Solutions Manual to Accompany

Time Series Analysis

with Applications in R, Second Edition

by Jonathan D. Cryer and Kung-Sik Chan

Solutions by Jonathan Cryer and Xuemiao Hao, updated 7/28/08

CHAPTER 1

Exercise 1.1 Use software to produce the time series plot shown in Exhibit (1.2), page 2. The following R code will produce the graph.

```
> library(TSA); data(larain); win.graph(width=3,height=3,pointsize=8)
> plot(y=larain,x=zlag(larain),ylab='Inches',xlab='Previous Year Inches')
```

Exercise 1.2 Produce the time series plot displayed in Exhibit (1.3), page 3. Use the R code

```
> data(color); plot(color,ylab='Color Property',xlab='Batch',type='o')
```

Exercise 1.3 Simulate a completely random process of length 48 with independent, normal values. Repeat this exercise several times with a new simulation, that is, a new seed, each time.

```
> plot(ts(rnorm(n=48)),type='o') # If you repeat this command R will use a new "random
    numbers" each time. If you want to reproduce the same simulation first use the
    command set.seed(#########) where ######### is an integer of your choice.
```

Exercise 1.4 Simulate a completely random process of length 48 with independent, chi-square distributed values each with 2 degrees of freedom. Use the same R code as in the solution of **Exercise 1.3** but replace `rnorm(n=48)` with `rchisq(n=48,df=2)`.

Exercise 1.5 Simulate a completely random process of length 48 with independent, *t*-distributed values each with 5 degrees of freedom. Construct the time series plot. Use the same R code as in the solution of **Exercise 1.3** but replace `rnorm(n=48)` with `rt(n=48,df=5)`.

Exercise 1.6 Construct a time series plot with monthly plotting symbols for the Dubuque temperature series as in Exhibit (1.7), page 6. (Make the plot full screen so that you can see all of detail.)

```
> data(tempdub); plot(tempdub,ylab='Temperature')
> points(y=tempdub,x=time(tempdub), pch=as.vector(season(tempdub)))
```

CHAPTER 2

Exercise 2.1 Suppose $E(X) = 2$, $Var(X) = 9$, $E(Y) = 0$, $Var(Y) = 4$, and $Corr(X,Y) = 0.25$. Find:

(a) $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X,Y) = 9 + 4 + 2(3*2*0.25) = 16$

(b) $Cov(X, X + Y) = Cov(X,X) + Cov(X,Y) = 9 + ((3*2*0.25) = 9 + 3/2 = 10.5$

(c) $Corr(X + Y, X - Y)$. As in part (a), $Var(X-Y) = 10$. Then $Cov(X + Y, X - Y) = Cov(X,X) - Cov(Y,Y) + Cov(X,Y) - Cov(X,Y) = Var(X) - Var(Y) = 9 - 4 = 5$. So

$$Corr(X+Y, X-Y) = \frac{Cov(X+Y, X-Y)}{\sqrt{Var(X+Y)Var(X-Y)}} = \frac{5}{\sqrt{16 \times 10}} = \frac{5}{4\sqrt{10}} = 0.39528471$$

Exercise 2.2 If X and Y are dependent but $Var(X) = Var(Y)$, find $Cov(X + Y, X - Y)$.

$$Cov(X + Y, X - Y) = Cov(X,X) - Cov(Y,Y) + Cov(X,Y) - Cov(Y,X) = Var(X) - Var((Y) = 0$$

Exercise 2.3 Let X have a distribution with mean μ and variance σ^2 and let $Y_t = X$ for all t.

(a) Show that $\{Y_t\}$ is strictly and weakly stationary. Let $t_1, t_2,\dots, t_n$ be any set of time points and k any time lag.

Then

$$
\begin{aligned}
Pr(Y_{t_1} < y_{t_1}, Y_{t_2} \le y_{t_2}, \ldots, Y_{t_n} \le y_{t_n}) &= Pr(X < y_{t_1}, X \le y_{t_2}, \ldots, X \le y_{t_n}) \\
&= Pr(Y_{t_1-k} < y_{t_1}, Y_{t_2-k} \le y_{t_2}, \ldots, Y_{t_n-k} \le y_{t_n})
\end{aligned}
$$

as required for strict stationarity. Since the autocovariance clearly exists, (see part (b)), the process is also weakly stationary.

(b) Find the autocovariance function for $\{Y_t\}$. $Cov(Y_t,Y_{t-k}) = Cov(X,X) = \sigma^2$ for all t and k, free of t (and k).

(c) Sketch a "typical" time plot of Y_t. The plot will be a horizontal "line" (really a discrete-time horizontal line) at the height of the observed X.

Exercise 2.4 Let $\{e_t\}$ be a zero mean white noise processes. Suppose that the observed process is $Y_t = e_t + \theta e_{t-1}$ where θ is either 3 or 1/3.

(a) Find the autocorrelation function for $\{Y_t\}$ both when $\theta = 3$ and when $\theta = 1/3$. $E(Y_t) = E(e_t + \theta e_{t-1}) = 0$.

Also $Var(Y_t) = Var(e_t + \theta e_{t-1}) = \sigma^2 + \theta^2\sigma^2 = \sigma^2(1 + \theta^2)$. Also $Cov(Y_t,Y_{t-1}) = Cov(e_t + \theta e_{t-1}, e_{t-1} + \theta e_{t-2}) = \theta\sigma^2$ free of t. Now for $k > 1$, $Cov(Y_t,Y_{t-k}) = Cov(e_t + \theta e_{t-1}, e_{t-k} + \theta e_{t-k-1}) = 0$ since all of these error terms are uncorrelated. So

$$
Corr(Y_t, Y_{t-k}) = \frac{Cov(Y_t, Y_{t-k})}{\sqrt{Var(Y_t)Var(Y_{t-k})}} = \begin{cases} 1 & \text{for } k = 0 \\ \dfrac{\theta\sigma^2}{\sigma^2(1+\theta^2)} = \dfrac{\theta}{1+\theta^2} & \text{for } k = 1 \\ 0 & \text{for } k > 1 \end{cases}
$$

But $3/(1+3^2) = 3/10$ and $(1/3)/[1+(1/3)^2] = 3/10$. So the autocorrelation functions are identical.

(b) You should have discovered that the time series is stationary regardless of the value of θ and that the autocorrelation functions are the same for $\theta = 3$ and $\theta = 1/3$. For simplicity, suppose that the process mean is known to be zero and the variance of Y_t is known to be 1. You observe the series $\{Y_t\}$ for $t = 1, 2,\ldots, n$ and suppose that you can produce good estimates of the autocorrelations ρ_k. Do you think that you could determine which value of θ is correct (3 or 1/3) based on the estimate of ρ_k? Why or why not?

Exercise 2.5 Suppose $Y_t = 5 + 2t + X_t$ where $\{X_t\}$ is a zero mean stationary series with autocovariance function γ_k.

(a) Find the mean function for $\{Y_t\}$. $E(Y_t) = E(5 + 2t + X_t) = 5 + 2t + E(X_t) = 5 + 2t$.

(b) Find the autocovariance function for $\{Y_t\}$.

$$Cov(Y_t,Y_{t-k}) = Cov(5 + 2t + X_t, 5 + 2(t-k) + X_{t-k}) = Cov(X_t,X_{t-k}) = \gamma_k \text{ free of } t.$$

(c) Is $\{Y_t\}$ stationary? (Why or why not?) In spite of part (b), The process $\{Y_t\}$ is *not* stationary since its mean varies with time.

Exercise 2.6 Let $\{X_t\}$ be a stationary time series and define

$$
Y_t = \begin{cases} X_t & \text{for } t \text{ odd} \\ X_t + 3 & \text{for } t \text{ even} \end{cases}
$$

(a) Show that $Cov(Y_t, Y_{t-k})$ is free of t for all lags k.

$Cov(Y_t,Y_{t-k}) = Cov(X_t + 3, X_{t-k} + 3) = Cov(X_t,X_{t-k})$ is free of t since $\{X_t\}$ is stationary.

(b) Is $\{Y_t\}$ stationary? $\{Y_t\}$ is *not* stationary since $E(Y_t) = E(X_t) = \mu_X$ for t odd but $E(Y_t) = E(X_t + 3) = \mu_X + 3$ for t even.

Exercise 2.7 Suppose that $\{Y_t\}$ is stationary with autocovariance function γ_k.

(a) Show that $W_t = \nabla Y_t = Y_t - Y_{t-1}$ is stationary by finding the mean and autocovariance function for $\{W_t\}$.

$E(W_t) = E(Y_t - Y_{t-1}) = E(Y_t) - E(Y_{t-1}) = 0$ since $\{Y_t\}$ is stationary. Also

$$
\begin{aligned}
Cov(Y_t,Y_{t-k}) &= Cov(Y_t - Y_{t-1}, Y_{t-k} - Y_{t-k-1}) = Cov(Y_t,Y_{t-k}) - Cov(Y_t, Y_{t-k-1}) - Cov(Y_{t-1},Y_{t-k}) + Cov(Y_{t-1},Y_{t-k-1}) \\
&= \gamma_k - \gamma_{k+1} - \gamma_{k-1} + \gamma_k = 2\gamma_k - \gamma_{k+1} - \gamma_{k-1}, \text{ free of } t.
\end{aligned}
$$

(b) Show that $U_t = \nabla^2 Y_t = \nabla[Y_t - Y_{t-1}] = Y_t - 2Y_{t-1} + Y_{t-2}$ is stationary. (You need not find the mean and autocovariance function for $\{U_t\}$.) U_t is the first difference of the process $\{\nabla Y_t\}$. By part (a), $\{\nabla Y_t\}$ is stationary. So U_t is the difference of a stationary process and, again by part (a), is itself stationary.

Exercise 2.8 Suppose that $\{Y_t\}$ is stationary with autocovariance function γ_k. Show that for any fixed positive integer n and any constants $c_1, c_2,..., c_n$, the process $\{W_t\}$ defined by $W_t = c_1 Y_t + c_2 Y_{t-1} + \cdots + c_n Y_{t-n+1}$ is stationary. First

$$E(W_t) = c_1 EY_t + c_2 EY_{t-1} + \cdots + c_n EY_{t-n+1} = (c_1 + c_2 + \cdots + c_n)\mu_Y \text{ free of } t. \text{ Also}$$

$$\begin{aligned} Cov(W_t, W_{t-k}) &= Cov(c_1 Y_t + c_2 Y_{t-1} + \cdots + c_n Y_{t-n+1}, c_1 Y_{t-k} + c_2 Y_{t-1-k} + \cdots + c_n Y_{t-n+1-k}) \\ &= \sum_{j=0}^{n} \sum_{i=0}^{n} c_j c_i Cov(Y_{t-j}, Y_{t-k-i}) = \sum_{j=0}^{n} \sum_{i=0}^{n} c_j c_i \gamma_{j-k-i} \qquad \text{free of } t. \end{aligned}$$

Exercise 2.9 Suppose $Y_t = \beta_0 + \beta_1 t + X_t$ where $\{X_t\}$ is a zero mean stationary series with autocovariance function γ_k and β_0 and β_1 are constants.

(a) Show that $\{Y_t\}$ is not stationary but that $W_t = \nabla Y_t = Y_t - Y_{t-1}$ is stationary. $\{Y_t\}$ is *not* stationary since its mean, $\beta_0 + \beta_1 t$, varies with t. However, $E(W_t) = E(Y_t - Y_{t-1}) = (\beta_0 + \beta_1 t) - (\beta_0 + \beta_1(t-1)) = \beta_1$, free of t. The argument in the solution of Exercise 2.7 shows that the covariance function for $\{W_t\}$ is free of t.

(b) In general, show that if $Y_t = \mu_t + X_t$ where $\{X_t\}$ is a zero mean stationary series and μ_t is a polynomial in t of degree d, then $\nabla^m Y_t = \nabla(\nabla^{m-1} Y_t)$ is stationary for $m \geq d$ and nonstationary for $0 \leq m < d$.

Use part (a) and proceed by induction.

Exercise 2.10 Let $\{X_t\}$ be a zero-mean, unit-variance stationary process with autocorrelation function ρ_k. Suppose that μ_t is a nonconstant function and that σ_t is a positive-valued nonconstant function. The observed series is formed as $Y_t = \mu_t + \sigma_t X_t$

(a) Find the mean and covariance function for the $\{Y_t\}$ process.

Notice that $Cov(X_t,X_{t-k}) = Corr(X_t,X_{t-k})$ since $\{X_t\}$ has unit variance. $E(Y_t) = E(\mu_t + \sigma_t X_t) = \mu_t + \sigma_t E(X_t) = \mu_t$. Now $Cov(Y_t,Y_{t-k}) = Cov(\mu_t + \sigma_t X_t, \mu_{t-k} + \sigma_{t-k} X_{t-k}) = \sigma_t \sigma_{t-k} Cov(X_t,X_{t-k}) = \sigma_t \sigma_{t-k} \rho_k$. Notice that $Var(Y_t) = (\sigma_t)^2$.

(b) Show that autocorrelation function for the $\{Y_t\}$ process depends only on time lag. Is the $\{Y_t\}$ process stationary? $Corr(Y_t,Y_{t-k}) = \sigma_t \sigma_{t-k} \rho_k / [\sigma_t \sigma_{t-k}] = \rho_k$ but $\{Y_t\}$ is not necessarily stationary since $E(Y_t) = \mu_t$.

(c) Is it possible to have a time series with a constant mean and with $Corr(Y_t,Y_{t-k})$ free of t but with $\{Y_t\}$ *not* stationary? If μ_t is constant but σ_t varies with t, this will be the case.

Exercise 2.11 Suppose $Cov(X_t,X_{t-k}) = \gamma_k$ is free of t but that $E(X_t) = 3t$.

(a) Is $\{X_t\}$ stationary? No since $E(X_t)$ varies with t.

(b) Let $Y_t = 7 - 3t + X_t$. Is $\{Y_t\}$ stationary? Yes, since the covariances are unchanged but now $E(X_t) = 7 - 3t + 3t = 7$, free of t.

Exercise 2.12 Suppose that $Y_t = e_t - e_{t-12}$. Show that $\{Y_t\}$ is stationary and that, for $k > 0$, its autocorrelation function is nonzero only for lag $k = 12$.

$E(Y_t) = E(e_t - e_{t-12}) = 0$. Also $Cov(Y_t,Y_{t-k}) = Cov(e_t - e_{t-12}, e_{t-k} - e_{t-12-k}) = -Cov(e_{t-12}, e_{t-k}) = -(\sigma_e)^2$ when $k = 12$. It is nonzero only for $k = 12$ since, otherwise, all of the error terms involved are uncorrelated.

Exercise 2.13 Let $Y_t = e_t - \theta e_{t-1}^2$. For this exercise, assume that the white noise series is normally distributed.

(a) Find the autocorrelation function for $\{Y_t\}$. First recall that for a zero-mean normal distribution $E(e_{t-1}^3) = 0$ and $E(e_{t-1}^4) = 3\sigma_e^4$. Then $E(Y_t) = -\theta Var(e_{t-1}) = -\theta\sigma_e^2$ which is constant in t and

$$\begin{aligned} Var(Y_t) &= Var(e_t) + \theta^2 Var(e_{t-1}^2) = \sigma_e^2 + \theta^2\{E(e_{t-1}^4) - [E(e_{t-1}^2)]^2\} \\ &= \sigma_e^2 + \theta^2\{3\sigma_e^4 - [\sigma_e^2]^2\} \\ &= \sigma_e^2 + 2\theta^2\sigma_e^4 \end{aligned}$$

Also $Cov(Y_t, Y_{t-1}) = Cov(e_t - \theta e_{t-1}^2, e_{t-1} - \theta e_{t-2}^2) = Cov(-\theta e_{t-1}^2, e_{t-1}) = -\theta E(e_{t-1}^3) = 0$

All other covariances are also zero.

(b) Is $\{Y_t\}$ stationary? Yes, in fact, it is a non-normal white noise in disguise!

Exercise 2.14 Evaluate the mean and covariance function for each of the following processes. In each case determine whether or not the process is stationary.

(a) $Y_t = \theta_0 + te_t$. The mean is θ_0 but it is not stationary since $Var(Y_t) = t^2 Var(e_t) = t^2\sigma^2$ is not free of t.

(b) $W_t = \nabla Y_t$ where Y_t is as given in part (a). $W_t = \nabla Y_t = (\theta_0 + te_t) - (\theta_0 + (t-1)e_{t-1}) = te_t - (t-1)e_{t-1}$

So the mean of W_t is zero. However, $Var(W_t) = [t^2 + (t-1)^2](\sigma_e)^2$ which depends on t and W_t is not stationary.

(c) $Y_t = e_t e_{t-1}$. (You may assume that $\{e_t\}$ is normal white noise.) The mean of Y_t is clearly zero. Lag one is the only lag at which there might be correlation. However, $Cov(Y_t, Y_{t-1}) = E(e_t e_{t-1} e_{t-1} e_{t-2})$ $= E(e_t)\, E[e_{t-1}]^2 E(e_{t-2}) = 0$. So the process $Y_t = e_t e_{t-1}$ is stationary and is a non-normal white noise!

Exercise 2.15 Suppose that X is a random variable with zero mean. Define a time series by $Y_t = (-1)^t X$.

(a) Find the mean function for $\{Y_t\}$. $E(Y_t) = (-1)^t E(X) = 0$.

(b) Find the covariance function for $\{Y_t\}$. $Cov(Y_t, Y_{t-k}) = Cov[(-1)^t X, (-1)^{t-k} X] = (-1)^{2t-k} Cov(X,X) = (-1)^k (\sigma_X)^2$

(c) Is $\{Y_t\}$ stationary? Yes, the mean is constant and the covariance only depends on lag.

Exercise 2.16 Suppose $Y_t = A + X_t$ where $\{X_t\}$ is stationary and A is random but independent of $\{X_t\}$. Find the mean and covariance function for $\{Y_t\}$ in terms of the mean and autocovariance function for $\{X_t\}$ and the mean and variance of A. First $E(Y_t) = E(A) + E(X_t) = \mu_A + \mu_X$, free of t. Also, since $\{X_t\}$ and A are independent,

$$Cov(Y_t, Y_{t-k}) = Cov(A + X_t, A + X_{t-k}) = Cov(A, A) + Cov(X_t, X_{t-k}) = Var(A) + \gamma_k^X \qquad \text{free of } t$$

Exercise 2.17 Let $\{Y_t\}$ be stationary with autocovariance function γ_k. Let $\bar{Y} = \frac{1}{n}\sum_{t=1}^{n} Y_t$. Show that

$$Var(\bar{Y}) = \frac{\gamma_0}{n} + \frac{2}{n}\sum_{k=1}^{n-1}\left(1 - \frac{k}{n}\right)\gamma_k$$

$$= \frac{1}{n}\sum_{k=-n+1}^{n-1}\left(1 - \frac{|k|}{n}\right)\gamma_k$$

$$Var(\bar{Y}) = \frac{1}{n^2} Var\left[\sum_{t=1}^{n} Y_t\right] = \frac{1}{n^2} Cov\left[\sum_{t=1}^{n} Y_t, \sum_{s=1}^{n} Y_s\right] = \frac{1}{n^2}\sum_{t=1}^{n}\sum_{s=1}^{n}\gamma_{t-s}$$

Now make the change of variable $t - s = k$ and $t = j$ in the double sum. The range of the summation $\{1 \le t \le n,\ 1 \le s \le n\}$ is transformed into $\{1 \le j \le n,\ 1 \le j - k \le n\} = \{k + 1 \le j \le n + k,\ 1 \le j \le n\}$ which may be written $\{k > 0, k + 1 \le j \le n\} \cup \{k \le 0, 1 \le j \le n + k\}$. Thus

$$Var(\bar{Y}) = \frac{1}{n^2}\left[\sum_{k=1}^{n-1}\sum_{j=k+1}^{n}\gamma_k + \sum_{k=-n+1}^{0}\sum_{j=1}^{n+k}\gamma_k\right]$$

$$= \frac{1}{n^2}\left[\sum_{k=1}^{n-1}(n-k)\gamma_k + \sum_{k=-n+1}^{0}(n+k)\gamma_k\right]$$

$$= \frac{1}{n}\sum_{k=-n+1}^{n-1}\left(1 - \frac{|k|}{n}\right)\gamma_k$$

Use $\gamma_k = \gamma_{-k}$ to get the first expression in the exercise.

