Homework 5

Benjamin Noland

1. (Cryer & Chan, Exercise 5.2)

   a. We have

   $$\nabla Y_t = Y_t - Y_{t-1} = 3 + e_t - 0.75 e_{t-1}.$$

   Therefore $\{\nabla Y_t\}$ is an MA(1) process, hence stationary. Its MA characteristic polynomial is $\theta(x) = 1 - 0.75x$, which has root $x = 4/3 > 1$, so that the process is invertible.

   Since $\{e_t\}$ is a (mean zero) white noise process, we get

   $$\mathrm{E}(\nabla Y_t) = \mathrm{E}(3 + e_t - 0.75 e_{t-1}) = 3,$$

   and

   $$\begin{aligned}\mathrm{Var}(\nabla Y_t) &= \mathrm{Var}(3 + e_t - 0.75 e_{t-1}) \\ &= \mathrm{Var}(e_t) + 0.75^2 \mathrm{Var}(e_{t-1}) \\ &= \sigma_e^2 + 0.75^2 \sigma_e^2 \\ &= \frac{25}{16}\sigma_e^2,\end{aligned}$$

   where σ_e^2 denotes the variance of the white noise process $\{e_t\}$.

   b. We have

   $$\begin{aligned}\nabla Y_t &= Y_t - Y_{t-1} \\ &= 10 + 0.25(Y_{t-1} - Y_{t-2}) + e_t - 0.1 e_{t-1} \\ &= 10 + 0.25 \nabla Y_{t-1} + e_t - 0.1 e_{t-1}.\end{aligned}$$

   The AR characteristic polynomial of this process is $\phi(x) = 1 - 0.25x$, which has root $x = 4 > 1$, so that the process is stationary. Moreover, its MA characteristic polynomial $\theta(x) = 1 - 0.1x$ has root $x = 10 > 1$, and so the process is invertible. We therefore see that $\{\nabla Y_t\}$ is an ARMA(1, 1) process.

   Note that $\{e_t\}$ is a (mean zero) white noise process. Furthermore, since $\{\nabla Y_t\}$ is stationary, its mean μ is constant in t. Therefore,

   $$\mathrm{E}(\nabla Y_t) = \mathrm{E}(10 + 0.25 \nabla Y_{t-1} + e_t - 0.1 e_{t-1}) = \mu = 10 + 0.25\mu,$$

   so that $\mu = 40/3$. Using the expression for γ_0 for a stationary ARMA(1, 1) process (see page 78), we get

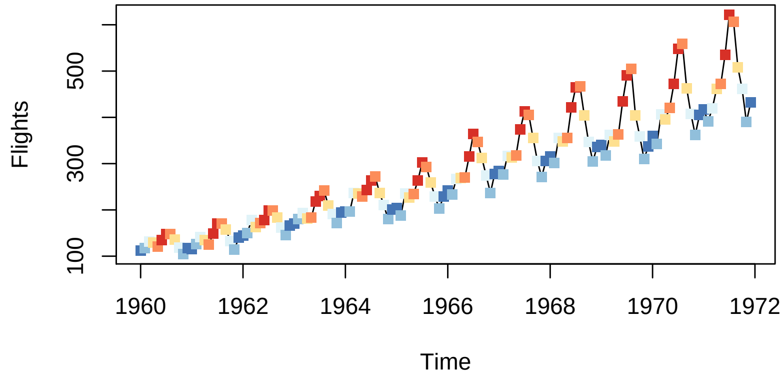
   $$\begin{aligned}\mathrm{Var}(\nabla Y_t) &= \mathrm{Var}(10 + 0.25 \nabla Y_{t-1} + e_t - 0.1 e_{t-1}) \\ &= \mathrm{Var}(0.25 \nabla Y_{t-1} + e_t - 0.1 e_{t-1}) \\ &= \gamma_0 \\ &= \frac{1 - 2(0.25)(0.1) + 0.1^2}{1 - 0.25^2}\sigma_e^2 \\ &= \frac{128}{125}\sigma_e^2,\end{aligned}$$

where σ_e^2 denotes the variance of the white noise process $\{e_t\}$.

