therefore, we can consider MA(1) for the differenced data, or ARIMA(0,1,1). We may also try ARIMA(1,1,1) as a model, since the PACF does not seem to fit the MA model well for the differenced model.

4.Fit the suggested ARIMA model to the annual minimum temperatures series. Write down your fitted model

We will fit the simplest model, which is ARIMA(0,1,1)

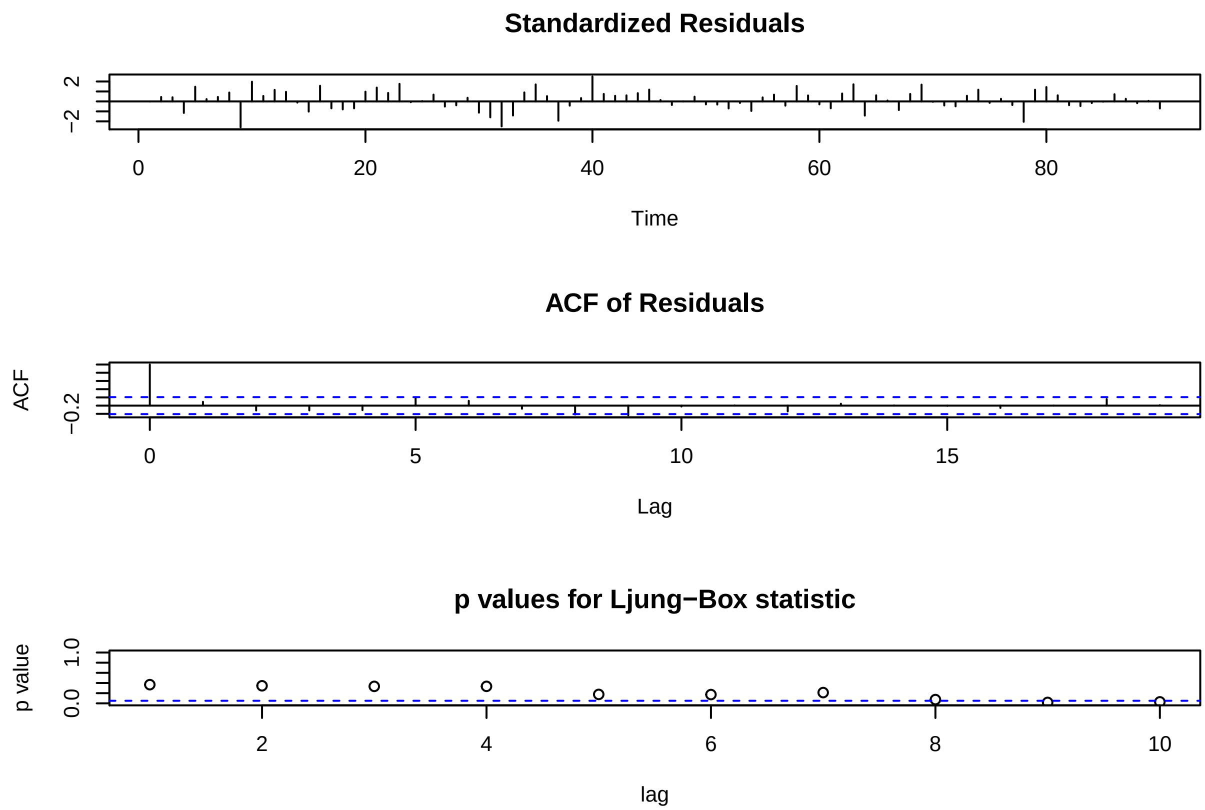
```r
fitted_data = arima(annual, order=c(0,1,1))
fitted_data
```

```
##
## Call:
## arima(x = annual, 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 parameter is: ma1=-0.7504 with a s.e. of 0.0892. We also see that this model's aic (269.29) is less than the aic from the non-differenced model (282.99), which represents that this model is a better fit compared to that of the non-differenced model. The model is: X_t = -0.7504Z_{t-1}

5. Use the tsdiag() function to see diagnostic plots for the model you have fitted. How well does the model appear to fit?

```
tsdiag(fitted_data)
```



From the diagnostic plots, we can see that the acf looks like there is no obvious serial correlation. The p-values are insignificant, until lags 8<M<10. This could be judged as acceptable.

6. When comparing the arima(1,0,0) model (undifferenced) and the arima(0,1,1) model (differenced), we observe that the differenced model has a smaller aic(269.29) compared to the undifferenced (282.99). Therefore, we would choose the arima (0,1,1) model.