Exercise 2.18 Let $\{Y_t\}$ be stationary with autocovariance function γ_k. Define the sample variance as

$$S^2 = \frac{1}{n-1}\sum_{t=1}^{n}(Y_t - \bar{Y})^2 .$$

(a) First show that $\sum_{t=1}^{n}(Y_t - \mu)^2 = \sum_{t=1}^{n}(Y_t - \bar{Y})^2 + n(\bar{Y} - \mu)^2$.

$$\sum_{t=1}^{n}(Y_t - \mu)^2 = \sum_{t=1}^{n}(Y_t - \bar{Y} + \bar{Y} - \mu)^2 = \sum_{t=1}^{n}(Y_t - \bar{Y})^2 + \sum_{t=1}^{n}(\bar{Y} - \mu)^2 + 2\sum_{t=1}^{n}(Y_t - \bar{Y})(\bar{Y} - \mu)$$

$$= \sum_{t=1}^{n}(Y_t - \bar{Y})^2 + n(\bar{Y} - \mu)^2 + 2(\bar{Y} - \mu)\sum_{t=1}^{n}(Y_t - \bar{Y}) = \sum_{t=1}^{n}(Y_t - \bar{Y})^2 + n(\bar{Y} - \mu)^2$$

(b) Use part (a) to show that $E(S^2) = \frac{n}{n-1}\gamma_0 - \frac{n}{n-1}Var(\bar{Y}) = \gamma_0 - \frac{2}{n-1}\sum_{k=1}^{n-1}\left(1 - \frac{k}{n}\right)\gamma_k$.
(Use the results of Exercise (2.17) for the last expression.)

$$E(S^2) = E\left(\frac{1}{n-1}\sum_{t=1}^{n}(Y_t-\bar{Y})^2\right) = \frac{1}{n-1}E\left(\sum_{t=1}^{n}(Y_t-\mu)^2 - n(\bar{Y}-\mu)^2\right) = \frac{1}{n-1}\left[\sum_{t=1}^{n}E[(Y_t-\mu)^2] - nE(\bar{Y}-\mu)^2\right]$$

$$= \frac{1}{n-1}[n\gamma_0 - nVar(\bar{Y})] = \frac{1}{n-1}\left[n\gamma_0 - n\left\{\frac{\gamma_0}{n}+\frac{2}{n}\sum_{k=1}^{n-1}\left(1-\frac{k}{n}\right)\gamma_k\right\}\right] = \gamma_0 - \frac{2}{n-1}\sum_{k=1}^{n-1}\left(1-\frac{k}{n}\right)\gamma_k$$

- **(c)** If $\{Y_t\}$ is a white noise process with variance γ_0, show that $E(S^2) = \gamma_0$. This follows since for white noise $\gamma_k = 0$ for $k > 0$.

Exercise 2.19 Let $Y_1 = \theta_0 + e_1$ and then for $t > 1$ define Y_t recursively by $Y_t = \theta_0 + Y_{t-1} + e_t$. Here θ_0 is a constant. The process $\{Y_t\}$ is called a **random walk with drift**.

- **(a)** Show that Y_t may be rewritten as $Y_t = t\theta_0 + e_t + e_{t-1} + \cdots + e_1$. Substitute $Y_{t-1} = \theta_0 + Y_{t-2} + e_{t-1}$ into $Y_t = \theta_0 + Y_{t-1} + e_t$ and repeat until you get back to e_1.
- **(b)** Find the mean function for Y_t. $E(Y_t) = E(t\theta_0 + e_t + e_{t-1} + \cdots + e_1) = t\theta_0$
- **(c)** Find the autocovariance function for Y_t.

$$\begin{aligned} Cov(Y_t, Y_{t-k}) &= Cov[t\theta_0 + e_t + e_{t-1} + \cdots + e_1, (t-k)\theta_0 + e_{t-k} + e_{t-1-k} + \cdots + e_1] \\ &= Cov[e_{t-k} + e_{t-1-k} + \cdots + e_1, e_{t-k} + e_{t-1-k} + \cdots + e_1] \\ &= Var(e_{t-k} + e_{t-1-k} + \cdots + e_1) = (t-k)\sigma_e^2 \qquad \text{for } t \ge k \end{aligned}$$

Exercise 2.20 Consider the standard random walk model where $Y_t = Y_{t-1} + e_t$ with $Y_1 = e_1$.

- **(a)** Use the above representation of Y_t to show that $\mu_t = \mu_{t-1}$ for $t > 1$ with initial condition $\mu_1 = E(e_1) = 0$. Hence show that $\mu_t = 0$ for all t. Clearly, $\mu_1 = E(Y_1) = E(e_1) = 0$. Then $E(Y_t) = E(Y_{t-1} + e_t) = E(Y_{t-1}) + E(e_t) = E(Y_{t-1})$ or $\mu_t = \mu_{t-1}$ for $t > 1$ and the result follows by induction.
- **(b)** Similarly, show that $Var(Y_t) = Var(Y_{t-1}) + \sigma_e^2$, for $t > 1$ with $Var(Y_1) = \sigma_e^2$, and, hence $Var(Y_t) = t\sigma_e^2$.

$Var(Y_1) = \sigma_e^2$ is immediate. Then $Var(Y_t) = Var(Y_{t-1} + e_t) = Var(Y_{t-1}) + Var(e_t) = Var(Y_{t-1}) + \sigma_e^2$. Recursion or induction on t yields $Var(Y_t) = t\sigma_e^2$.

- **(c)** For $0 \le t \le s$, use $Y_s = Y_t + e_{t+1} + e_{t+2} + \cdots + e_s$ to show that $Cov(Y_t, Y_s) = Var(Y_t)$ and, hence, that $Cov(Y_t, Y_s) = \min(t, s)\sigma_e^2$. For $0 \le t \le s$,

$Cov(Y_t, Y_s) = Cov(Y_t, Y_t + e_{t+1} + e_{t+2} + \ldots + e_s) = Cov(Y_t, Y_t) = Var(Y_t) = t\sigma_e^2$ and hence the result.

Exercise 2.21 A random walk with random starting value. Let $Y_t = Y_0 + e_t + e_{t-1} + \cdots + e_1$ for $t > 0$ where Y_0 has a distribution with mean μ_0 and variance σ_0^2. Suppose further that $Y_0, e_1, \ldots, e_t$ are independent.

- **(a)** Show that $E(Y_t) = \mu_0$ for all t.

$E(Y_t) = E(Y_0 + e_t + e_{t-1} + \cdots + e_1) = E(Y_0) + E(e_t) + E(e_{t-1}) + \cdots + E(e_1) = E(Y_0) = \mu_0$.

- **(b)** Show that $Var(Y_t) = t\sigma_e^2 + \sigma_0^2$.

$Var(Y_t) = Var(Y_0 + e_t + e_{t-1} + \cdots + e_1) = Var(Y_1) + Var(e_t) + Var(e_{t-1}) + \ldots + Var(e_1) = \sigma_0^2 + t\sigma_e^2$

- **(c)** Show that $Cov(Y_t, Y_s) = \min(t, s)\sigma_e^2 + \sigma_0^2$. Let t be less than s. Then, as in the previous exercise,

$$Cov(Y_t, Y_s) = Cov(Y_t, Y_t + e_{t+1} + e_{t+2} + \ldots + e_s) = Var(Y_t) = \sigma_0^2 + t\sigma_e^2$$

- **(d)** Show that $Corr(Y_t, Y_s) = \sqrt{\frac{t\sigma_a^2 + \sigma_0^2}{s\sigma_a^2 + \sigma_0^2}}$ for $0 \le t \le s$. Just use the results of parts (b) and (c).

Exercise 2.22 Let $\{e_t\}$ be a zero-mean white noise process and let c be a constant with $|c| < 1$. Define Y_t recursively by $Y_t = cY_{t-1} + e_t$ with $Y_1 = e_1$.

This exercise can be solved using the recursive definition of Y_t or by expressing Y_t explicitly using repeated substitution as $Y_t = c(cY_{t-2} + e_{t-1}) + e_t = \cdots = e_t + ce_{t-1} + c^2e_{t-2} + \cdots + c^{t-1}e_1$. Parts (c), (d), and (e) essentially assume you are working with the recursive version of Y_t but they can also be solved using this explicit representation.

(a) Show that $E(Y_t) = 0$. First $E(Y_1) = E(e_1) = 0$. Then $E(Y_t) = cE(Y_{t-1}) + E(e_t) = cE(Y_{t-1})$ and the result follows by induction on t.

(b) Show that $Var(Y_t) = \sigma_e^2 (1 + c^2 + c^4 + \cdots + c^{2t-2})$. Is $\{Y_t\}$ stationary?

$Var(Y_t) = Var(e_t + ce_{t-1} + c^2 e_{t-2} + \cdots + c^{t-1} e_1) = \sigma_e^2(1 + c^2 + c^4 + \cdots + c^{2(t-1)}) = \sigma_e^2\left(\frac{1-c^{2t}}{1-c^2}\right)$. $\{Y_t\}$ is *not* stationary since $Var(Y_t)$ depends on t.

Alternatively,
$$\begin{aligned} Var(Y_t) &= Var(cY_{t-1} + e_t) = c^2 Var(Y_{t-1}) + \sigma_e^2 = c^2 Var(cY_{t-2} + e_{t-1}) + \sigma_e^2 \\ &= c^3 Var(Y_{t-2}) + \sigma_e^2(1 + c^2) = \cdots = \sigma_e^2(1 + c^2 + c^4 + \cdots + c^{2(t-1)}) \end{aligned}$$

(c) Show that $Corr(Y_t, Y_{t-1}) = c\sqrt{\frac{Var(Y_{t-1})}{Var(Y_t)}}$ and, in general,

$$Corr(Y_t, Y_{t-k}) = c^k \sqrt{\frac{Var(Y_{t-k})}{Var(Y_t)}} \qquad \text{for } k > 0$$

(Hint: Argue that Y_{t-1} is independent of e_t. Then use $Cov(Y_t, Y_{t-1}) = Cov(cY_{t-1} + e_t, Y_{t-1})$

$$Cov(Y_t, Y_{t-1}) = Cov(cY_{t-1} + e_t, Y_{t-1}) = cVar(Y_{t-1})$$

So $Corr(Y_t, Y_{t-1}) = \dfrac{cVar(Y_{t-1})}{\sqrt{Var(Y_t)Var(Y_{t-1})}} = c\sqrt{\dfrac{Var(Y_{t-1})}{Var(Y_t)}}$. Next

$$\begin{aligned} Cov(Y_t, Y_{t-k}) &= Cov(cY_{t-1} + e_t, Y_{t-k}) = cCov(cY_{t-2} + e_{t-1}, Y_{t-k}) = c^2 Cov(Y_{t-2}, Y_{t-k}) \\ &= c^2 Cov(cY_{t-3} + e_{t-2}, Y_{t-k}) = \cdots = c^k Var(Y_{t-k}) \end{aligned}$$

So $Corr(Y_t, Y_{t-k}) = \dfrac{c^k Var(Y_{t-k})}{\sqrt{Var(Y_t)Var(Y_{t-k})}} = c^k\sqrt{\dfrac{Var(Y_{t-k})}{Var(Y_t)}}$ as required.

(d) For large t, argue that

$$Var(Y_t) \approx \frac{\sigma_e^2}{1-c^2} \qquad \text{and} \qquad Corr(Y_t, Y_{t-k}) \approx c^k \qquad \text{for } k > 0$$

so that $\{Y_t\}$ could be called **asymptotically stationary**. These two results follow from parts (b) and (c).

(e) Suppose now that we alter the initial condition and put $Y_1 = e_1/(\sqrt{1-c^2})$. Show that now $\{Y_t\}$ is stationary. This part can be solved using repeated substitution to express Y_t explicitly as

$$Y_t = c(cY_{t-2} + e_{t-1}) + e_t = \cdots = e_t + ce_{t-1} + c^2 e_{t-2} + \cdots + c^{t-2} e_2 + \frac{c^{t-1}}{\sqrt{1-c^2}} e_1$$

Then show that $Var(Y_t) = \dfrac{\sigma_e^2}{1-c^2}$ and $Corr(Y_t, Y_{t-k}) = c^k$ for $k > 0$.

Exercise 2.23 Two processes $\{Z_t\}$ and $\{Y_t\}$ are said to be **independent** if for any time points $t_1, t_2, \ldots, t_m$ and $s_1, s_2, \ldots, s_n$, the random variables $\{Z_{t_1}, Z_{t_2}, \ldots, Z_{t_m}\}$ are independent of the random variables $\{Y_{s_1}, Y_{s_2}, \ldots, Y_{s_n}\}$. Show that if $\{Z_t\}$ and $\{Y_t\}$ are independent stationary processes, then $W_t = Z_t + Y_t$ is stationary.

First, $E(W_t) = E(Z_t) + E(Y_t) = \mu_Z + \mu_Y$. Then $Cov(W_t, W_{t-k}) = Cov(Z_t + Y_t, Z_{t-k} + Y_{t-k}) = Cov(Z_t, Z_{t-k}) + Cov(Y_t, Y_{t-k})$ which is free of t since both $\{Z_t\}$ and $\{Y_t\}$ are stationary.

Exercise 2.24 Let $\{X_t\}$ be a time series in which we are interested. However, because the measurement process itself is not perfect, we actually observe $Y_t = X_t + e_t$. We assume that $\{X_t\}$ and $\{e_t\}$ are independent processes. We call X_t the **signal** and e_t the **measurement noise** or **error process**.

If $\{X_t\}$ is stationary with autocorrelation function ρ_k, show that $\{Y_t\}$ is also stationary with

$$Corr(Y_t, Y_{t-k}) = \frac{\rho_k}{1 + \sigma_e^2/\sigma_X^2} \qquad \text{for } k \ge 1$$

We call σ_X^2/σ_e^2 the **signal-to-noise ratio**, or SNR. Note that the larger the SNR, the closer the autocorrelation function of the observed process $\{Y_t\}$ is to the autocorrelation function of the desired signal $\{X_t\}$.

First, $E(Y_t) = E(X_t) + E(e_t) = \mu_X$ free of t. Next, for $k \geq 1$, $Cov(Y_t,Y_{t-k}) = Cov(X_t + e_t, X_{t-k} + e_{t-k}) = Cov(X_t,X_{t-k}) + Cov(e_t,e_{t-k}) = Cov(X_t,X_{t-k}) = Var(X_t)\rho_k$ which is free of t. Finally,

$$Corr(Y_t, Y_{t-k}) = \frac{Cov(Y_t, Y_{t-k})}{Var(Y_t)} = \frac{Var(X_t)\rho_k}{Var(X_t) + \sigma_e^2} = \frac{\sigma_X^2\rho_k}{\sigma_X^2 + \sigma_e^2} = \frac{\rho_k}{1 + \sigma_e^2/\sigma_X^2} \qquad \text{for } k \geq 1$$

Exercise 2.25 Suppose $Y_t = \beta_0 + \sum_{i=1}^{k} [A_i\cos(2\pi f_i t) + B_i\sin(2\pi f_i t)]$ where $\beta_0, f_1, f_2, \ldots, f_k$ are constants and $A_1, A_2, \ldots, A_k, B_1, B_2, \ldots, B_k$ are independent random variables with zero means and variances $Var(A_i) = Var(B_i) = \sigma_i^2$. Show that $\{Y_t\}$ is stationary and find its covariance function. Compare this exercise with the results for the Random Cosine wave on page 18. First

$$E(Y_t) = \beta_0 + \sum_{i=1}^{k} [E(A_i)\cos(2\pi f_i t) + E(B_i)\sin(2\pi f_i t)] = \beta_0$$

Next using the independence of $A_1, A_2, \ldots, A_k, B_1, B_2, \ldots, B_k$ and some trig identities we have

$$Cov(Y_t, Y_s) = Cov\left\{\sum_{i=1}^{k} [A_i\cos(2\pi f_i t) + B_i\sin(2\pi f_i t)], \sum_{j=1}^{k} [A_j\cos(2\pi f_j s) + B_j\sin(2\pi f_j s)]\right\}$$

$$= \sum_{i=1}^{k} Cov\{A_i\cos((2\pi f_i t), A_i\cos(2\pi f_i s))\} + \sum_{i=1}^{k} Cov\{B_i\sin(2\pi f_j t), B_i\sin(2\pi f_j s)\}$$

$$= \sum_{i=1}^{k} \{\cos(2\pi f_i t)\cos(2\pi f_i s)\} Var(A_i) + \sum_{i=1}^{k} \{\sin(2\pi f_j t)\sin(2\pi f_j s)\} Var(B_i)$$

$$= \sum_{i=1}^{k} \{\cos(2\pi f_i(t-s)) + \cos(2\pi f_i(t+s))\}\frac{\sigma_i^2}{2} + \sum_{i=1}^{k} \{\cos(2\pi f_i(t-s)) - \cos(2\pi f_i(t+s))\}\frac{\sigma_i^2}{2}$$

$$= \sum_{i=1}^{k} \cos(2\pi f_i(t-s))\sigma_i^2 \qquad \text{hence the process is stationary.}$$

Exercise 2.26 Define the function $\Gamma_{t,s} = \frac{1}{2}E[(Y_t - Y_s)^2]$. In geostatistics, $\Gamma_{t,s}$ is called the *semivariogram*.

(a) Show that for a stationary process $\Gamma_{t,s} = \gamma_0 - \gamma_{|t-s|}$. Without loss of generality, we may assume that the stationary process has a zero mean. Then

$$\Gamma_{t,s} = \tfrac{1}{2}E[(Y_t - Y_s)^2] = \tfrac{1}{2}E[Y_t^2 - 2Y_tY_s + Y_s^2] = \tfrac{1}{2}E[Y_t^2] + \tfrac{1}{2}E[Y_s^2] - \tfrac{1}{2}E[2Y_tY_s] = \gamma_0 - \gamma_{|t-s|}$$

(b) A process is said to be *intrinsically stationary* if $\Gamma_{t,s}$ depends only on the time difference $|t-s|$. Show that the random walk process is intrinsically stationary.

For the random walk for $t > s$ we have $Y_t = e_t + e_{t-1} + \cdots + e_1$ so that

$$Y_t - Y_s = (e_t + e_{t-1} + \cdots + e_1) - (e_s + e_{s-1} + \cdots + e_1) = (e_t + e_{t-1} + \cdots + e_{s+1}) \text{ and}$$

$$\Gamma_{t,s} = \tfrac{1}{2}E[(Y_t - Y_s)^2] = \tfrac{1}{2}Var(Y_t - Y_s) = \tfrac{1}{2}Var(e_t + e_{t-1} + \cdots + e_{s+1}) = \tfrac{1}{2}(t-s)\sigma_e^2 \text{ as required.}$$

Exercise 2.27 For a fixed, positive integer r and constant ϕ, consider the time series defined by

$$Y_t = e_t + \phi e_{t-1} + \phi^2 e_{t-2} + \cdots + \phi^r e_{t-r}.$$

(a) Show that this process is stationary for any value of ϕ. The mean is clearly zero and

$$Cov(Y_t, Y_{t-k}) = Cov(e_t + \phi e_{t-1} + \phi^2 e_{t-2} + \cdots + \phi^r e_{t-r}, e_{t-k} + \phi e_{t-k-1} + \phi^2 e_{t-k-2} + \cdots + \phi^r e_{t-k-r})$$
$$= Cov(e_t + \cdots + \phi^k e_{t-k} + \phi^{k+1} e_{t-k-1} + \cdots + \phi^r e_{t-r}, e_{t-k} + \phi e_{t-k-1} + \cdots + \phi^r e_{t-k-r})$$
$$= (\phi^k + \phi^{k+2} + \phi^{k+4} + \ldots + \phi^{k+2(r-k)})\sigma_e^2 = (1 + \phi^2 + \phi^4 + \ldots + \phi^{2(r-k)})\sigma_e^2\phi^k$$

(b) Find the autocorrelation function. We have

$$Var(Y_t) = Var(e_t + \phi e_{t-1} + \phi^2 e_{t-2} + \cdots + \phi^r e_{t-r}) = (1 + \phi^2 + \phi^4 + \ldots + \phi^{2r})\sigma_e^2 \text{ so that}$$

$$Corr(Y_t, Y_{t-k}) = \frac{(1 + \phi^2 + \phi^4 + \ldots + \phi^{2(r-k)})\phi^k}{(1 + \phi^2 + \phi^4 + \ldots + \phi^{2r})}$$

The results in parts (a) and (b) can be simplified for $\phi \neq 1$ and separately for $\phi = 1$.

Exercise 2.28 (Random cosine wave extended) Suppose that

$$Y_t = R\cos(2\pi(ft + \Phi)) \qquad \text{for } t = 0, \pm1, \pm2, \ldots$$

where $0 < f < ½$ is a fixed frequency and R and Φ are uncorrelated random variables and with Φ uniformly distributed on the interval (0,1).

(a) Show that $E(Y_t) = 0$ for all t. $E(Y_t) = E\{R\cos(2\pi(ft + \Phi))\} = E(R)E\{\cos(2\pi(ft + \Phi))\}$ But $E\{\cos(2\pi(ft + \Phi))\} = 0$ with a calculation entirely similar to the one on page 18.