2. (Cryer & Chan, Exercise 5.13)

    a. The following R code produces a time series plot of the data:

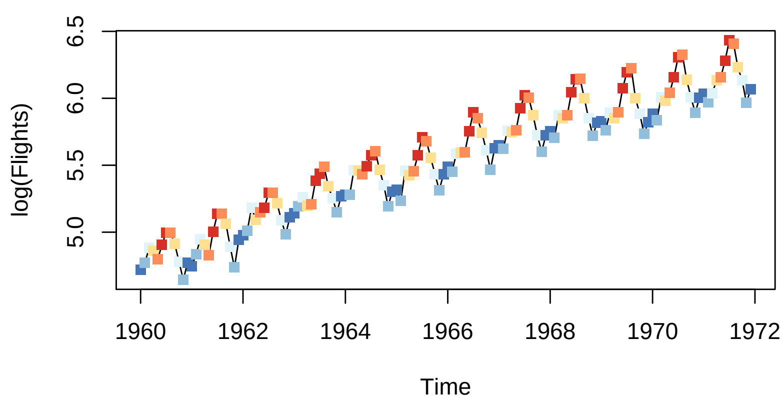
```
temp_color = c(rev(brewer.pal(6, 'RdYlBu')), brewer.pal(6, 'RdYlBu'))

plot(airpass, ylab = "Flights", type = "l")
points(y = airpass, x = time(airpass), col = temp_color, pch = 15)
```



The plot seems to indicate a strong seasonal trend in the time series, with more flights in the warmer months than in the colder months. Moreover, the mean number of flights per year is steadily increasing over the whole time span of the series, as well as the variability in monthly flights within a given year – later years display considerably more variability than earlier years in the series.

    b. The following R code produces a time series plot of the log-transformed series:

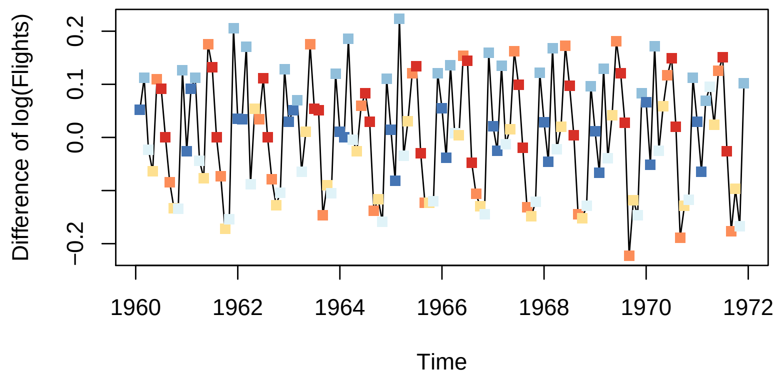
```
plot(log(airpass), ylab = "log(Flights)", type = "l")
points(y = log(airpass), x = time(airpass), col = temp_color, pch = 15)
```



The log transformation seems to have considerably dampened the yearly variability in the series. However, as in the untransformed case, the yearly means of the transformed data are steadily increasing, and the series displays the same seasonal trend.

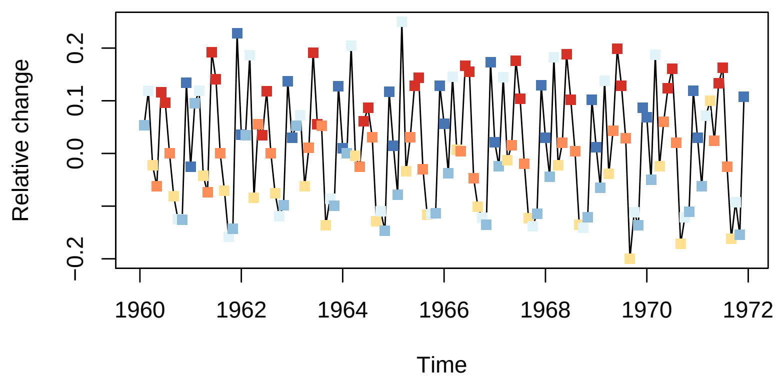
c. The following R code produces a time series plot of the differences of the log-transformed data:

```
diff_log <- diff(log(airpass))
plot(diff_log, ylab = "Difference of log(Flights)", type = "l")
points(y = diff_log, x = time(diff_log), col = temp_color, pch = 15)
```



