Lab 6: STAT 443

Caden Hewlett

2024-02-16

Question 1

The dataset `TempPG.csv` includes minimum temperatures measured at Prince George, BC, from 1919 to 2008. Read the data into R using either `read.table()` or `read.csv()` commands.

The column labelled "Summer" contains the summer minimum temperatures.

Part 1:

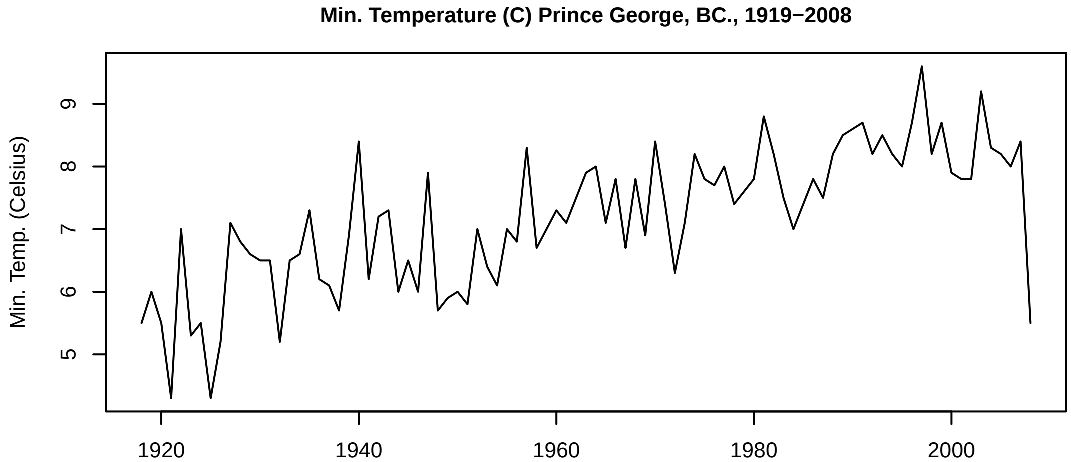
Extract those data, and coerce them into a time series object.

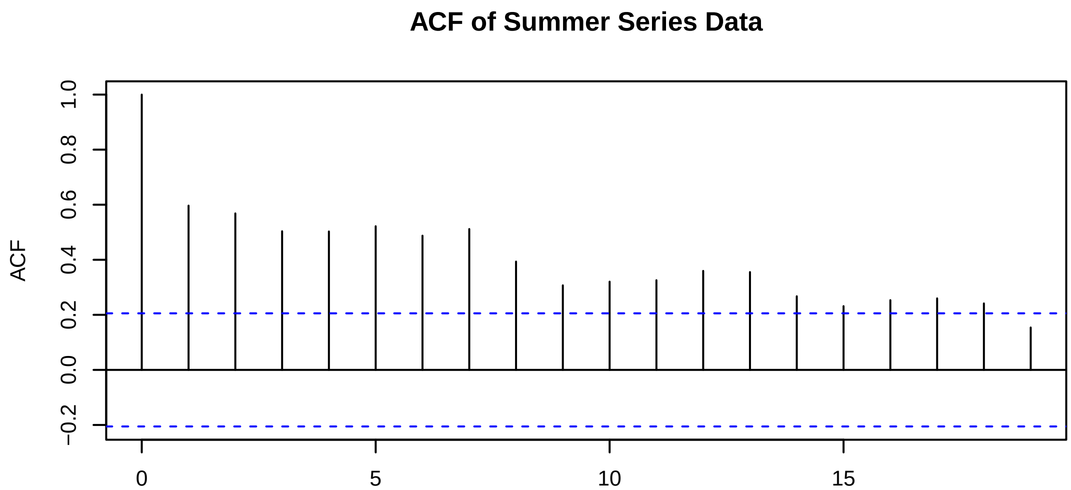
```
df = read.csv("TempPG.csv")
summer = df$Summer

summerts = ts(summer, start = 1918,
              end = 2008, freq = 1)
```