(b) Show that the process is stationary with $\gamma_k = \frac{1}{2}E(R^2)\cos(2\pi fk)$

$$\gamma_k = E[R^2\cos(2\pi(f(t-k) + \Phi))\cos(2\pi(f(t-k) + \Phi))] = E(R^2)E[\cos(2\pi(ft + \Phi))\cos(2\pi(f(t-k) + \Phi))]$$

and then use the calculations leading up to Equation (2.3.4), page 19 to show that

$$E[\cos(2\pi(ft + \Phi))\cos(2\pi(f(t-k) + \Phi))] = \cos(2\pi fk)$$

Exercise 2.29 (Random cosine wave extended more) Suppose that

$$Y_t = \sum_{j=1}^{m} R_j\cos[2\pi(f_j t + \Phi_j)] \qquad \text{for } t = 0, \pm1, \pm2, \ldots$$

where $0 < f_1 < f_2 < \cdots < f_m < ½$ are m fixed frequencies, $R_1, \Phi_1, R_2, \Phi_2, \ldots, R_m, \Phi_m$ are uncorrelated random variables and with each Φ_j uniformly distributed on the interval (0,1).

(a) Show that $E(Y_t) = 0$ for all t.

(b) Show that the process is stationary with $\gamma_k = \frac{1}{2}\sum_{j=1}^{m} E(R_j^2)\cos(2\pi f_j k)$.

Parts (a) and (b) follow directly from the solution of Exercise (2.28) using the independence.

Exercise 2.30 (Mathematical statistics required) Suppose that

$$Y_t = R\cos[2\pi(ft + \Phi)] \qquad \text{for } t = 0, \pm1, \pm2, \ldots$$

where R and Φ are independent random variables and f is a fixed frequency. The phase Φ is assumed to be uniformly distributed on $(0,1)$, and the amplitude R has a Rayleigh distribution with pdf $f(r) = re^{-r^2/2}$ for $r > 0$. Show that for each time point t, Y_t has a normal distribution. (Hint: Let $Y = R\cos[2\pi(ft + \Phi)]$ and $X = R\sin[2\pi(ft + \Phi)]$. Now find the joint distribution of X and Y. It can also be shown that all of the finite dimensional distributions are multivariate normal and hence the process is strictly stationary.)

For fixed t and f consider the one-to-one transformation defined by

$$Y = R\cos[2\pi(ft + \Phi)], X = R\sin[2\pi(ft + \Phi)]$$

The range for (X,Y) will be $\{-\infty < X < \infty, -\infty < Y < \infty\}$. Also $X^2 + Y^2 = R^2$. Furthermore,

$$\begin{bmatrix} \frac{\partial X}{\partial R} & \frac{\partial X}{\partial \Phi} \\ \frac{\partial Y}{\partial R} & \frac{\partial Y}{\partial \Phi} \end{bmatrix} = \begin{bmatrix} \cos[2\pi(ft + \Phi)] & 2\pi R\sin[2\pi(ft + \Phi)] \\ \sin[2\pi(ft + \Phi)] & 2\pi R\cos[2\pi(ft + \Phi)] \end{bmatrix} \text{ and the Jacobian is } -2\pi R = -2\pi\sqrt{X^2 + Y^2}$$

with inverse Jacobian $1/(-2\pi\sqrt{X^2 + Y^2})$. The joint density for R and Φ is $f(r, \phi) = re^{-r^2/2}$ for $0 < r$ and $0 < \phi < 1$. Hence the joint density for X and Y is given by

$$f(x, y) = \frac{\sqrt{x^2+y^2}e^{-(x^2+y^2)/2}}{2\pi\sqrt{x^2+y^2}} = \frac{e^{-x^2/2}}{\sqrt{2\pi}}\frac{e^{-y^2/2}}{\sqrt{2\pi}} \text{ for } -\infty < x < \infty, -\infty < y < \infty \text{ as required.}$$

CHAPTER 3

Exercise 3.1 Verify Equation (3.3.2), page 30, for the least squares estimates of β_0 and of β_1 when the model $Y_t = \beta_0 + \beta_1 t + X_t$ is considered. This is a standard calculation in many first courses in statistics usually using calculus. Here we give an algebraic solution. Without loss of generality, we assume that Y_t and t have each been standardized so that $\Sigma Y_t = \Sigma t = 0$ and $\Sigma(Y_t)^2 = \Sigma t^2 = n - 1$. Then we have

$$\begin{aligned} Q(\beta_0, \beta_1) &= \sum_{t=1}^{n}[Y_t - (\beta_0 + \beta_1 t)]^2 \\ &= n\beta_0^2 + \sum_{t=1}^{n}[Y_t]^2 + \beta_1^2\sum_{t=1}^{n}t^2 - 2\beta_0\sum_{t=1}^{n}Y_t + 2\beta_0\beta_1\sum_{t=1}^{n}t - 2\beta_1\sum_{t=1}^{n}tY_t \\ &= n\beta_0^2 + \sum_{t=1}^{n}[Y_t]^2 + \beta_1^2\sum_{t=1}^{n}t^2 - 2\beta_1\sum_{t=1}^{n}tY_t \\ &= n\beta_0^2 + (n-1)(1+\beta_1^2) - 2\beta_1\sum_{t=1}^{n}tY_t \\ &= n\beta_0^2 + (n-1) + (n-1)\left[\beta_1 - \frac{1}{(n-1)}\sum_{t=1}^{n}tY_t\right]^2 - (n-1)\left[\frac{1}{(n-1)}\sum_{t=1}^{n}tY_t\right]^2 \end{aligned}$$

This is clearly smallest when $\hat{\beta}_0 = 0$ and

$$\hat{\beta}_1 = \frac{1}{(n-1)}\sum_{t=1}^{n}tY_t$$

When these results are translated back to (unstandardized) original terms, we obtain the usual ordinary least squares regression results. In addition, by looking at the minimum value of Q we have

$$Q(\hat{\beta}_0, \hat{\beta}_1) = (n-1)(1-r^2)$$

where r is the correlation coefficient between Y and t. Since $Q \geq 0$, this also provides a proof that correlations are always between −1 and +1.

Exercise 3.2 Suppose $Y_t = \mu + e_t - e_{t-1}$. Find $Var(\overline{Y})$. Note any unusual results. In particular, compare your answer to what would have been obtained if $Y_t = \mu + e_t$. (Hint: You may avoid Equation (3.2.3), page 28, by first doing some algebraic simplification on $\sum_{t=1}^{n}(e_t - e_{t-1})$.)

$$\overline{Y} = \left[\mu + \frac{1}{n}\sum_{t=1}^{n}(e_t - e_{t-1})\right] = \mu + \frac{1}{n}(e_n - e_0) \text{ so } Var(\overline{Y}) = \frac{1}{n^2}Var(e_n - e_0) = \frac{2}{n^2}\sigma_e^2$$

The denominator of n^2 is very unusual. We expect a denominator of n in the variance of a sample mean. The negative autocorrelation at lag one makes it easier to estimate the process mean when compared with estimating the mean of a white noise process.

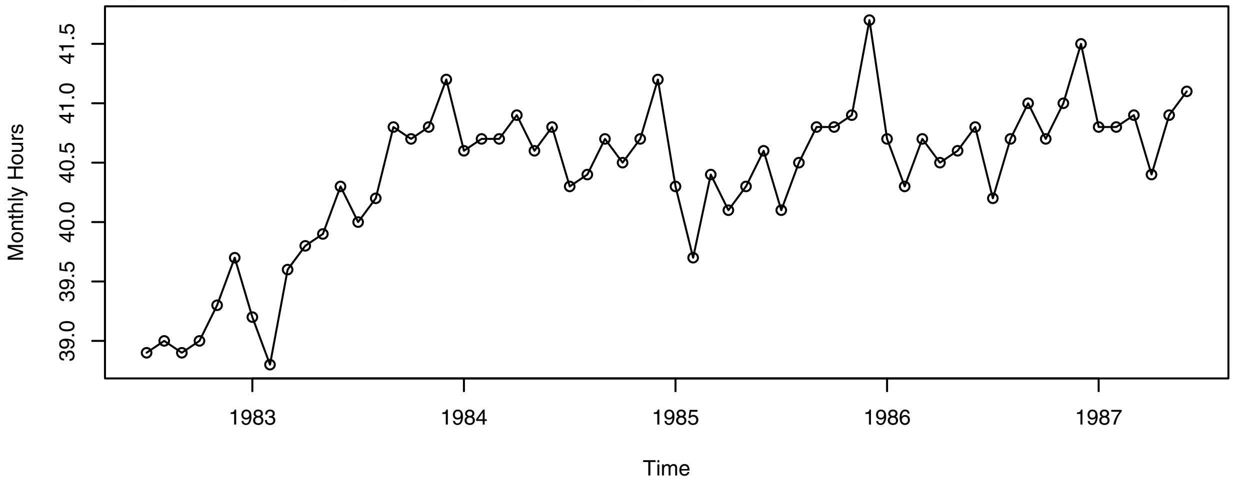
Exercise 3.3 Suppose $Y_t = \mu + e_t + e_{t-1}$. Find $Var(\overline{Y})$. Compare your answer to what would have been obtained if $Y_t = \mu + e_t$. Describe the effect that the autocorrelation in $\{Y_t\}$ has on $Var(\overline{Y})$.

$$\sum_{t=1}^{n}(e_t + e_{t-1}) = e_n + e_0 + 2\sum_{t=1}^{n-1}e_t \text{ so } Var(\overline{Y}) = \frac{1}{n^2}[\sigma_e^2 + \sigma_e^2 + 4(n-1)\sigma_e^2] = \frac{2(2n-1)}{n^2}\sigma_e^2$$

If $Y_t = \mu + e_t$ we would have $Var(\overline{Y}) = (1/n)\sigma_e^2$ but in our present case $Var(\overline{Y}) \approx (4/n)\sigma_e^2$, approximately four times larger. The positive autocorrelation at lag one makes it more difficult to estimate the process mean compared with estimating the mean of a white noise process.

Exercise 3.4 The data file hours contains monthly values of the average hours worked per week in the U.S. manufacturing sector for July 1982 through June 1987.

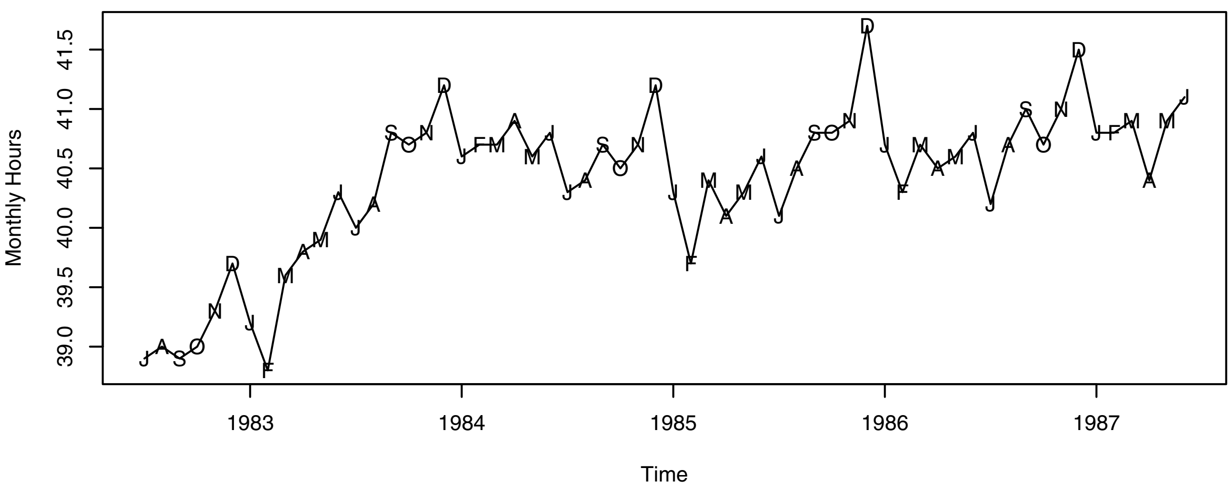
(a) Display the time series plot for these data and interpret.




```
> data(hours); plot(hours,ylab='Monthly Hours',type='o')
```

The plot displays an upward "trend," in the first half of the series. However, there is certainly no distinct pattern in the display.

(b) Now construct a time series plot that uses separate plotting symbols for the various months. Does your interpretation change from that in part (a)?

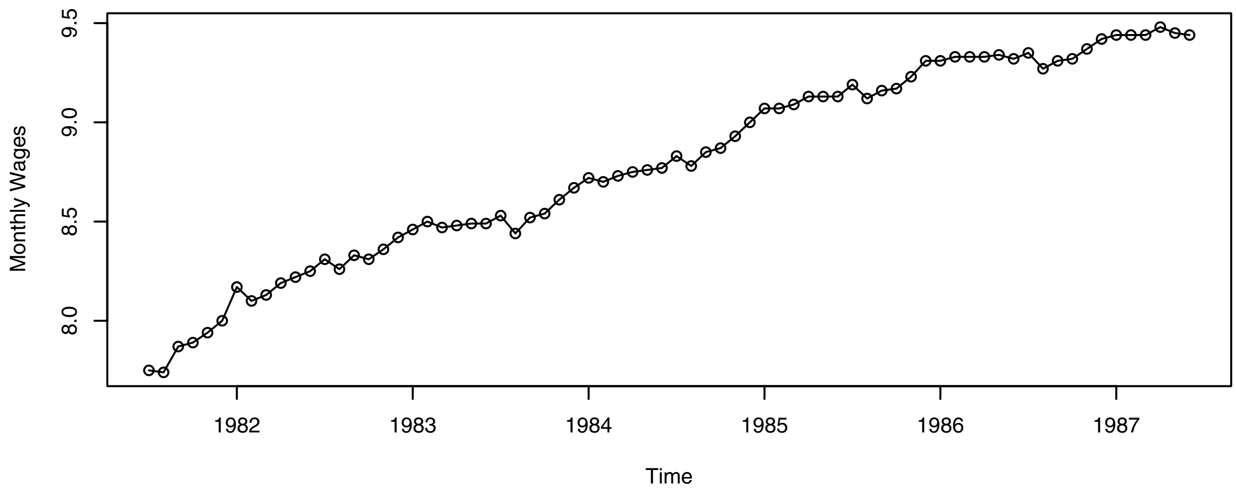



```
> plot(hours,ylab='Monthly Hours',type='l')
> points(y=hours,x=time(hours), pch=as.vector(season(hours)))
```

The most distinct pattern in this plot is that Decembers are nearly always high months relative the others. Decembers stick out.

Exercise 3.5 The data file wages contains monthly values of the average hourly wages ($) for workers in the U.S. apparel and textile products industry for July 1981 through June 1987.

(a) Display the time series plot for these data and interpret.




```
> data(wages); plot(wages,ylab='Monthly Wages',type='o')
```

This plot shows a strong increasing "trend," perhaps linear or curved.

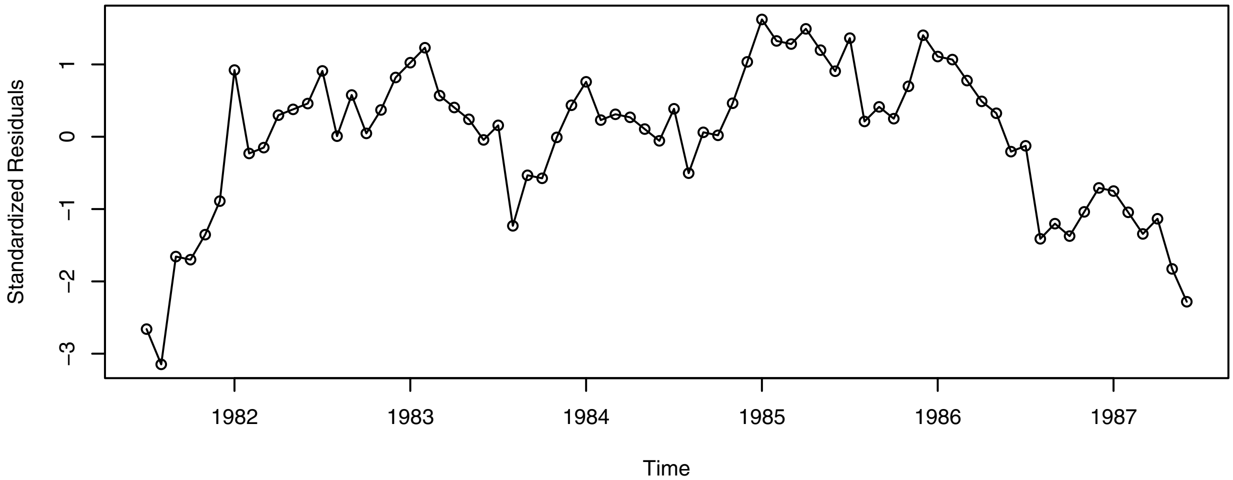
(b) Use least squares to fit a linear time trend to this time series. Interpret the regression output. Save the standardized residuals from the fit for further analysis.

```
> wages.lm=lm(wages~time(wages)); summary(wages.lm); y=rstudent(wages.lm)
```

```
Call:
lm(formula = wages ~ time(wages))

Residuals:
     Min       1Q   Median       3Q      Max
-0.23828 -0.04981  0.01942  0.05845  0.13136

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -5.490e+02  1.115e+01  -49.24   <2e-16 ***
time(wages)  2.811e-01  5.618e-03   50.03   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.08257 on 70 degrees of freedom
Multiple R-Squared: 0.9728,     Adjusted R-squared: 0.9724
F-statistic:  2503 on 1 and 70 DF,  p-value: < 2.2e-16
```

With a multiple R-squared of 97% and highly significant regression coefficients, it "appears" as if we might have an excellent model. However,...

(c) Construct and interpret the time series plot of the standardized residuals from part (b).




```
> plot(y,x=as.vector(time(wages)),ylab='Standardized Residuals',type='o')
```

This plot does not look “random” at all. It has, generally, an upside down U shape and suggests that perhaps we should try a quadratic fit.

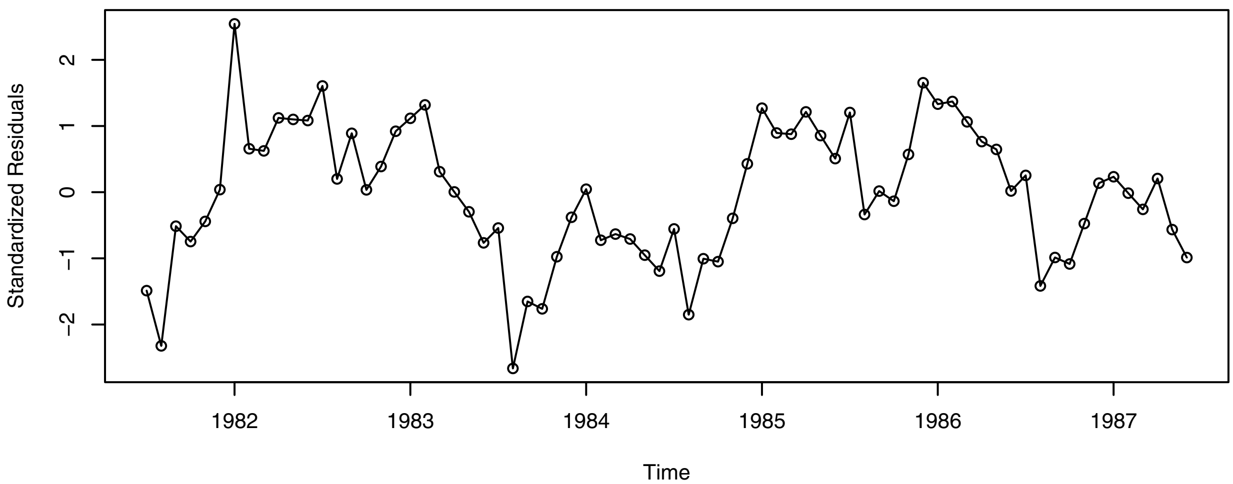
(d) Use least squares to fit a quadratic time trend to the wages time series. Interpret the regression output. Save the standardized residuals from the fit for further analysis.

```
> wages.lm2=lm(wages~time(wages)+I(time(wages)^2))
> summary(wages.lm2); y=rstudent(wages.lm)
```

```
Call:
lm(formula = wages ~ time(wages) + I(time(wages)^2))

Residuals:
      Min        1Q    Median        3Q       Max
-0.148318 -0.041440  0.001563  0.050089  0.139839

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)      -8.495e+04  1.019e+04  -8.336 4.87e-12 ***
time(wages)       8.534e+01  1.027e+01   8.309 5.44e-12 ***
I(time(wages)^2) -2.143e-02  2.588e-03  -8.282 6.10e-12 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.05889 on 69 degrees of freedom
Multiple R-Squared: 0.9864,     Adjusted R-squared: 0.986
F-statistic:  2494 on 2 and 69 DF,  p-value: < 2.2e-16
```

Again, based on the regression summary and a 99% R-squared, it “appears” as if we might have an excellent model.