Next, we have a time series plot of the (fractional) relative differences:

```
rel_change <- (airpass / zlag(airpass)) - 1
plot(rel_change, ylab = "Relative change", type = "l")
points(y = rel_change, x = time(rel_change), col = temp_color, pch = 15)
```



The two plots seem to display the same, relatively stable trend. Assuming the true model is of the form

$$Y_t = (1 + X_t)Y_{t-1},$$

with $|X_t|$ small for every t, then

$$\nabla \log Y_t = \log\left(\frac{Y_t}{Y_{t-1}}\right) = \log(1 + X_t) \approx X_t$$

for every t. Moreover,

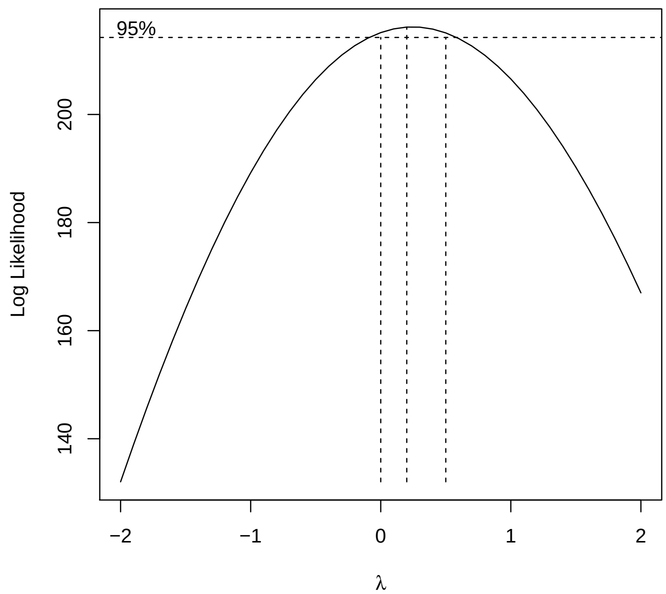
$$\frac{Y_t - Y_{t-1}}{Y_{t-1}} = \frac{Y_t}{Y_{t-1}} - 1 = (1 + X_t) - 1 = X_t.$$

Thus, under this model, we would expect the two plots to look about the same.

3. (Cryer & Chan, Exercise 5.14)

a. The following R code generates a plot of log-likelihood against λ.

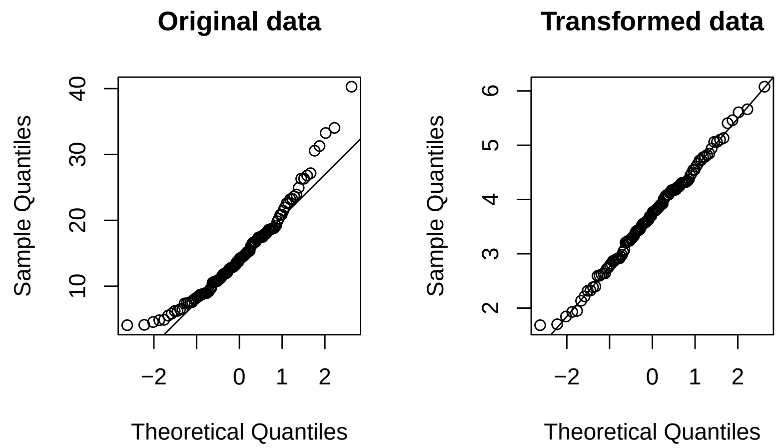
```r
BoxCox.ar(larain)
```



The vertical dashed lines denote a 95% confidence interval for λ. In particular, from the plot we see that the value $\lambda = 0.25$ lies within this confidence interval. For the sake of potential interpretability, we will take this to be the value of λ for the power transformation.

b. The following R code produces two plots: the left plot is a normal Q-Q plot of the original (untransformed) data, and the right plot is one of the transformed data.