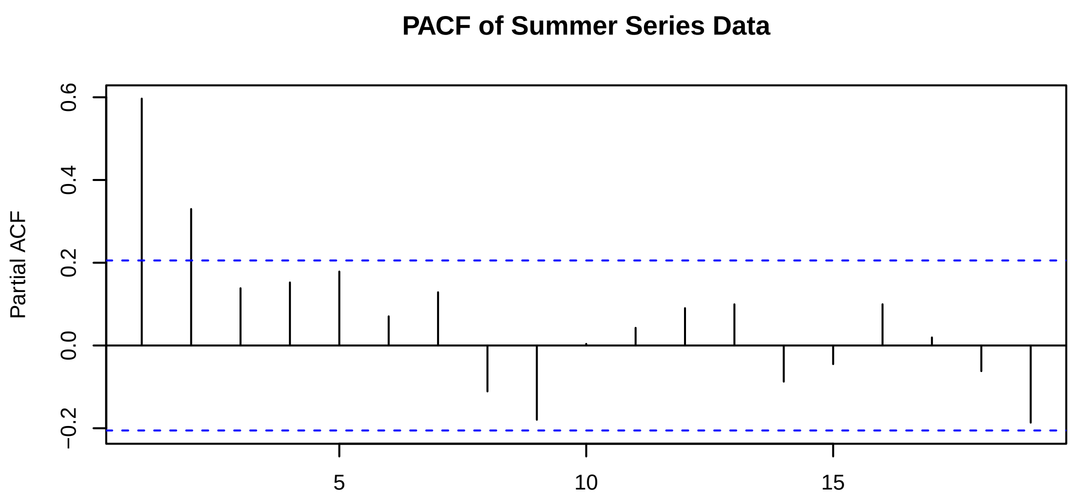
Part 2

Plot the time series, the sample acf and pacf.

```
par(mfrow = c(3, 1), mai = c(0.4, 0.6, 0.4, 0.4))
plot.ts(
  summerts,
  main = "Min. Temperature (C) Prince George, BC., 1919-2008",
  cex.main = 1.0,
  xlab = "Year",
  ylab = "Min. Temp. (Celsius)"
)
acf(summerts, main = "ACF of Summer Series Data")
pacf(summerts, main = "PACF of Summer Series Data")
```







Part 3

Comment on what you observe in these plots.

We see a slow (exponential) decline in magnitude in the ACF, implying that an AR model might be applicable. In an inspection of the PACF, we see that the final $\hat{\alpha}_{kk}$ value for significant coefficients is $k = 2$, this implies that if we were to fit an AR component it would have $p = 2$. However, there might be an additional (positive) trend component in the data (as seen in the first plot), that is causing us to see other values in the ACF. So, without first attempting to de-trend and perhaps difference the series, any conclusions made here are rudimentary.

Part 4

What ARMA model would you fit?

I would fit an ARMA$(p = 2, q = 0)$ model based on this analysis, for the reasons explained in the comments above; namely, that there appears to be an autoregressive component of order 2, and no visible moving average component.

Question 2

Fit the ARMA model you proposed above using the `arima()` command. Note that in the output of the arima command, 'intercept' refers to the mean of the process, which we denote by μ in class

```
fitted = arima(summerts, order = c(2, 0, 0))
fitted
```

```
##
## Call:
## arima(x = summerts, order = c(2, 0, 0))
##
## Coefficients:
##          ar1     ar2  intercept
##       0.4026  0.3446     7.0669
## s.e.  0.1029  0.1041     0.3263
##
## sigma^2 estimated as 0.6672:  log likelihood = -111.07,  aic = 230.15
```

Write down your fitted model.

Our original model (in pure theory) would be:

$$(X_t - \mu) = \alpha_1(X_{t-1} - \mu) + \alpha_2(X_{t-2} - \mu) + Z_t$$

Where $Z_t \sim \text{WN}(0, \sigma^2)$.

We now would have:

$$(X_t - \hat{\mu}) = \hat{\alpha}_1(X_{t-1} - \mu) + \hat{\alpha}_2(X_{t-2} - \hat{\mu}) + Z_t, \text{ where } Z_t \sim \text{WN}(0, \hat{\sigma}^2)$$

From the results, we see $\hat{\alpha}_1 = 0.4026$, $\hat{\alpha}_2 = 0.3446$, $\hat{\mu} = 7.0669$ and $\hat{\sigma}^2 = 0.6672$. In other words,

$$(X_t - 7.0669) = 0.4026(X_{t-1} - 7.0669) + 0.3446(X_{t-2} - 7.0669) + Z_t$$

Where $Z_t \sim \text{WN}(0, 0.6672)$.

Question 3

Use the `confint()` command to find 95% confidence intervals for relevant parameters.

(*NOTE*: We adopt the assumption of non-simultaneous confidence intervals, and hence do not apply a Bonferroni correction.)

The following will generate the CIs for $\hat{\alpha}_1$, $\hat{\alpha}_2$ and $\hat{\mu}$.

```
confint(fitted, level = 0.95)
```

```
##                2.5 %    97.5 %
## ar1        0.2008682 0.6043640
## ar2        0.1404444 0.5486957
## intercept 6.4273702 7.7063323
```

If we wished, we could manually construct the confidence interval for the $\hat{\sigma}^2$ estimate, by assuming it takes a χ^2 distribution on $n - k$ degrees of freedom, where $k = 4$ for our 4 estimated parameters (estimates of α_1, α_2, μ and σ^2 itself.)