(e) Construct and interpret the time series plot of the standardized residuals from part (d).

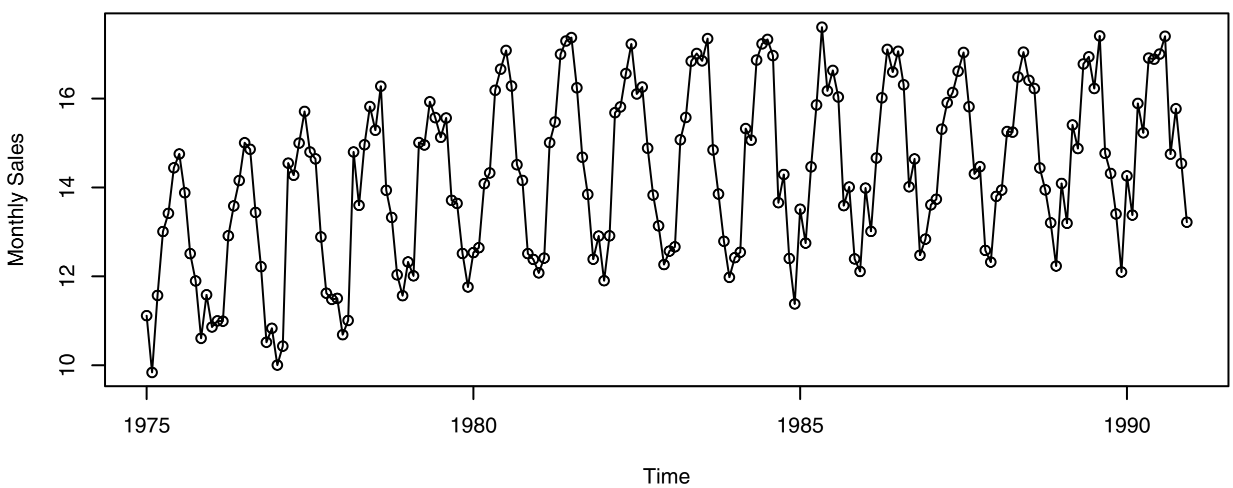



```
> plot(y,x=as.vector(time(wages)),ylab='Standardized Residuals',type='o')
```

This plot does not look “random” either. It hangs together too much—it is too smooth. See **Exercise** 3.11.

Exercise 3.6 The data file beersales contains monthly U.S. beer sales (in millions of barrels) for the period January 1975 through December 1990.

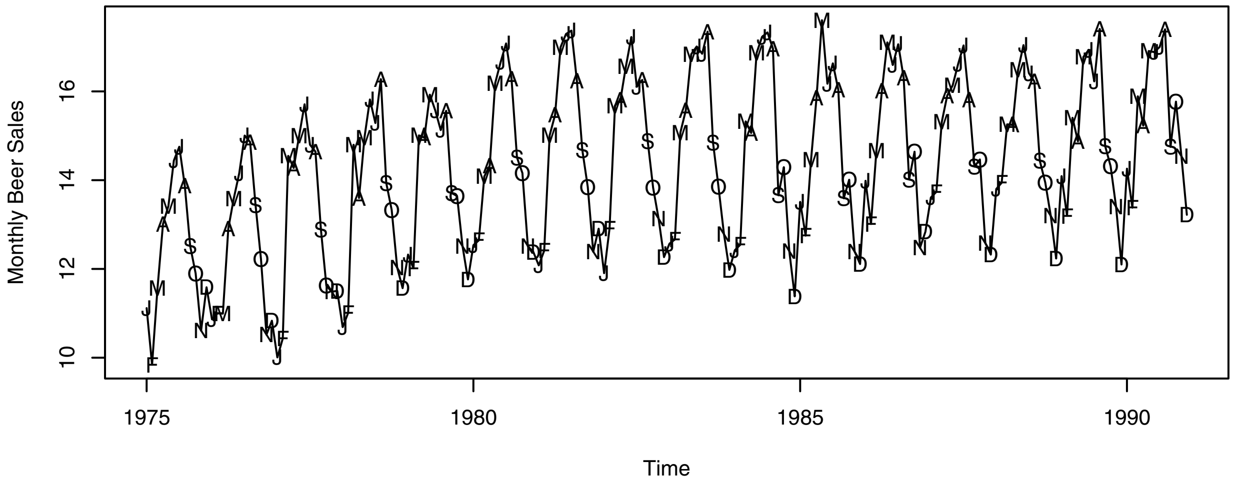
(a) Display the time series plot for these data and interpret the plot.




```
> data(beersales); plot(beersales,ylab='Monthly Sales',type='o')
```

In addition to a possible seasonality in the series, there is a general upward “trend” in the first part of the series. However, this effect “levels off” in the latter years.

(b) Now construct a time series plot that uses separate plotting symbols for the various months. Does your interpretation change from that in part (a)?




```
> plot(beersales,ylab='Monthly Beer Sales',type='l')
> points(y=beersales,x=time(beersales), pch=as.vector(season(beersales)))
```

Now the seasonality is quite clear with higher sales in the summer months and lower sales in the winter.

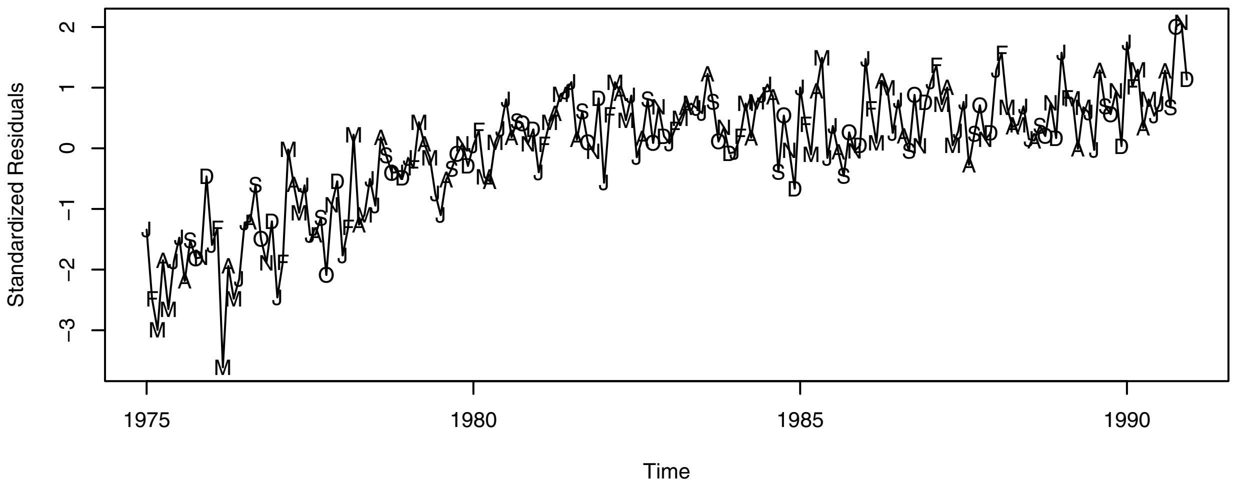
(c) Use least squares to fit a seasonal-means trend to this time series. Interpret the regression output. Save the standardized residuals from the fit for further analysis.

```
> month.=season(beersales);beersales.lm=lm(beersales~month.);summary(beersales.lm)
```

```
Call:
lm(formula = beersales ~ month.)

Residuals:
    Min      1Q  Median      3Q     Max
-3.5745 -0.4772  0.1759  0.7312  2.1023

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)     12.48568    0.26392  47.309  < 2e-16 ***
month.February  -0.14259    0.37324  -0.382 0.702879
month.March      2.08219    0.37324   5.579 8.77e-08 ***
month.April      2.39760    0.37324   6.424 1.15e-09 ***
month.May        3.59896    0.37324   9.643  < 2e-16 ***
month.June       3.84976    0.37324  10.314  < 2e-16 ***
month.July       3.76866    0.37324  10.097  < 2e-16 ***
month.August     3.60877    0.37324   9.669  < 2e-16 ***
month.September  1.57282    0.37324   4.214 3.96e-05 ***
month.October    1.25444    0.37324   3.361 0.000948 ***
month.November  -0.04797    0.37324  -0.129 0.897881
month.December  -0.42309    0.37324  -1.134 0.258487
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.056 on 180 degrees of freedom
Multiple R-Squared: 0.7103,     Adjusted R-squared: 0.6926
F-statistic: 40.12 on 11 and 180 DF,  p-value: < 2.2e-16
```

This model leaves out the January term so all of the other effects are in comparison to January. The multiple R-squared is rather large at 71% and all the terms except November, December, and February are significantly different from January.

(d) Construct and interpret the time series plot of the standardized residuals from part (c). Be sure to use proper plotting symbols to check on seasonality in the standardized residuals.




```
> plot(y=rstudent(beersales.lm),x=as.vector(time(beersales)),type='l',
   ylab='Standardized Residuals')
> points(y=rstudent(beersales.lm),x=as.vector(time(beersales)),
   pch=as.vector(season(beersales)))
```

Display this plot full screen to see the detail. However, it is clear that this model does not capture the structure of this time series and we proceed to look for a more adequate model.

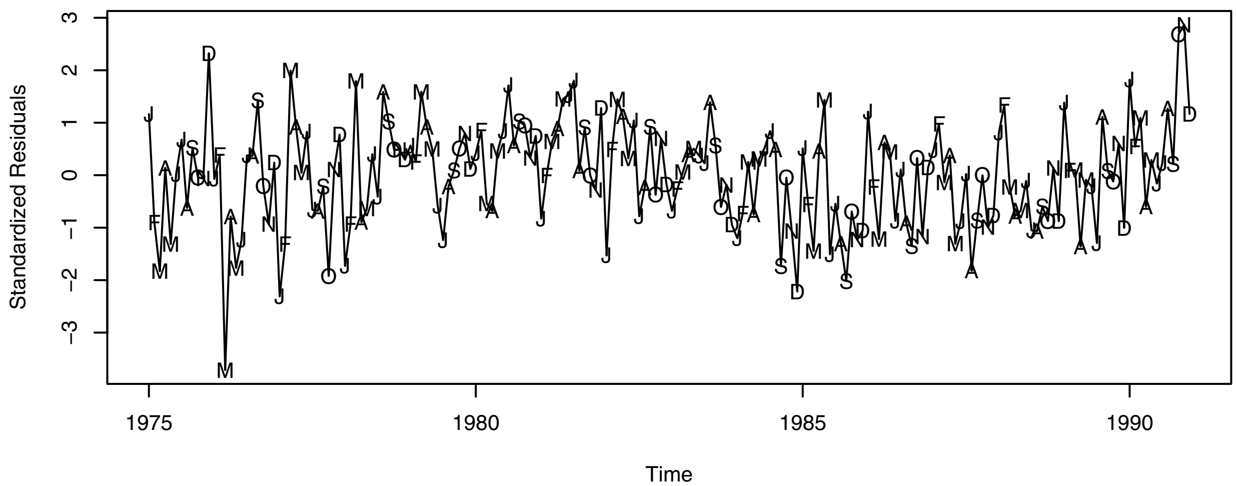
(e) Use least squares to fit seasonal-means plus quadratic time trend to the beer sales time series. Interpret the regression output. Save the standardized residuals from the fit for further analysis.

```
> beersales.lm2=lm(beersales~month.+time(beersales)+I(time(beersales)^2))
> summary(beersales.lm2)
```

```
Call:
lm(formula = beersales ~ month. + time(beersales) + I(time(beersales)^2))

Residuals:
     Min       1Q   Median       3Q      Max
-2.03203 -0.43118  0.04977  0.34509  1.57572

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)
(Intercept)         -7.150e+04  8.791e+03  -8.133 6.93e-14 ***
month.February      -1.579e-01  2.090e-01  -0.755  0.45099
month.March          2.052e+00  2.090e-01   9.818  < 2e-16 ***
month.April          2.353e+00  2.090e-01  11.256  < 2e-16 ***
month.May            3.539e+00  2.090e-01  16.934  < 2e-16 ***
month.June           3.776e+00  2.090e-01  18.065  < 2e-16 ***
month.July           3.681e+00  2.090e-01  17.608  < 2e-16 ***
month.August         3.507e+00  2.091e-01  16.776  < 2e-16 ***
month.September      1.458e+00  2.091e-01   6.972 5.89e-11 ***
month.October        1.126e+00  2.091e-01   5.385 2.27e-07 ***
month.November      -1.894e-01  2.091e-01  -0.906  0.36622
month.December      -5.773e-01  2.092e-01  -2.760  0.00638 **
time(beersales)      7.196e+01  8.867e+00   8.115 7.70e-14 ***
I(time(beersales)^2) -1.810e-02  2.236e-03  -8.096 8.63e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5911 on 178 degrees of freedom
Multiple R-Squared: 0.9102,     Adjusted R-squared: 0.9036
F-statistic: 138.8 on 13 and 178 DF,  p-value: < 2.2e-16
```

This model seems to do a better job than the seasonal means alone but we should reserve judgement until we look next at the residuals.

(f) Construct and interpret the time series plot of the standardized residuals from part (e). Again use proper plotting symbols to check for any remaining seasonality in the residuals.

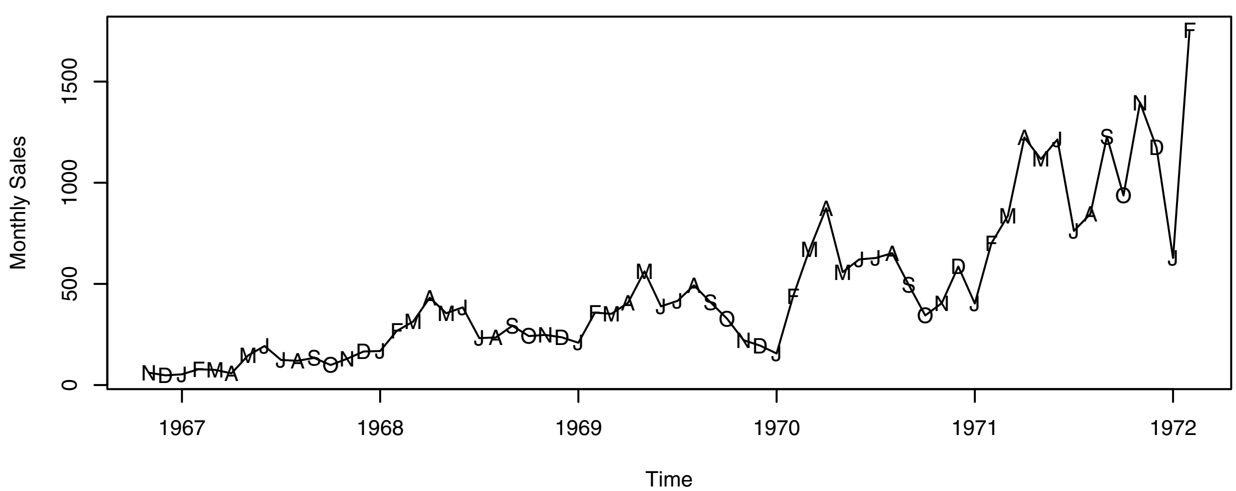



```
> plot(y=rstudent(beersales.lm2),x=as.vector(time(beersales)),type='l',
>   ylab='Standardized Residuals')
> points(y=rstudent(beersales.lm2),x=as.vector(time(beersales)),
   pch=as.vector(season(beersales)))
```

This model does much better than the previous one but we would be hard pressed to convince anyone that the underlying quadratic "trend" makes sense. Notice that the coefficient on the square term is negative so that in the future sales will decrease substantially and even eventually go negative!

Exercise 3.7 The data file winnebago contains monthly unit sales of recreational vehicles from Winnebago, Inc. from November 1966 through February 1972.

(a) Display and interpret the time series plot for these data.



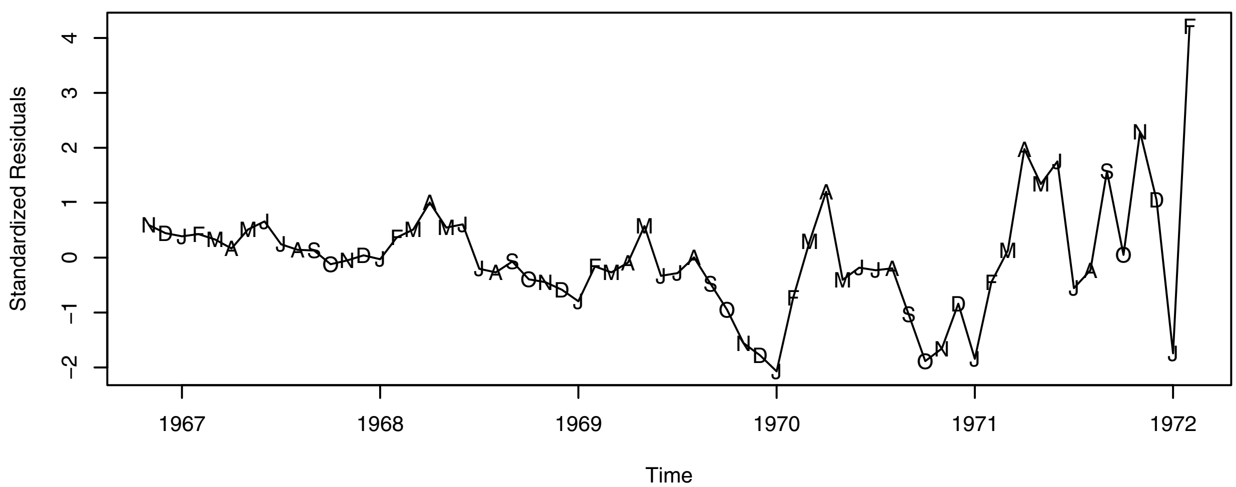

```
> data(winnebago); plot(winnebago,ylab='Monthly Sales',type='l')
> points(y=winnebago,x=time(winnebago), pch=as.vector(season(winnebago)))
```

As we would expect with recreational vehicles in the U.S., there is substantial seasonality in the series. However, there is also a general upward "trend" with increasing variation at the higher levels of the series.

(b) Use least squares to fit a line to these data. Interpret the regression output. Plot the standardized residuals from the fit as a time series. Interpret the plot.

```
> winnebago.lm=lm(winnebago~time(winnebago)); summary(winnebago.lm)
```

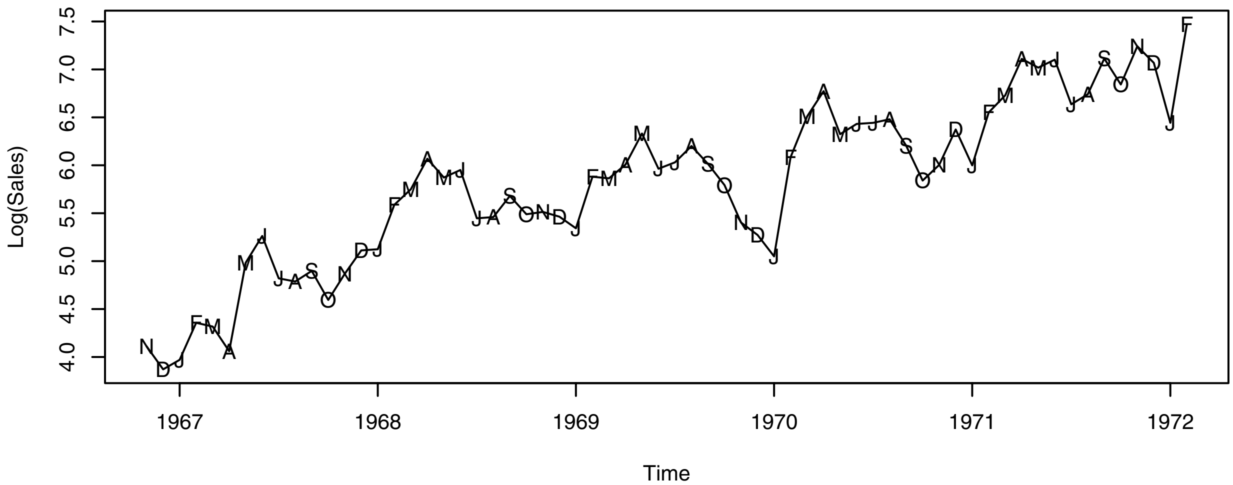
```
Call:
lm(formula = winnebago ~ time(winnebago))

Residuals:
    Min      1Q  Median      3Q     Max
-419.58  -93.13  -12.78   94.96  759.21

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)     -394885.68   33539.77  -11.77   <2e-16 ***
time(winnebago)     200.74      17.03   11.79   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 209.7 on 62 degrees of freedom
Multiple R-Squared: 0.6915,     Adjusted R-squared: 0.6865
F-statistic: 138.9 on 1 and 62 DF,  p-value: < 2.2e-16
```




```
> plot(y=rstudent(winnebago.lm),x=as.vector(time(winnebago)),type='l',
  ylab='Standardized Residuals')
> points(y=rstudent(winnebago.lm),x=as.vector(time(winnebago)),
  pch=as.vector(season(winnebago)))
```

Although the “trend” has been removed, this clearly is not an acceptable model and we move on.