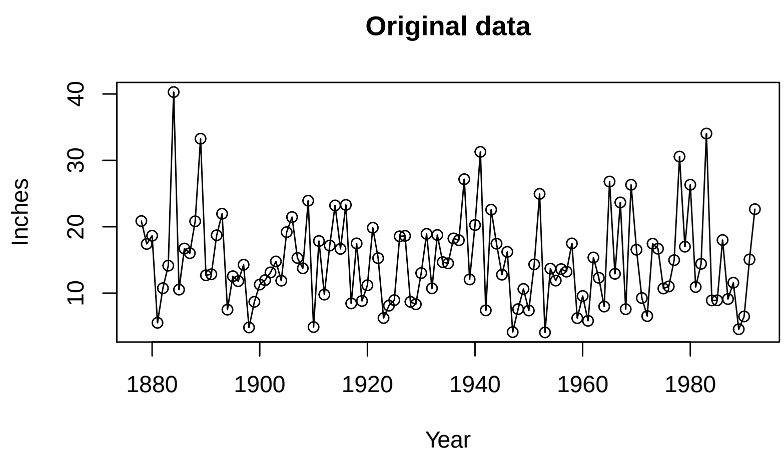
```r
lambda <- 0.25
transformed <- (larain^lambda - 1) / lambda

old_par <- par(mfrow = c(1, 2))
qqnorm(larain, main = "Original data")
qqline(larain)

qqnorm(transformed, main = "Transformed data")
qqline(transformed)
par(old_par)
```

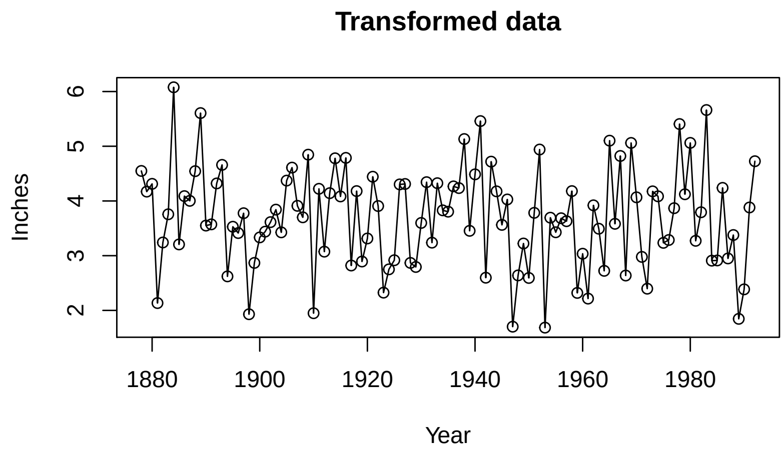


The normal Q-Q plot of the transformed data adheres very closely to the theoretical Q-Q plot, indicated by the solid black line. Thus, in contrast with the original data, the transformed data appear to be normal.

c. The following R code produces time series plots of both the original (untransformed) data and the transformed data:

```
plot(larain, type = "o", main = "Original data",
     xlab = "Year", ylab = "Inches")
plot(transformed, type = "o", main = "Transformed data",
     xlab = "Year", ylab = "Inches")
```

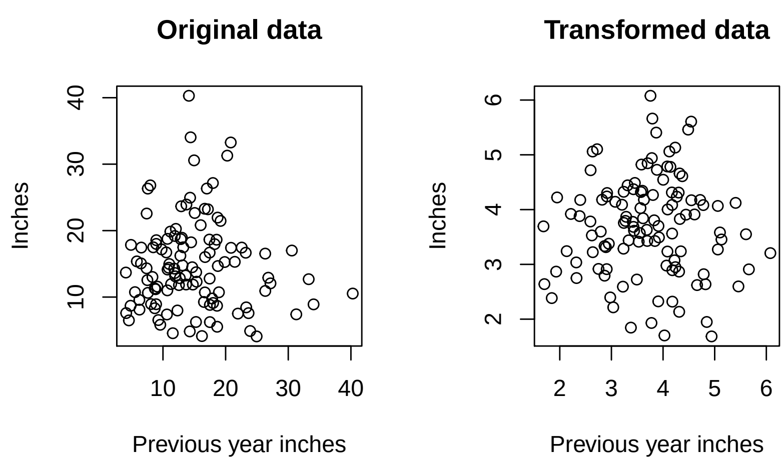




Notice that the correlations between the data points appear to be similar in both plots.

d. The following R code plots Y_t vs. Y_{t-1} for both the original (untransformed) data and the transformed data.

```
old_par <- par(mfrow = c(1, 2))
plot(y = larain, x = zlag(larain), main = "Original data",
     xlab = "Previous year inches", ylab = "Inches")
plot(y = transformed, x = zlag(transformed), main = "Transformed data",
     xlab = "Previous year inches", ylab = "Inches")
par(old_par)
```