Let $\nu = n - k$. Recall that under the assumptions of the variance taking a chi-square distirbution,

$$\text{CI} = \left\{ \frac{\nu\hat{\sigma}^2}{\chi^2_{\nu;1-(\alpha/2)}}, \frac{\nu\hat{\sigma}^2}{\chi^2_{\nu;(\alpha/2)}} \right\}$$

In `R`, this is:

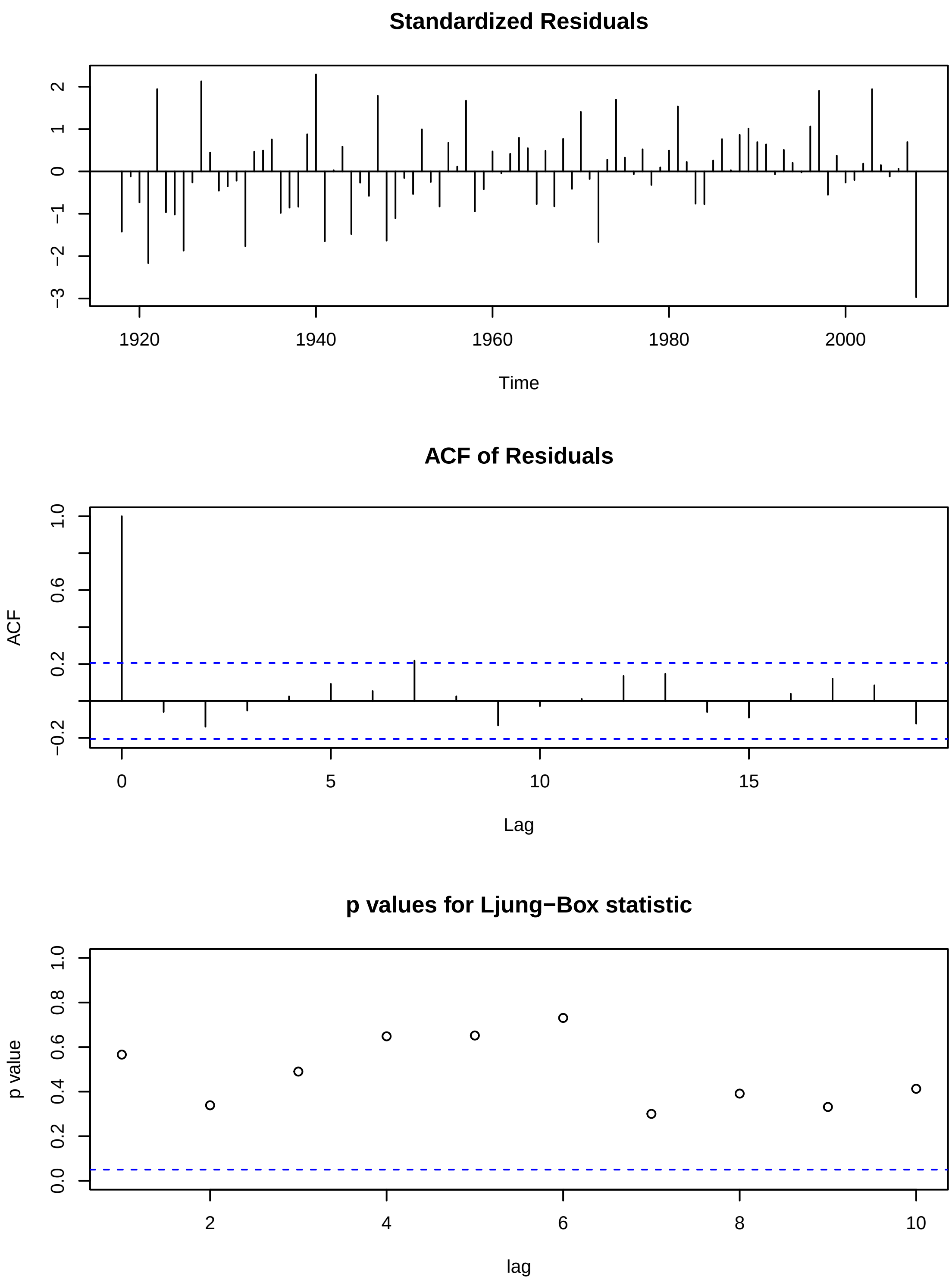
```
degf = length(summerts) - 4
c((fitted$sigma2 * degf)/qchisq(0.975, degf),
   (fitted$sigma2 * degf)/qchisq(0.025, degf))
```

```
## [1] 0.5061079 0.9200760
```

Question 4

Use the `tsdiag()` function to see diagnostic plots for the model you have fitted (remember to include “fig.height” option for a better display of your plots)

```
tsdiag(fitted)
```


Standardized Residuals
−3
−2
−1
0
1
2
1920
1940
1960
1980
2000
Time
ACF of Residuals
ACF
−0.2
0.2
0.6
1.0
0
5
10
15
Lag
p values for Ljung−Box statistic
p value
0.0
0.2
0.4
0.6
0.8
1.0
2
4
6
8
10
lag