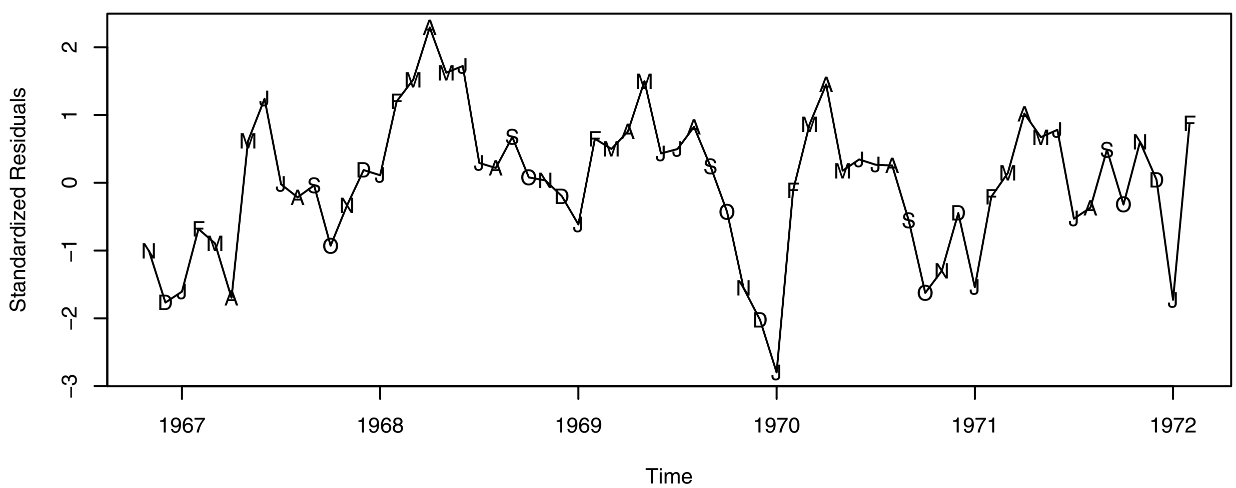
(c) Now take natural logarithms of the monthly sales figures and display and interpret the time series plot of the transformed values.




```
> plot(log(winnebago),ylab='Log(Sales)',type='l')
> points(y=log(winnebago),x=time(winnebago), pch=as.vector(season(winnebago)))
```

In this we see that the seasonality is still present but that now the upward trend is accompanied by much more equal variation around that trend.

(d) Use least squares to fit a line to the logged data. Display the time series plot of the standardized residuals from this fit and interpret.




```
> logwinnebago.lm=lm(log(winnebago)~time(log(winnebago))); summary(logwinnebago.lm)
> plot(y=rstudent(logwinnebago.lm),x=as.vector(time(winnebago)),type='l',
   ylab='Standardized Residuals')
> points(y=rstudent(logwinnebago.lm),x=as.vector(time(winnebago)),
   pch=as.vector(season(winnebago)))
```

The residual plot looks much more acceptable now but we still need to model the seasonality.

(e) Now use least squares to fit a seasonal-means plus linear time trend to the logged sales time series and save the standardized residuals for further analysis. Check the statistical significance of each of the regression coefficients in the model.

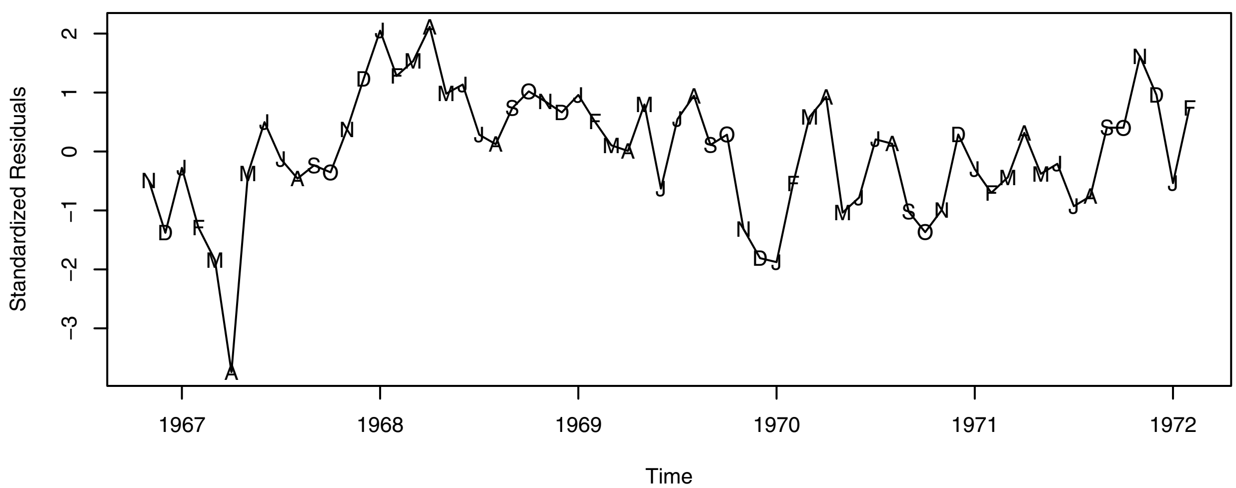
```
> month.=season(winnebago)
> logwinnebago.lm2=lm(log(winnebago)~month.+time(log(winnebago)))
```

```
> summary(logwinnebago.lm2)
```

```
Call:
lm(formula = log(winnebago) ~ month. + time(log(winnebago)))

Residuals:
     Min       1Q   Median       3Q      Max
-0.92501 -0.16328  0.03344  0.20757  0.57388

Coefficients:
                       Estimate Std. Error t value Pr(>|t|)
(Intercept)          -997.33061   50.63995 -19.695  < 2e-16 ***
month.February          0.62445    0.18182   3.434  0.00119 **
month.March             0.68220    0.19088   3.574  0.00078 ***
month.April             0.80959    0.19079   4.243 9.30e-05 ***
month.May               0.86953    0.19073   4.559 3.25e-05 ***
month.June              0.86309    0.19070   4.526 3.63e-05 ***
month.July              0.55392    0.19069   2.905  0.00542 **
month.August            0.56989    0.19070   2.988  0.00431 **
month.September         0.57572    0.19073   3.018  0.00396 **
month.October           0.26349    0.19079   1.381  0.17330
month.November          0.28682    0.18186   1.577  0.12095
month.December          0.24802    0.18182   1.364  0.17853
time(log(winnebago))    0.50909    0.02571  19.800  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3149 on 51 degrees of freedom
Multiple R-Squared: 0.8946,     Adjusted R-squared: 0.8699
F-statistic: 36.09 on 12 and 51 DF,  p-value: < 2.2e-16
```

This model explains a large percentage of the variation in sales but, as always, we should also look at the residuals.

(f) Display the time series plot of the standardized residuals obtained in part (e). Interpret the plot.

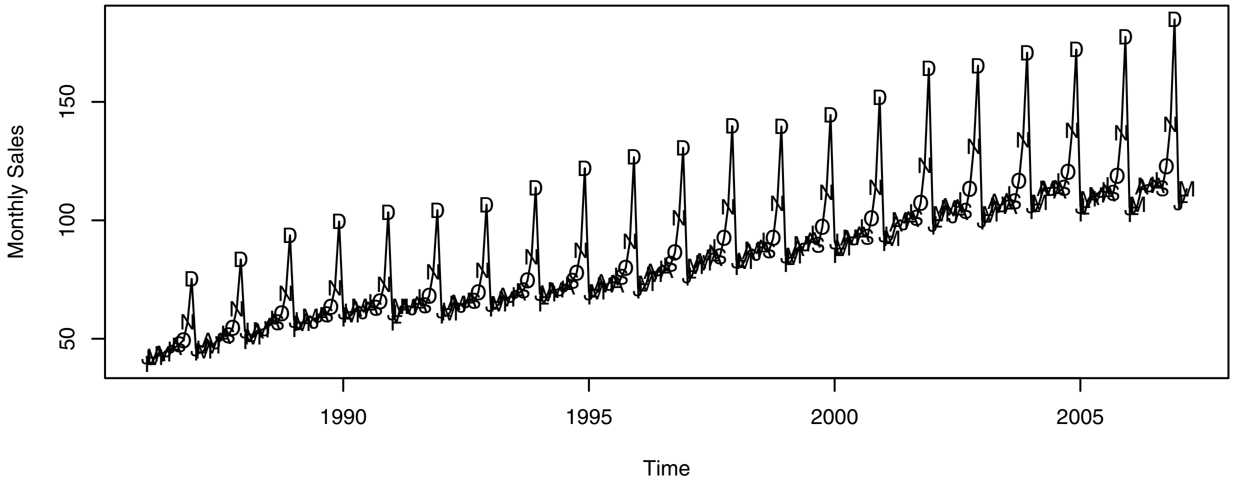



```
> plot(y=rstudent(logwinnebago.lm2),x=as.vector(time(winnebago)),type='l',
>  ylab='Standardized Residuals')
> points(y=rstudent(logwinnebago.lm2),x=as.vector(time(winnebago)),
>  pch=as.vector(season(winnebago)))
```

This residual plot is the best we have seen for models of this series but perhaps there are better models to be explored later.

Exercise 3.8 The data file retail lists total UK (United Kingdom) retail sales (in billions of pounds) from January 1986 through March 2007. The data are not "seasonally adjusted" and year 2000 = 100 is the base year.

(a) Display and interpret the time series plot for these data. Be sure to use plotting symbols that permit you to look for seasonality.




```
> data(retail); plot(retail,ylab='Monthly Sales',type='l')
> points(y=retail,x=time(retail), pch=as.vector(season(retail)))
```

There is a clear upward trend with at least some seasonality. Large holiday sales in November and especially December are striking. There is some tendency for increased variation at the higher levels of the series.

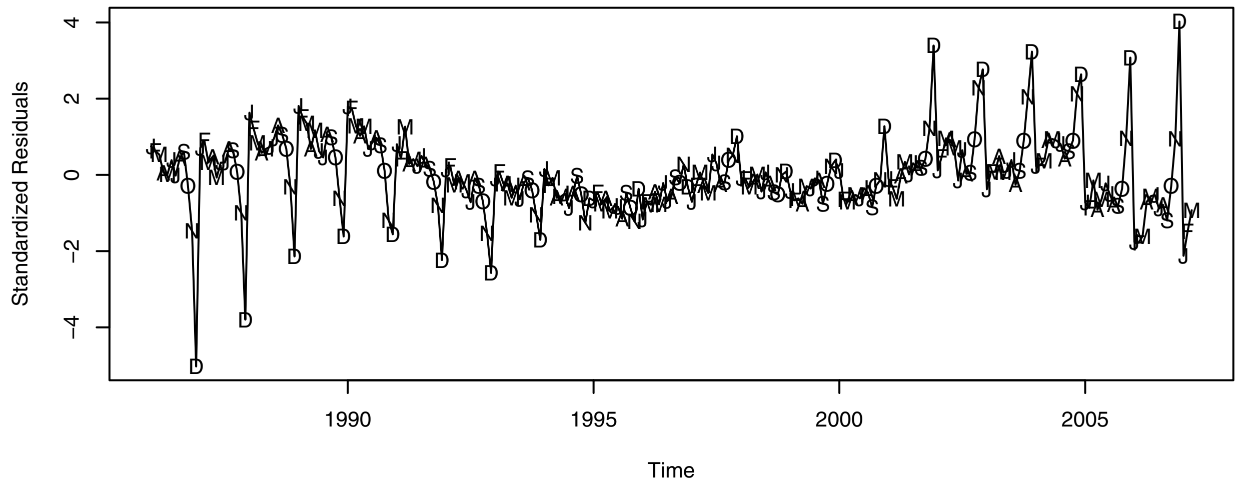
(b) Use least squares to fit a seasonal-means plus linear time trend to this time series. Interpret the regression output and save the standardized residuals from the fit for further analysis.

```
> month.=season(retail)
> retail.lm=lm(retail~month.+time(retail))
> summary(retail.lm)
```

```
Call:
lm(formula = retail ~ month. + time(retail))

Residuals:
     Min       1Q   Median       3Q      Max
-19.8950  -2.4440  -0.3518   2.1971  16.2045

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)     -7.249e+03  8.724e+01 -83.099  < 2e-16 ***
month.February  -3.015e+00  1.290e+00  -2.337  0.02024 *
month.March      7.469e-02  1.290e+00   0.058  0.95387
month.April      3.447e+00  1.305e+00   2.641  0.00880 **
month.May        3.108e+00  1.305e+00   2.381  0.01803 *
month.June       3.074e+00  1.305e+00   2.355  0.01932 *
month.July       6.053e+00  1.305e+00   4.638 5.76e-06 ***
month.August     3.138e+00  1.305e+00   2.404  0.01695 *
month.September  3.428e+00  1.305e+00   2.626  0.00919 **
month.October    8.555e+00  1.305e+00   6.555 3.34e-10 ***
month.November   2.082e+01  1.305e+00  15.948  < 2e-16 ***
month.December   5.254e+01  1.305e+00  40.255  < 2e-16 ***
time(retail)     3.670e+00  4.369e-02  83.995  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.278 on 242 degrees of freedom
Multiple R-Squared: 0.9767,     Adjusted R-squared: 0.9755
F-statistic:   845 on 12 and 242 DF,  p-value: < 2.2e-16
```

All but the March effect is statistically significant at the usual levels and the R-square is very large. Let's consider the residuals.

(c) Construct and interpret the time series plot of the standardized residuals from part (b). Be sure to use proper plotting symbols to check on seasonality.

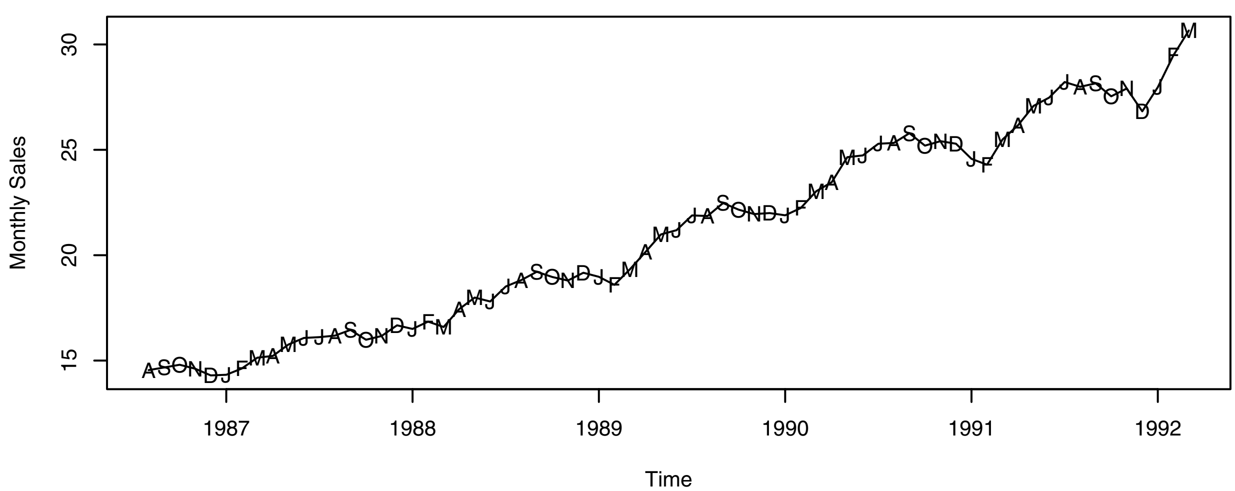



```
> plot(y=rstudent(retail.lm),x=as.vector(time(retail)),type='l',
   ylab='Standardized Residuals', xlab='Time')
> points(y=rstudent(retail.lm),x=as.vector(time(retail)),
   pch=as.vector(season(retail)))
```

Clearly this model still leaves a lot to be desired.

Exercise 3.9 The data file prescrip gives monthly U.S. prescription costs for the months August 1986 to March 1992. These data are from the State of New Jersey's Prescription Drug Program and are the cost per prescription claim.

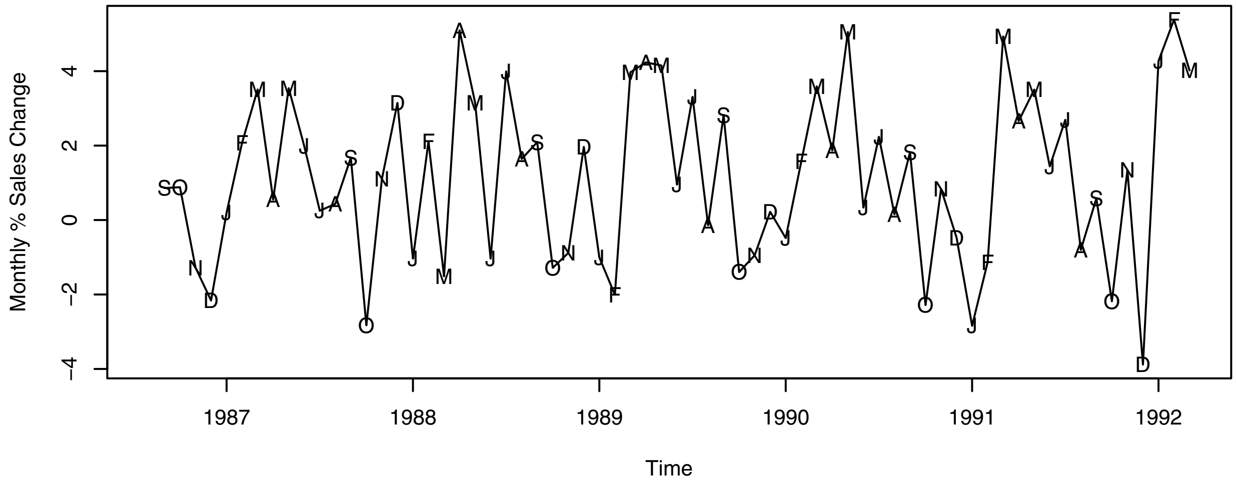
(a) Display and interpret the time series plot for these data. Use plotting symbols that permit you to look for seasonality.




```
> data(prescrip); plot(prescrip,ylab='Monthly Sales',type='l')
> points(y=prescrip,x=time(prescrip), pch=as.vector(season(prescrip)))
```

This plot shows a generally linear upward trend with possible seasonality as Septembers are generally peaks in all years.

(b) Calculate and plot the sequence of month-to-month percentage changes in the prescription costs. Again, use plotting symbols that permit you to look for seasonality.




```
> perprescrip=na.omit(100*(prescrip-zlag(prescrip))/zlag(prescrip))
> plot(perprescrip,ylab='Monthly % Sales Change',type='l')
> points(y=perprescrip,x=time(perprescrip), pch=as.vector(season(perprescrip)))
```

The percentage changes look reasonably stable and stationary with perhaps some subtle seasonality.

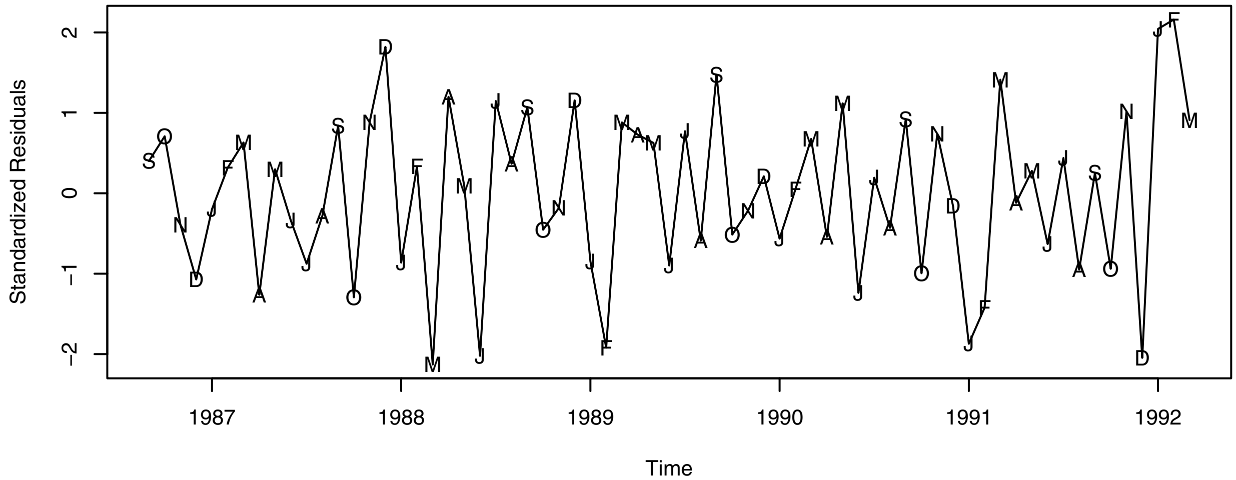
(c) Use least squares to fit a cosine trend with fundamental frequency 1/12 to the percentage change series. Interpret the regression output. Save the standardized residuals.

```
> har.=harmonic(perprescrip); prescrip.lm=lm(perprescrip~har.); summary(prescrip.lm)
```

```
Call:
lm(formula = perprescrip ~ har.)

Residuals:
    Min      1Q  Median      3Q     Max
-3.8444 -1.3742  0.1697  1.4069  3.8980

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)       1.2217     0.2325   5.254 1.82e-06 ***
har.cos(2*pi*t)  -0.6538     0.3298  -1.982   0.0518 .
har.sin(2*pi*t)   1.6596     0.3269   5.077 3.54e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.897 on 64 degrees of freedom
Multiple R-Squared: 0.3148,     Adjusted R-squared: 0.2933
F-statistic:  14.7 on 2 and 64 DF,  p-value: 5.584e-06
```

The cosine trend model is statistically significant but the R-squared is only 31%.