Both plots seem to indicate a lack of correlation between Y_t and Y_{t-1}. Moreover, we should not expect the transformation to change the dependence between data points in this series. This is a consequence of the following fact: if $Y_1, \dots, Y_n$ are random variables and g is a (measurable) function, then $g(Y_1), \dots, g(Y_n)$ are random variables, and they are mutually independent if and only if $Y_1, \dots, Y_n$ are mutually independent. In this case, $Y_1, \dots, Y_n$ are elements of the time series, and $g(x) = (x^\lambda - 1)/\lambda$, where $\lambda = 0.25$.

4. Bartlett's Theorem states that, subject to some regularity conditions (which are satisified

for any stationary ARMA process – see page 110 of the text), the joint distribution

$$(\sqrt{n}(r_1 - \rho_1), \ldots, \sqrt{n}(r_m - \rho_m)) \xrightarrow{\mathcal{D}} \mathrm{N}_m(0, C),$$

where $C = (c_{ij})$ is an $m \times m$ matrix defined by

$$c_{ij} = \sum_{k=-\infty}^{\infty} (\rho_{k+i}\rho_{k+j} + \rho_{k-i}\rho_{k+j} - 2\rho_i\rho_k\rho_{k+j} - 2\rho_j\rho_k\rho_{k+i} + 2\rho_i\rho_j\rho_k^2).$$

i. In particular, for any $m \geq 1$, we can apply Bartlett's Theorem to get the asymptotic marginal distribution of r_m:

$$\sqrt{n}(r_m - \rho_m) \xrightarrow{\mathcal{D}} \mathrm{N}(0, c),$$

where

$$c = \sum_{k=-\infty}^{\infty} (\rho_{k+1}^2 + \rho_{k-1}\rho_{k+1} - 4\rho_1\rho_k\rho_{k+1} - 2\rho_1^2\rho_k^2).$$

For a white noise process, $\rho_0 = 1$ and $\rho_k = 0$ for every $k \neq 0$, and so this expression reduces to $c = \rho_0^2 = 1$. Thus, for large n, since $\rho_m = 0$, r_m is approximately distributed as $\mathrm{N}(0, 1/n)$.

ii. Let $i, j \geq 1$, with $i \neq j$. Then, for a white noise process, $c_{ij} = 0$. Applying Bartlett's Theorem to compute the asymptotic marginal distribution of (r_i, r_j), we conclude that for large n, (r_i, r_j) is approximately $\mathrm{N}_2(\rho, C/n)$, where $\rho = (\rho_1, \rho_2)$ and C is the 2×2 matrix

$$C = \begin{pmatrix} c_{ii} & c_{ij} \\ c_{ij} & c_{jj} \end{pmatrix}.$$

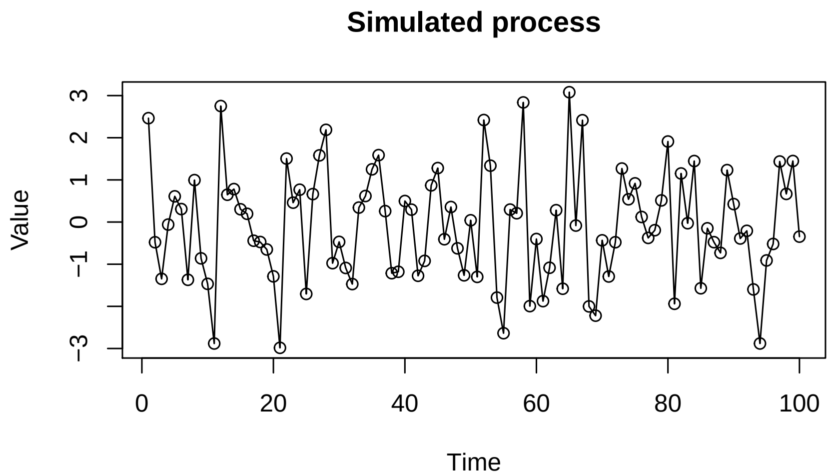
Thus, since $c_{ij} = 0$, we see that r_i and r_j are approximately uncorrelated for large n.