(d) Plot the sequence of standardized residuals to investigate the adequacy of the cosine trend model. Interpret the plot.




```
> plot(y=rstudent(prescrip.lm),x=as.vector(time(perprescrip)),type='l',
>   ylab='Standardized Residuals')
> points(y=rstudent(prescrip.lm),x=as.vector(time(perprescrip)),
>   pch=as.vector(season(perprescrip)))
```

These residuals look basically random and without any significant patterns.

Exercise 3.10 (Continuation of Exercise 3.4) Consider the hours time series again.

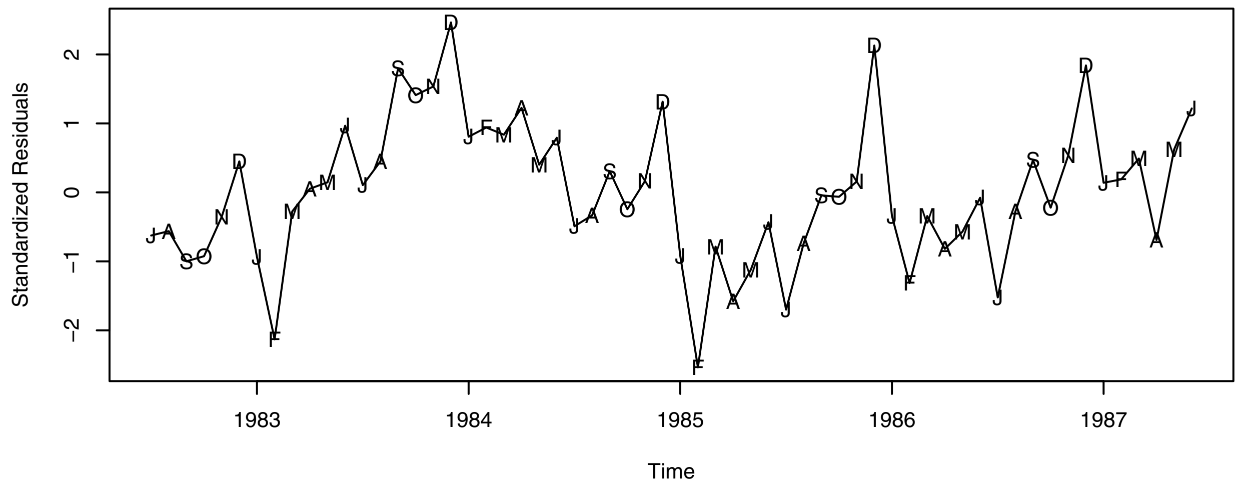
(a) Use least squares to fit a quadratic trend to these data. Interpret the regression output and save the standardized residuals for further analysis.

```
> data(hours); hours.lm=lm(hours~time(hours)+I(time(hours)^2)); summary(hours.lm)
```

```
Call:
lm(formula = hours ~ time(hours) + I(time(hours)^2))

Residuals:
     Min       1Q   Median       3Q      Max
-1.00603 -0.25431 -0.02267  0.22884  0.98358

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)      -5.122e+05  1.155e+05  -4.433 4.28e-05 ***
time(hours)       5.159e+02  1.164e+02   4.431 4.31e-05 ***
I(time(hours)^2) -1.299e-01  2.933e-02  -4.428 4.35e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.423 on 57 degrees of freedom
Multiple R-Squared: 0.5921,     Adjusted R-squared: 0.5778
F-statistic: 41.37 on 2 and 57 DF,  p-value: 7.97e-12
```

The quadratic fit looks to be quite significant, but, of course, the seasonality, if any, has not been accounted for.

(b) Display a sequence plot of the standardized residuals and interpret. Use monthly plotting symbols so that possible seasonality may be readily identified.




```
> plot(y=rstudent(hours.lm),x=as.vector(time(hours)),type='l',
   ylab='Standardized Residuals')
> points(y=rstudent(hours.lm),x=as.vector(time(hours)), pch=as.vector(season(hours)))
```

These residuals are too smooth for randomness and Decembers are all high.

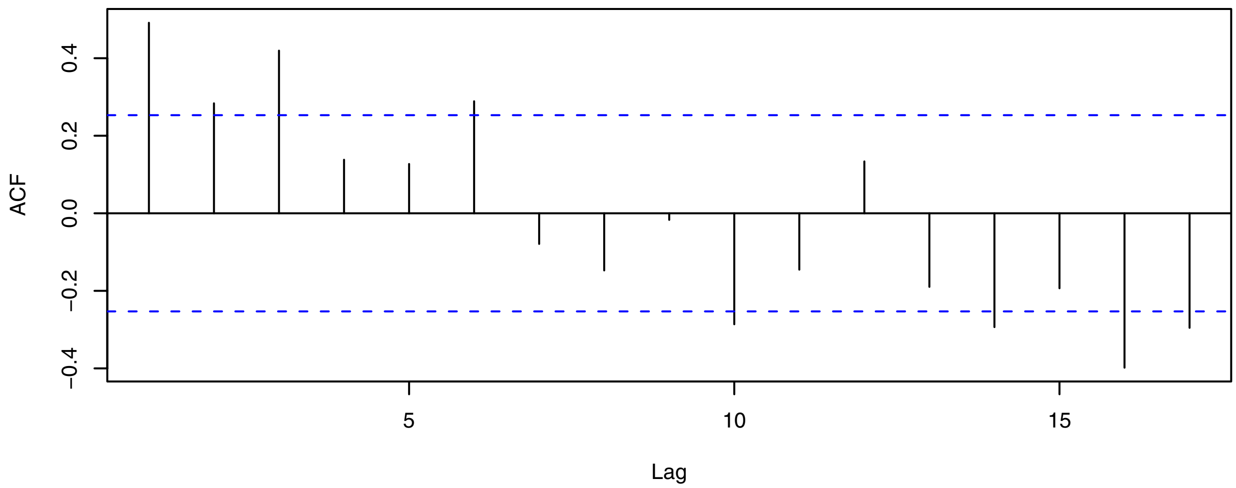
(c) Perform the Runs test of the standardized residuals and interpret the results.

```
> runs(rstudent(hours.lm))
```

```
$pvalue
[1] 0.00012
$observed.runs
[1] 16
$expected.runs
[1] 30.96667
$n1
[1] 31
$n2
[1] 29
$k
[1] 0
```

The *p*-value of 0.00012 indicates that our suspicion of nonrandomness is quite justified.

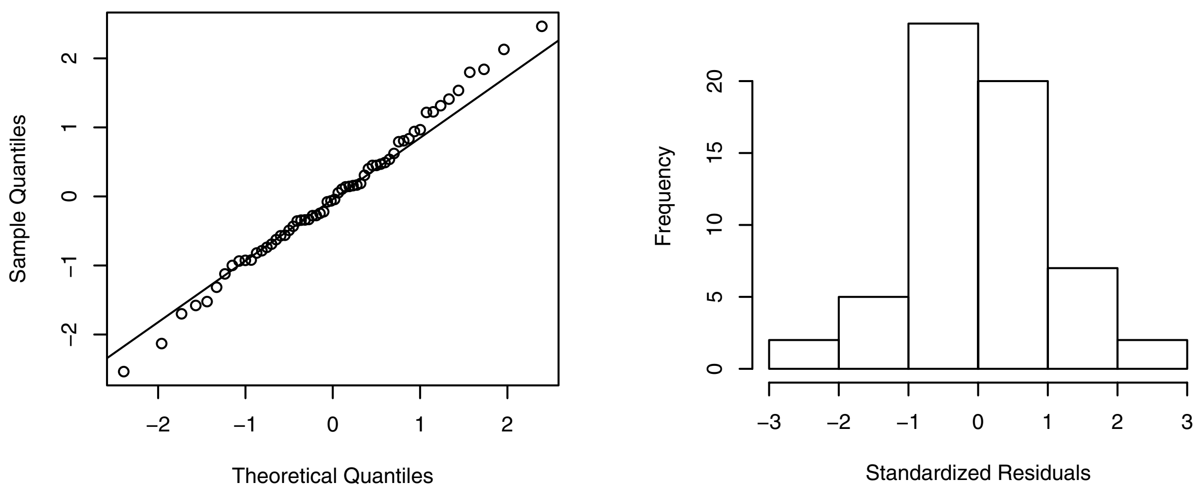
(d) Calculate the sample autocorrelations for the standardized residuals and interpret.




```
> acf(rstudent(hours.lm))
```

There is significant autocorrelation at lags 1, 3, 6, 10, 14, 16, and 17.

(e) Investigate the normality of the standardized residuals (error terms). Consider histograms and normal probability plots. Interpret the plots.




```
> qqnorm(rstudent(hours.lm)); qqline(rstudent(hours.lm))
> hist(rstudent(hours.lm),xlab='Standardized Residuals')
> shapiro.test(rstudent(hours.lm))
```

```
Shapiro-Wilk normality test
data:  rstudent(hours.lm)
W = 0.9939, p-value = 0.991
```

We have no evidence against normality for the error terms in this model.

Exercise 3.11 (Continuation of Exercise 3.5) Return to the wages series.

(a) Consider the residuals from a least squares fit of a quadratic time trend.

```
> wages.lm2=lm(wages~time(wages)+I(time(wages)^2)); summary(wages.lm2)
```

```
Call:
lm(formula = wages ~ time(wages) + I(time(wages)^2))
```

```
Residuals:
      Min        1Q    Median        3Q       Max
-0.148318 -0.041440  0.001563  0.050089  0.139839

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)      -8.495e+04  1.019e+04  -8.336 4.87e-12 ***
time(wages)       8.534e+01  1.027e+01   8.309 5.44e-12 ***
I(time(wages)^2) -2.143e-02  2.588e-03  -8.282 6.10e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.05889 on 69 degrees of freedom
Multiple R-Squared: 0.9864,     Adjusted R-squared: 0.986
F-statistic:  2494 on 2 and 69 DF,  p-value: < 2.2e-16
```

The quardratic fit is certainly statistically significant!

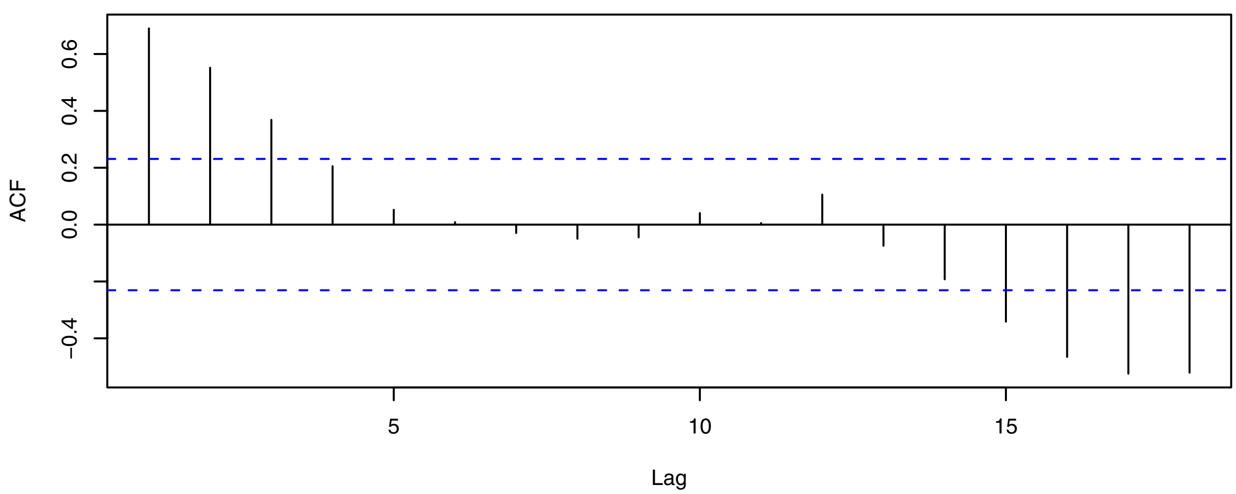
(b) Perform a Runs test on the standardized residuals and interpret the results.

```
> runs(rstudent(wages.lm2))
```

```
$pvalue
[1] 1.56e-07
$observed.runs
[1] 15
$expected.runs
[1] 36.75
$n1
[1] 33
$n2
[1] 39
$k
[1] 0
```

There are too few runs for these residuals (and hence error terms) to be considered random.

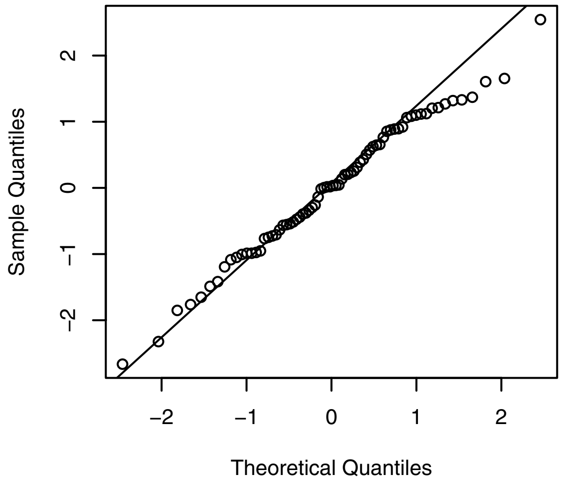
(c) Calculate the sample autocorrelations for the standardized residuals and interpret.

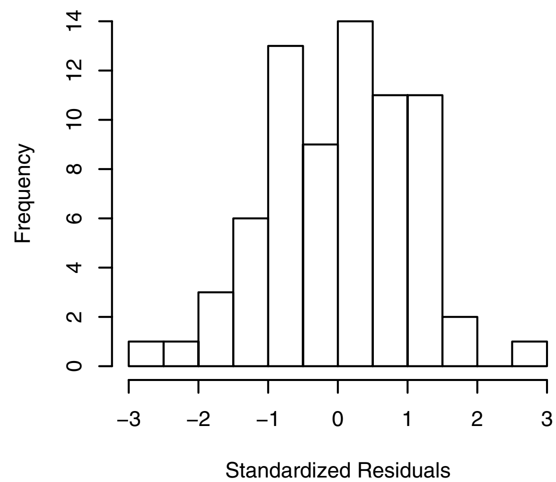



```
> acf(rstudent(wages.lm2))
```

The autocorrelations displayed reinforce the results from the runs test.

(d) Investigate the normality of the standardized residuals (error terms). Consider histograms and normal probability plots. Interpret the plots.






```
> qqnorm(rstudent(wages.lm2)); qqline(rstudent(wages.lm2))
> hist(rstudent(wages.lm2),xlab='Standardized Residuals')
> shapiro.test(rstudent(wages.lm2)
```

Although the distribution is rather mound shaped, normality is somewhat tenuous. However, the evidence is not sufficient to reject normality. See Shapiro-Wilk test results below.

```
Shapiro-Wilk normality test
data:  rstudent(wages.lm2)
W = 0.9887, p-value = 0.7693
```

Exercise 3.12 (Continuation of Exercise 3.6) Consider the time series in the data file beersales.

(a) Obtain the residuals from the least squares fit of the seasonal-means plus quadratic time trend model.

```
> beersales.lm2=lm(beersales~month.+time(beersales)+I(time(beersales)^2))
```

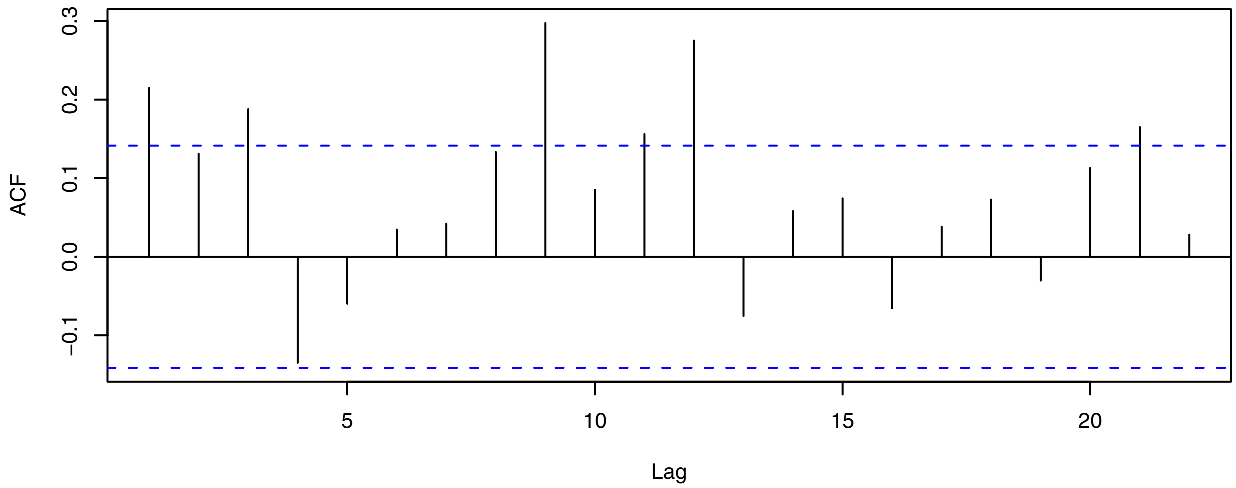
(b) Perform a Runs test on the standardized residuals and interpret the results.

```
> runs(rstudent(beersales.lm2))
```

```
$pvalue
[1] 0.0127
$observed.runs
[1] 79
$expected.runs
[1] 96.625
$n1
[1] 90
$n2
[1] 102
$k
[1] 0
```

We would reject independence of the error terms on the basis of these results.

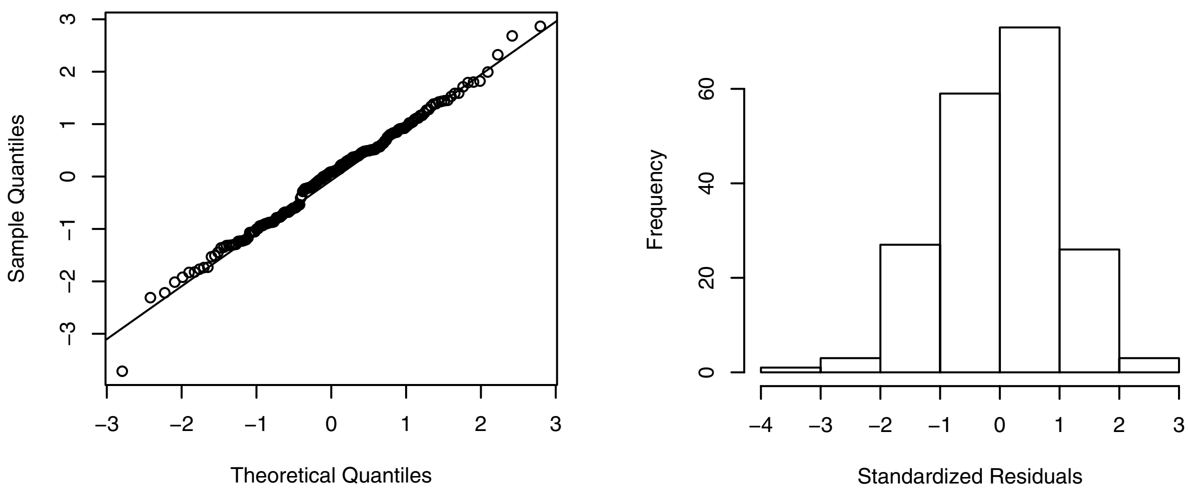
(c) Calculate the sample autocorrelations for the standardized residuals and interpret.




```
> acf(rstudent(beersales.lm2))
```

These results also show the lack of independence in the error terms of this model.

(d) Investigate the normality of the standardized residuals (error terms). Consider histograms and normal probability plots. Interpret the plots.




```
> qqnorm(rstudent(beersales.lm2)); qqline(rstudent(beersales.lm2))
> hist(rstudent(beersales.lm2),xlab='Standardized Residuals')
> shapiro.test(rstudent(beersales.lm2))
```

```
Shapiro-Wilk normality test
data:  rstudent(beersales.lm2)
W = 0.9924, p-value = 0.4139
```

All of these results provide good support for the assumption of normal error terms.

Exercise 3.13 (Continuation of Exercise 3.7) Return to the winnebago time series.

(a) Calculate the least squares residuals from a seasonal-means plus linear time trend model on the logarithms of the sales time series.

```
> month.=season(winnebago)
> logwinnebago.lm2=lm(log(winnebago)~month.+time(winnebago))
```

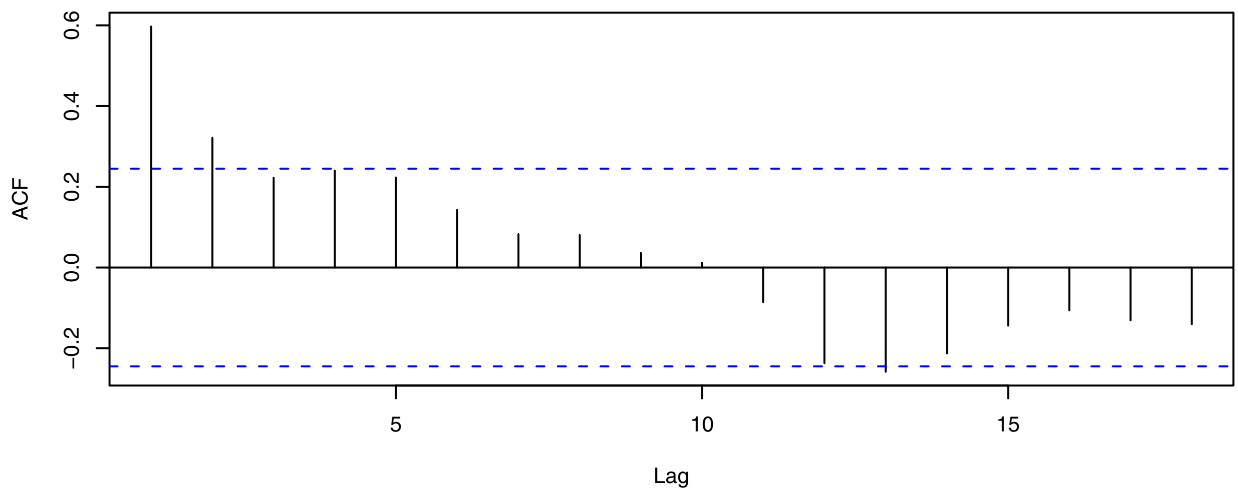
(b) Perform a Runs test on the standardized residuals and interpret the results.

```
> runs(rstudent(logwinnebago.lm2))
```

```
$pvalue
[1] 0.000243
$observed.runs
[1] 18
$expected.runs
[1] 32.71875
$n1
[1] 29
$n2
[1] 35
$k
[1] 0
```

These results suggest strongly that the error terms are not independent.

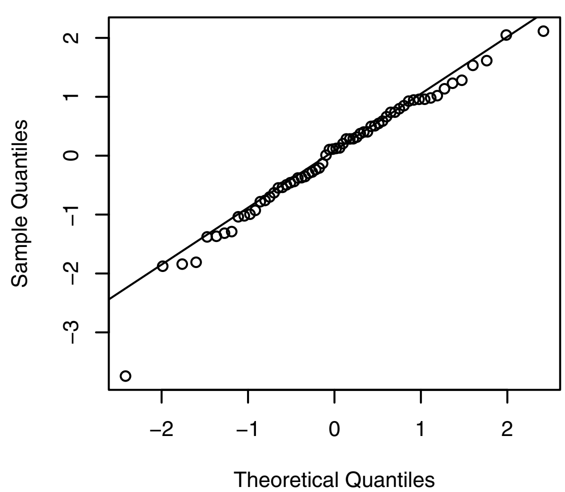
(c) Calculate the sample autocorrelations for the standardized residuals and interpret.

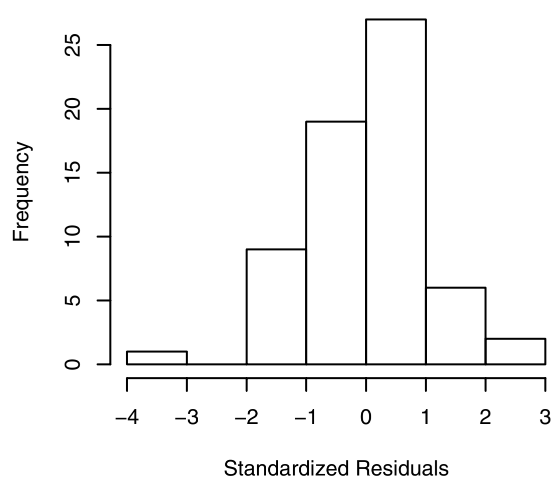



```
> acf(rstudent(logwinnebago.lm2))
```

More evidence that the error terms are not independent.

(d) Investigate the normality of the standardized residuals (error terms). Consider histograms and normal probability plots. Interpret the plots.






```
> qqnorm(rstudent(logwinnebago.lm2)); qqline(rstudent(logwinnebago.lm2))
```

```
> hist(rstudent(logwinnebago.lm2),xlab='Standardized Residuals')
> shapiro.test(rstudent(logwinnebago.lm2))
```

```
       Shapiro-Wilk normality test
       data:  rstudent(logwinnebago.lm2)
       W = 0.9704, p-value = 0.1262
```

There is evidence of an outlier on the low end but not enough to reject normality at this point.

Exercise 3.14 (Continuation of Exercise 3.8) The data file retail contains UK monthly retail sales figures.

(a) Obtain the least squares residuals from a seasonal-means plus linear time trend model.

```
> month.=season(retail); retail.lm=lm(retail~month.+time(retail))
```

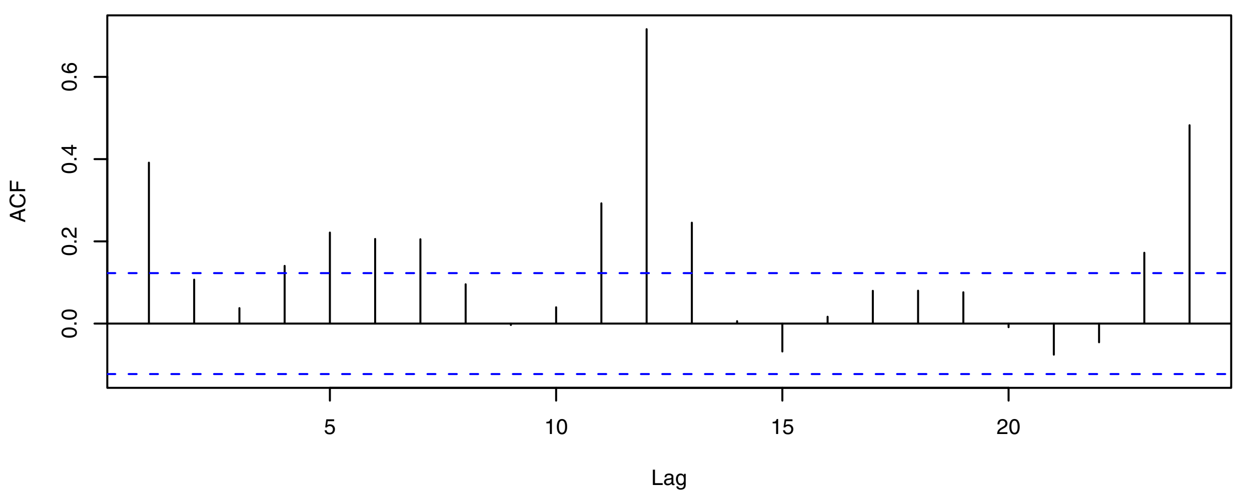
(b) Perform a Runs test on the standardized residuals and interpret the results.

```
> runs(rstudent(retail.lm))
```

```
       $pvalue
       [1] 9.19e-23
       $observed.runs
       [1] 52
       $expected.runs
       [1] 127.9333
       $n1
       [1] 136
       $n2
       [1] 119
       $k
       [1] 0
```

The runs test provides strong evidence against randomness in the error terms.

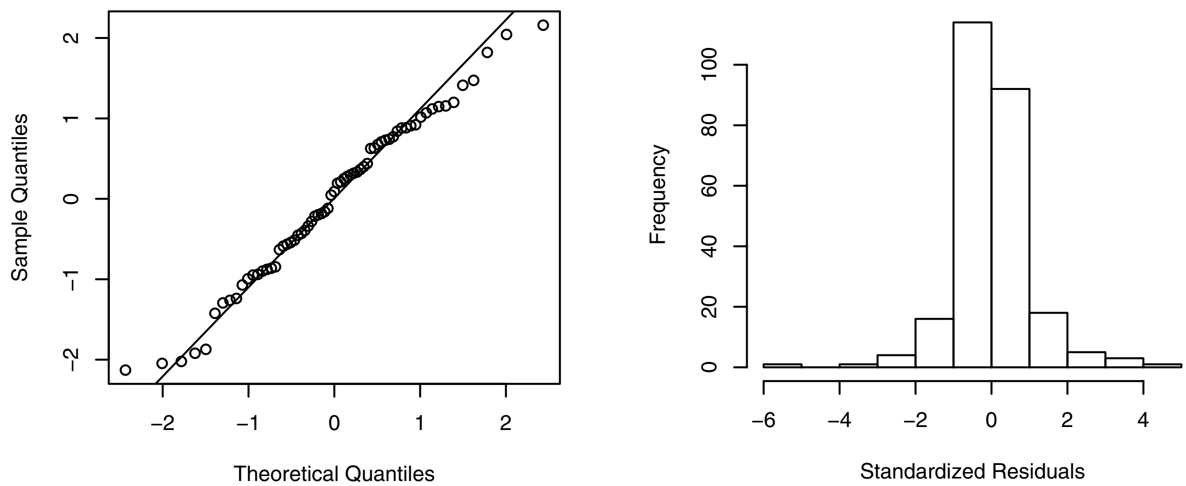
(c) Calculate the sample autocorrelations for the standardized residuals and interpret.




```
> acf(rstudent(retail.lm))
```

Here we have additional evidence that the error terms are not random—the substantial autocorrelation at the seasonal lag 12 is especially bothersome.

(d) Investigate the normality of the standardized residuals (error terms). Consider histograms and normal probability plots. Interpret the plots.




```
> qqnorm(rstudent(retail.lm)); qqline(rstudent(retail.lm))
> hist(rstudent(retail.lm),xlab='Standardized Residuals')
> shapiro.test(rstudent(retail.lm))
```

```
Shapiro-Wilk normality test
data:  rstudent(retail.lm)
W = 0.939, p-value = 8.534e-09
```

Here we see considerable evidence against normality of the error terms. The distribution is not spread out as much as a normal distribution.

Exercise 3.15 (Continuation of Exercise 3.9) Consider again the prescrip time series.

(a) Save the standardized residuals from a least squares fit of a cosine trend with fundamental frequency 1/12 to the percentage change time series.

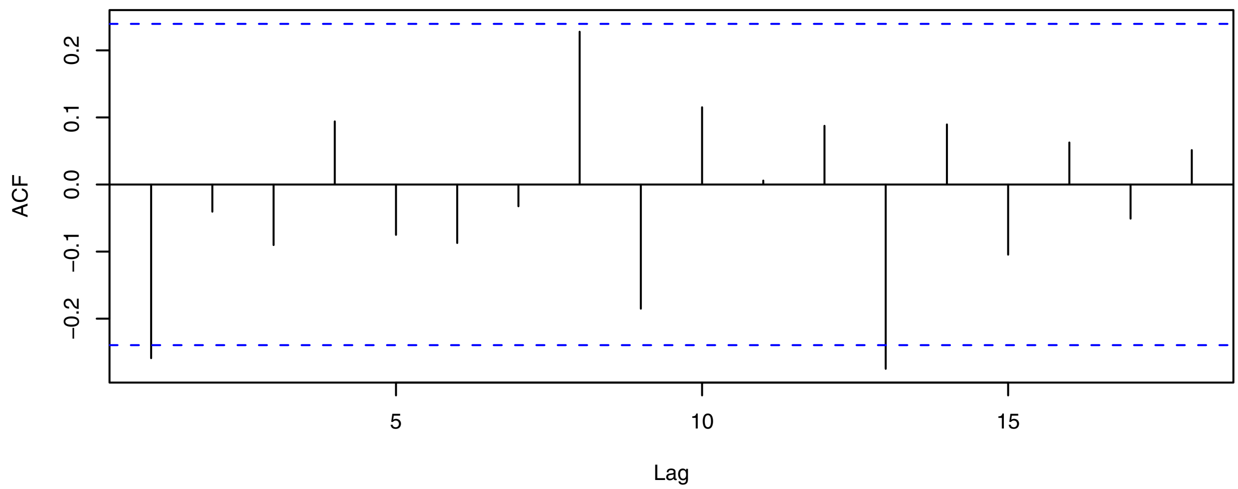
```
> perprescrip=na.omit(100*(prescrip-zlag(prescrip))/zlag(prescrip))
> har.=harmonic(perprescrip); prescrip.lm=lm(perprescrip~har.)
```

(b) Perform a Runs test on the standardized residuals and interpret the results.

```
> runs(rstudent(prescrip.lm))
      $pvalue
      [1] 0.0026
      $observed.runs
      [1] 47
      $expected.runs
      [1] 34.43284
      $n1
      [1] 32
      $n2
      [1] 35
      $k
      [1] 0
```

The runs test indicates lack of independence in the error terms of the model. The large number of runs suggests that the residuals oscillate back and forth across the median much more than a random series would.

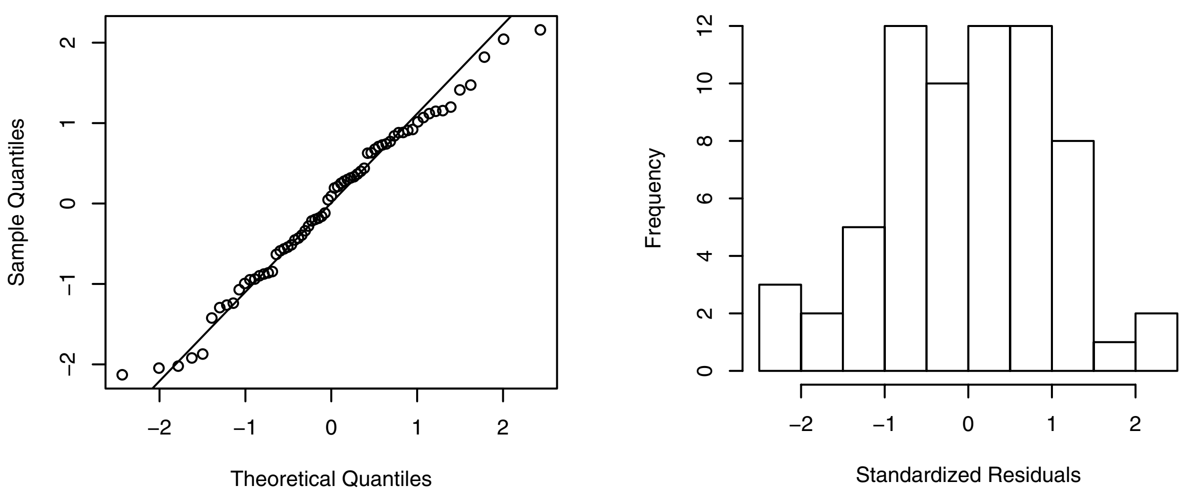
(c) Calculate the sample autocorrelations for the standardized residuals and interpret.




```
> acf(rstudent(prescrip.lm))
```

As we suspected from the runs test results, the residuals have statistically significant negative autocorrelation at lag one.

(d) Investigate the normality of the standardized residuals (error terms). Consider histograms and normal probability plots. Interpret the plots.




```
> qqnorm(rstudent(prescrip.lm)); qqline(rstudent(prescrip.lm))
> hist(rstudent(prescrip.lm),xlab='Standardized Residuals')
> shapiro.test(rstudent(prescrip.lm))
```

```
Shapiro-Wilk normality test
data:  rstudent(retail.lm)
W = 0.939, p-value = 8.534e-09
```

We have evidence against normality in the errors for this model. The distribution seems to have fatter tails” than a normal distribution would have.

Exercise 3.16 Suppose that a stationary time series, $\{Y_t\}$, has autocorrelation function of the form $\rho_k = \phi^k$ for $k > 0$ where ϕ is a constant in the range $(-1,+1)$.

(a) Show that $Var(\bar{Y}) = \frac{\gamma_0}{n}\left[\frac{1+\phi}{1-\phi} - \frac{2\phi}{n}\frac{(1-\phi^n)}{(1-\phi)^2}\right]$.

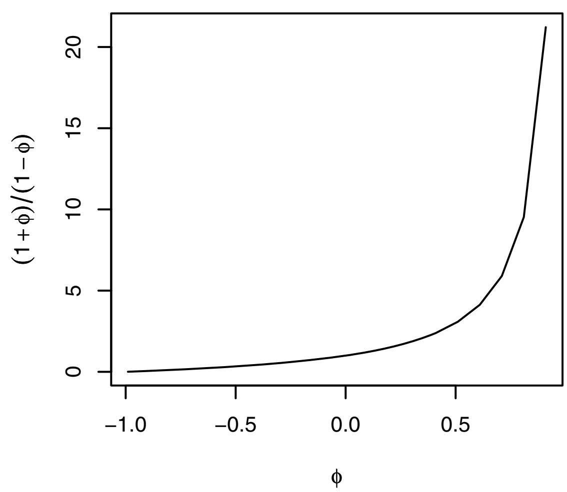
(Hint: Use Equation (3.2.3), page 28, the finite geometric sum $\sum_{k=0}^{n}\phi^k = \frac{1-\phi^{n+1}}{1-\phi}$, and the related sum $\sum_{k=0}^{n}k\phi^{k-1} = \frac{d}{d\phi}\left[\sum_{k=0}^{n}\phi^k\right]$.) First $\frac{d}{d\phi}\left[\sum_{k=0}^{n-1}\phi^k\right] = \frac{d}{d\phi}\left[\frac{1-\phi^n}{1-\phi}\right] = \left[\frac{(1-\phi)(-n\phi^{n-1})-(1-\phi)(-1)}{(1-\phi)^2}\right]$. Then

$$Var(\bar{Y}) = \frac{\gamma_0}{n}\left[1+2\sum_{k=1}^{n-1}\left(1-\frac{k}{n}\right)\rho_k\right] = \frac{\gamma_0}{n}\left[-1+2\sum_{k=0}^{n-1}\left(1-\frac{k}{n}\right)\phi^k\right] = \frac{\gamma_0}{n}\left[-1+2\sum_{k=0}^{n-1}\phi^k - \frac{2}{n}\sum_{k=0}^{n-1}k\phi^k\right]$$

$$= \frac{\gamma_0}{n}\left\{\left[-1+2\frac{(1-\phi^n)}{1-\phi}\right] - \frac{2}{n}\phi\left[\frac{(1-\phi)(-n\phi^{n-1})-(1-\phi)(-1)}{(1-\phi)^2}\right]\right\} = \frac{\gamma_0}{n}\left[\frac{1+\phi}{1-\phi} - \frac{2\phi}{n}\frac{(1-\phi^n)}{(1-\phi)^2}\right]$$

(b) If n is large argue that $Var(\bar{Y}) \approx \frac{\gamma_0}{n}\left[\frac{1+\phi}{1-\phi}\right]$. As n gets large, the term $\frac{2\phi}{n}\frac{(1-\phi^n)}{(1-\phi)^2}$ goes to zero and hence the required result.

(c) Plot $(1+\phi)/(1-\phi)$ for ϕ over the range -1 to $+1$. Interpret the plot in terms of the precision in estimating the process mean.




```
> phi=seq(from=-0.99,to=0.99,by=0.1)
> plot(y=(1+phi)/(1-phi),x=phi,type='l',xlab=expression(phi),
   ylab=expression((1+phi)/(1-phi)))
```

Negative values of ϕ imply better estimates of the mean compared with white noise but positive values imply worse estimates. This is especially true as ϕ approaches +1.