5. The following R code generates $n = 100$ observations from the MA(3) process in question using Gaussian noise:

```r
set.seed(1)  # Set a random seed for reproducability
process <- arima.sim(n = 100, model = list(ma = -c(0.7, 0.5, 0.6)))
```

The following is a time series plot of the the resulting simulated data:

```r
plot(process, type = "o", main = "Simulated process", ylab = "Value")
```



i. The following computes the sample autocorrelation coefficient r_4:

```r
r_4 <- acf(process, plot = FALSE)$acf[[4]]
r_4
```

```
## [1] -0.09994342
```

ii. For a general MA(q) process, we have

$$\mathrm{Var}(r_k) \approx \frac{1}{n}\left(1 + 2\sum_{j=1}^{q} \rho_j^2\right) \quad \text{for any } k > q$$

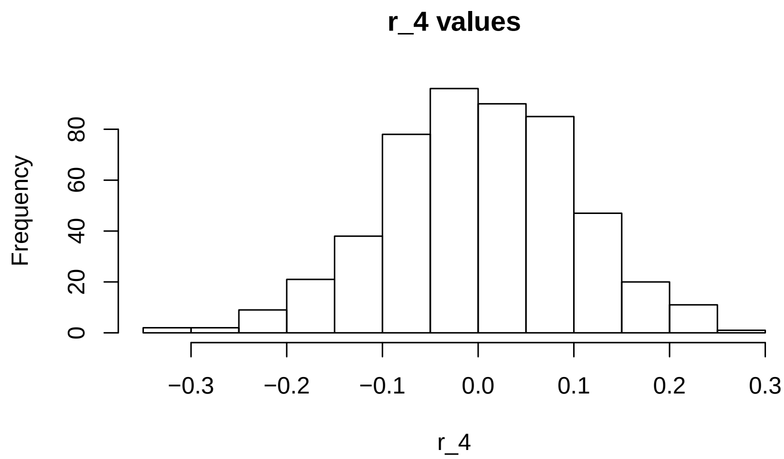
for large n. Using this expression, the following computes $\mathrm{Var}(r_4)$ for the model in question:

```r
rho_values <- ARMAacf(ma = -c(0.7, 0.5, 0.6))[2:4]
var_r_4 <- (1 + 2 * sum(rho_values^2)) / 100
var_r_4
```

```
## [1] 0.01167302
```

iii. The following R code simulates the process 500 times and plots a histogram of the resulting r_4 values:

```r
r_4_values <- numeric(500)
for (i in 1:length(r_4_values)) {
  p <- arima.sim(n = 100, model = list(ma = -c(0.7, 0.5, 0.6)))
  r_4_values[[i]] <- acf(p, plot = FALSE)$acf[[4]]
}

hist(r_4_values, main = "r_4 values", xlab = "r_4")
```



The variance of the simulated r_4 values is computed as

```
var(r_4_values)
```

```
## [1] 0.009895135
```

which is very close to the value computed in part (ii).

iv. We have the estimator

$$\widehat{\mathrm{Var}(r_4)} = \frac{1}{n}\left(1 + 2\sum_{k=1}^{3} r_k^2\right)$$

for $\mathrm{Var}(r_4)$. The following computes this estimator on the originally generated data:

```
acf_values <- acf(process, plot = FALSE)$acf[1:3]
var_r_4_est <- (1 + 2 * sum(acf_values^2)) / 100
var_r_4_est
```

```
## [1] 0.01219351
```

v. Since r_4 is approximately distributed as $\mathrm{N}(\rho_4, c/n)$ for large n, where

$$c = 1 + 2\sum_{j=1}^{3} \rho_j^2 \approx n\mathrm{Var}(r_4),$$

we can use the estimator from part (iv) to construct an approximate 95% confidence interval for ρ_4 as follows:

```
se <- sqrt(var_r_4_est) / 10
conf_int <- c(r_4 - 2 * se, r_4 + 2 * se)
conf_int
```

```
## [1] -0.12202827 -0.07785858
```

vi. The value 0 lies outside of the 95% confidence interval computed in part (v), and so we can reject the null hypothesis $H_0 : \rho_4 = 0$ at level $\alpha = 0.05$.