Exercise 3.17 Verify Equation (3.2.6), page 29. (Hint: You will need the fact that $\sum_{k=0}^{\infty}\phi^k = \frac{1}{1-\phi}$ for $-1 < \phi < +1$.)

$$Var(\bar{Y}) \approx \frac{\gamma_0}{n}\left[\sum_{k=-\infty}^{\infty}\rho_k\right] = \frac{\gamma_0}{n}\left[1+2\sum_{k=1}^{\infty}\phi^k\right] = \frac{\gamma_0}{n}\left[1+\frac{2\phi}{1-\phi}\right] = \frac{\gamma_0}{n}\left[\frac{1+\phi}{1-\phi}\right]$$

Exercise 3.18 Verify Equation (3.2.7), page 30. (Hint: You will need the two sums

$\sum_{t=1}^{n} t = \frac{n(n+1)}{2}$ and $\sum_{t=1}^{n} t^2 = \frac{n(n+1)(2n+1)}{6}$.)

$Var(\bar{Y}) = \sigma_e^2(2n+1)\frac{(n+1)}{6n}$ This is solved in the text on page 30. You could construct an alternative proof using Equation (2.2.7), page 12, and Equation (2.2.12), page 13, but it would be longer.

CHAPTER 4

Exercise 4.1 Use first principles to find the autocorrelation function for the stationary process defined by $Y_t = 5 + e_t - \frac{1}{2}e_{t-1} + \frac{1}{4}e_{t-2}$.

First $Var(Y_t) = Var(5 + e_t - \frac{1}{2}e_{t-1} + \frac{1}{4}e_{t-2}) = [1 + (\frac{1}{2})^2 + (\frac{1}{4})^2]\sigma_e^2 = \frac{21}{16}\sigma_e^2$ then

$$Cov(Y_t, Y_{t-1}) = Cov(e_t - \tfrac{1}{2}e_{t-1} + \tfrac{1}{4}e_{t-2}, e_{t-1} - \tfrac{1}{2}e_{t-2} + \tfrac{1}{4}e_{t-3}) = Cov(-\tfrac{1}{2}e_{t-1} + \tfrac{1}{4}e_{t-2}, e_{t-1} - \tfrac{1}{2}e_{t-2})$$
$$= Cov(-\tfrac{1}{2}e_{t-1}, e_{t-1}) + Cov(\tfrac{1}{4}e_{t-2}, -\tfrac{1}{2}e_{t-2}) = [-\tfrac{1}{2}(\tfrac{1}{4}) - \tfrac{1}{2}]\sigma_e^2 = -\tfrac{5}{8}\sigma_e^2$$

$Cov(Y_t, Y_{t-2}) = Cov(e_t - \frac{1}{2}e_{t-1} + \frac{1}{4}e_{t-2}, e_{t-2} - \frac{1}{2}e_{t-3} + \frac{1}{4}e_{t-4}) = Cov(\frac{1}{4}e_{t-2}, e_{t-2}) = \frac{1}{4}\sigma_e^2$, and

$Cov(Y_t, Y_{t-3}) = Cov(e_t - \frac{1}{2}e_{t-1} + \frac{1}{4}e_{t-2}, e_{t-3} - \frac{1}{2}e_{t-4} + \frac{1}{4}e_{t-5}) = 0$ and this persists for all lags 3 or more.

Therefore

$$\rho_k = \begin{cases} 1 & k = 0 \\ \dfrac{-\frac{5}{8}\sigma_e^2}{\frac{21}{16}\sigma_e^2} = -\frac{10}{21} & k = 1 \\ \dfrac{\frac{1}{4}\sigma_e^2}{\frac{21}{16}\sigma_e^2} = \frac{4}{21} & k = 2 \\ 0 & k > 2 \end{cases}$$

Exercise 4.2 Sketch the autocorrelation functions for the following MA(2) models with parameters as specified:

(a) $\theta_1 = 0.5$ and $\theta_2 = 0.4$

You could do these calculations using Equation (4.2.3), page 63, with a calculator. Or you could use the following R code:

```
> ARMAacf(ma=list(-.5,-.4))
```

```
            0          1          2          3
    1.0000000 -0.2127660 -0.2836879  0.0000000
```

(b) $\theta_1 = 1.2$ and $\theta_2 = -0.7$

```
> ARMAacf(ma=list(-1.2,.7))
```

```
        0          1          2          3
    1.0000000 -0.6962457  0.2389078  0.0000000
```

(c) $\theta_1 = -1$ and $\theta_2 = -0.6$

```
> ARMAacf(ma=list(1,.6))
```

```
        0         1         2         3
    1.0000000 0.6779661 0.2542373 0.0000000
```

Exercise 4.3 Verify that for an MA(1) process

$$\max_{-\infty < \theta < \infty} \rho_1 = 0.5 \qquad \text{and} \qquad \min_{-\infty < \theta < \infty} \rho_1 = -0.5$$

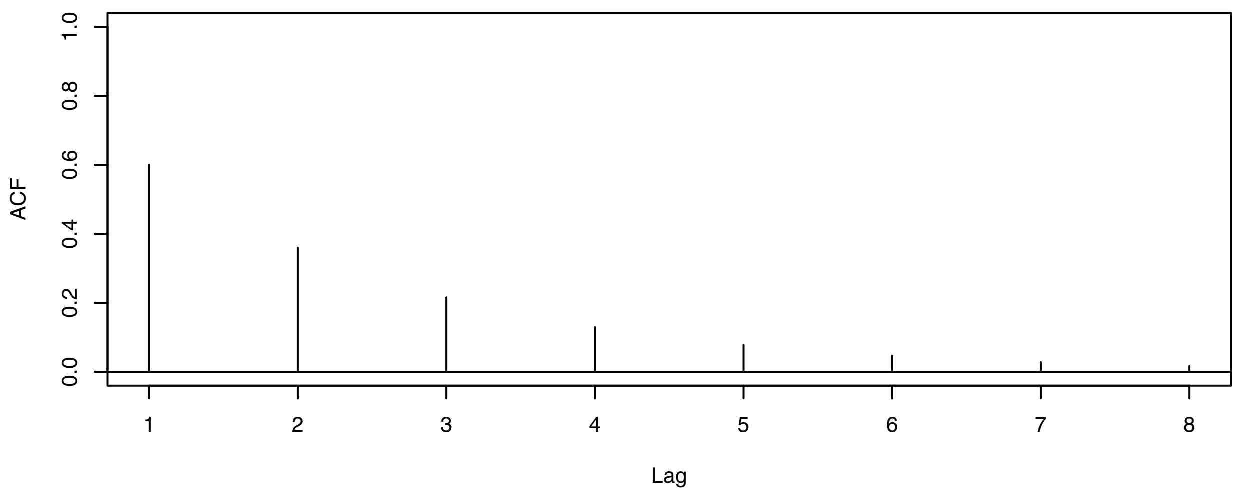
Recall that $\rho_1 = \dfrac{-\theta}{1+\theta^2}$ so that $\dfrac{d\rho_1}{d\theta} = \dfrac{\theta^2 - 1}{(1+\theta^2)^2}$ which is seen to be negative on the range $-1 < \theta < +1$ and zero at both -1 and $+1$. For $|\theta| > 1$, the derivative is positive. Taken together, these facts imply the desired results.

Exercise 4.4 Show that when θ is replaced by $1/\theta$, the autocorrelation function for an MA(1) process does not change.

$$\frac{-\frac{1}{\theta}}{1 + \left(\frac{1}{\theta}\right)^2} = \frac{-\theta}{1+\theta^2}$$ as required.

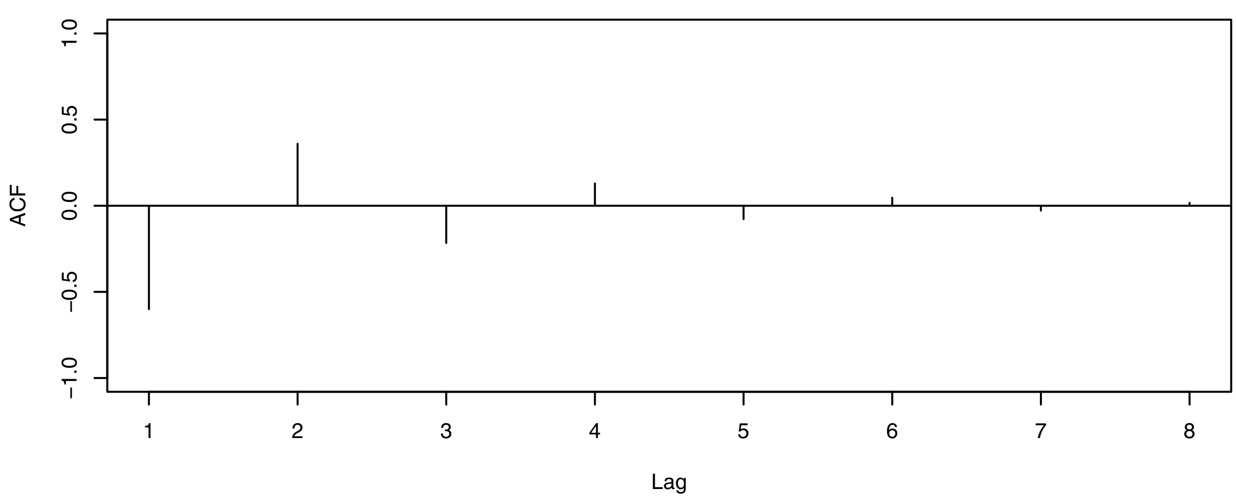
Exercise 4.5 Calculate and sketch the autocorrelation functions for each of the following AR(1) models. Plot for sufficient lags that the autocorrelation function has nearly died out.

(a) $\phi_1 = 0.6$



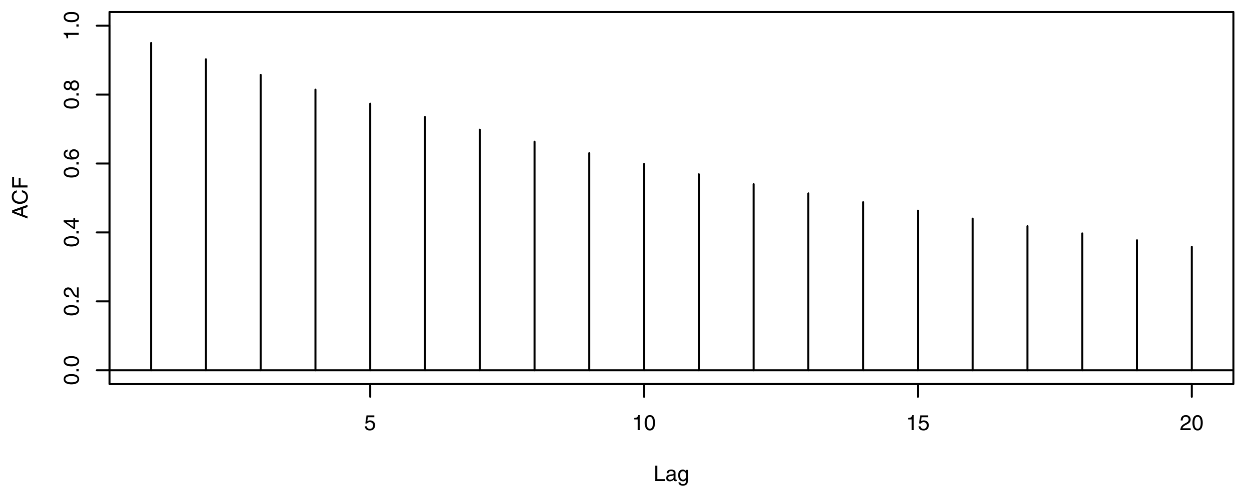

```
> ACF=ARMAacf(ar=.6,lag.max=8)
> plot(y=ACF[-1],x=1:8,xlab='Lag',ylab='ACF',type='h'); abline(h=0)
```

(b) $\phi_1 = -0.6$



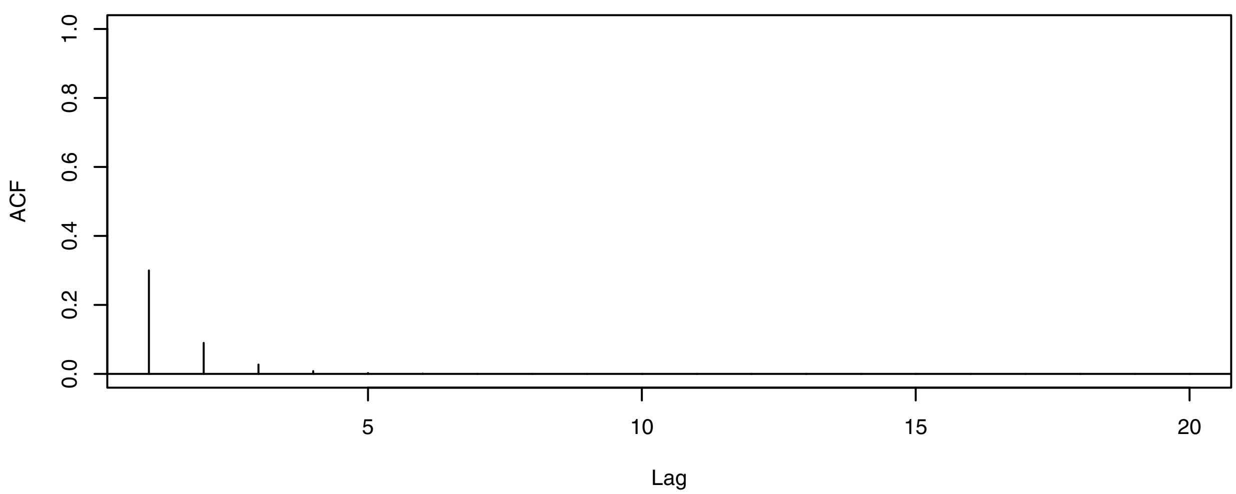
```
> ACF=ARMAacf(ar=-0.6,lag.max=8)
> plot(y=ACF[-1],x=1:8,xlab='Lag',ylab='ACF',type='h',ylim=c(-1,1)); abline(h=0)
```

(c) $\phi_1 = 0.95$ (Do out to 20 lags.)




```
> ACF=ARMAacf(ar=.95,lag.max=20)
> plot(y=ACF[-1],x=1:20,xlab='Lag',ylab='ACF',type='h',ylim=c(0,1)); abline(h=0)
```

(d) $\phi_1 = 0.3$




```
> ACF=ARMAacf(ar=.3,lag.max=20)
> plot(y=ACF[-1],x=1:20,xlab='Lag',ylab='ACF',type='h',ylim=c(0,1)); abline(h=0
```

Exercise 4.6 Suppose that $\{Y_t\}$ is an AR(1) process with $-1 < \phi < +1$.

(a) Find the autocovariance function for $W_t = \nabla Y_t = Y_t - Y_{t-1}$ in terms of ϕ and σ_e^2.

Recall that for $k \geq 0$ $Cov(Y_t, Y_{t-k}) = \phi^k \frac{\sigma_e^2}{1-\phi^2}$ Then for $k > 0$

$$\begin{aligned} Cov(W_t, W_{t-k}) &= Cov(Y_t - Y_{t-1}, Y_{t-k} - Y_{t-k-1}) \\ &= (Cov(Y_t, Y_{t-k}) + Cov(Y_t, -Y_{t-k-1}) + Cov(-Y_{t-1}, Y_{t-k}) + Cov(-Y_{t-1}, -Y_{t-k-1})) \end{aligned}$$

$$= [\phi^k - \phi^{k+1} - \phi^{k-1} + \phi^k]\frac{\sigma_e^2}{1-\phi^2} = [2\phi^k - \phi^{k+1} - \phi^{k-1}]\frac{\sigma_e^2}{1-\phi^2} = [2\phi - \phi^2 - 1]\phi^{k-1}\frac{\sigma_e^2}{1-\phi^2}$$

$$= -[(1-\phi)^2]\phi^{k-1}\frac{\sigma_e^2}{1-\phi^2} = -\left[\frac{1-\phi}{1+\phi}\right]\phi^{k-1}\sigma_e^2$$

(b) In particular, show that $Var(W_t) = 2\sigma_e^2/(1+\phi)$.

$$Var(W_t) = Var(Y_t - Y_{t-1}) = Var(Y_t) + Var(Y_{t-1}) - 2Cov(Y_t, Y_{t-1}) = 2(1-\phi)\frac{\sigma_e^2}{1-\phi^2} = \frac{2\sigma_e^2}{1+\phi}$$

Exercise 4.7 Describe the important characteristics of the autocorrelation function for the following models: (a) MA(1), (b) MA(2), (c) AR(1), (d) AR(2), and (e) ARMA(1,1).

(a) Has nonzero correlation only at lag 1. Could be positive or negative but must be between -0.5 and +0.5.

(b) Has nonzero correlation only at lags 1 and 2.

(c) Has exponentially decaying autocorrelations starting from lag 0. If $\phi > 0$, then all autocorrelations are positive. If $\phi < 0$, then autocorrelations alternate negative, positive negative, etc.

(d) Autocorrelations can have several patterns but if the roots of the charcteristic equation are complex numbers, then the pattern will be a cosine with a decaying magnitude.

(e) Has exponentially decaying autocorrelations starting from lag 1—but not from lag zero.

Exercise 4.8 Let $\{Y_t\}$ be an AR(2) process of the special form $Y_t = \phi_2 Y_{t-2} + e_t$. Use first principles to find the range of values of ϕ_2 for which the process is stationary.

If $\{Y_t\}$ is stationary, then $Var(Y_t) = \phi_2^2 Var(Y_{t-2}) + \sigma_e^2$. But, by stationarity $Var(Y_t) = Var(Y_{t-2})$ so solving the first equation gives

$Var(Y_t) = \dfrac{\sigma_e^2}{1-\phi_2^2}$ Hence, we must have $-1 < \phi_2 < +1$.

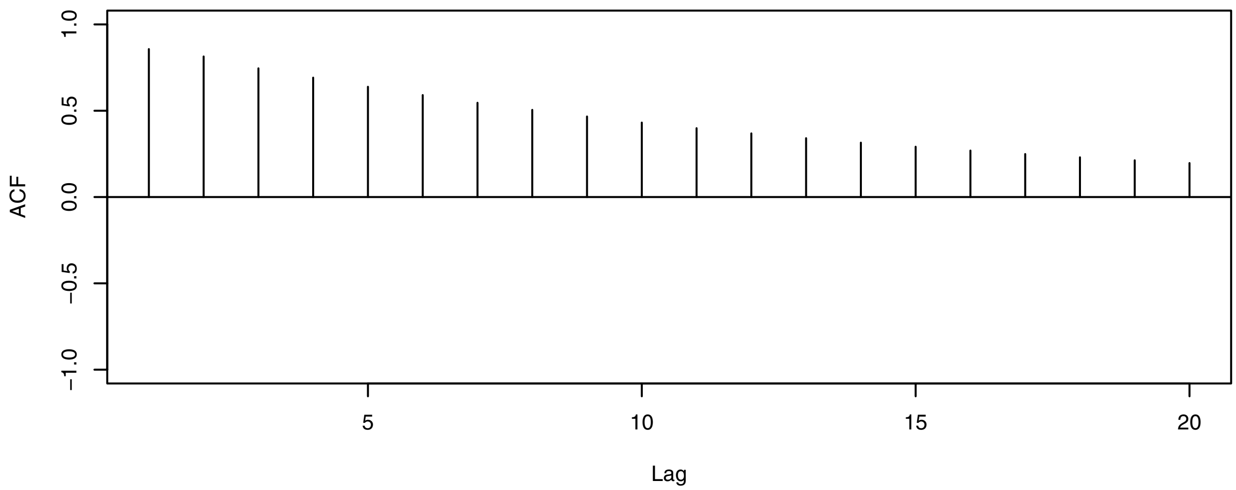
Exercise 4.9 Use the recursive formula of Equation (4.3.13), page 72, to calculate and then sketch the autocorrelation functions for the following AR(2) models with parameters as specified. In each case specify whether the roots of the characteristic equation are real or complex. If the roots are complex, find the damping factor, R, and frequency, Θ, for the corresponding autocorrelation function when expressed as in Equation (4.3.17), page 73.

(a) $\phi_1 = 0.6$ and $\phi_2 = 0.3$

These calculations could be done on a calculator or with the following R code

```
> rho=NULL; phi1=.6; phi2=.3; max.lag=20
> rho1=phi1/(1-phi2); rho2=(phi2*(1-phi2)+phi1^2)/(1-phi2)
> rho[1]=rho1; rho[2]=rho2
> for (k in 3:max.lag) rho[k]=phi1*rho[k-1]+phi2*rho[k-2]
> rho # to display the values
> plot(y=rho,x=1:max.lag,type='h',ylab='ACF',xlab='Lag',ylim=c(-1,+1)); abline(h=0)
```

```
> rho
 [1] 0.8571429 0.8142857 0.7457143 0.6917143 0.6387429 0.5907600 0.5460789
 [8] 0.5048753 0.4667488 0.4315119 0.3989318 0.3688126 0.3409671 0.3152241
[15] 0.2914246 0.2694220 0.2490806 0.2302749 0.2128891 0.1968159
```

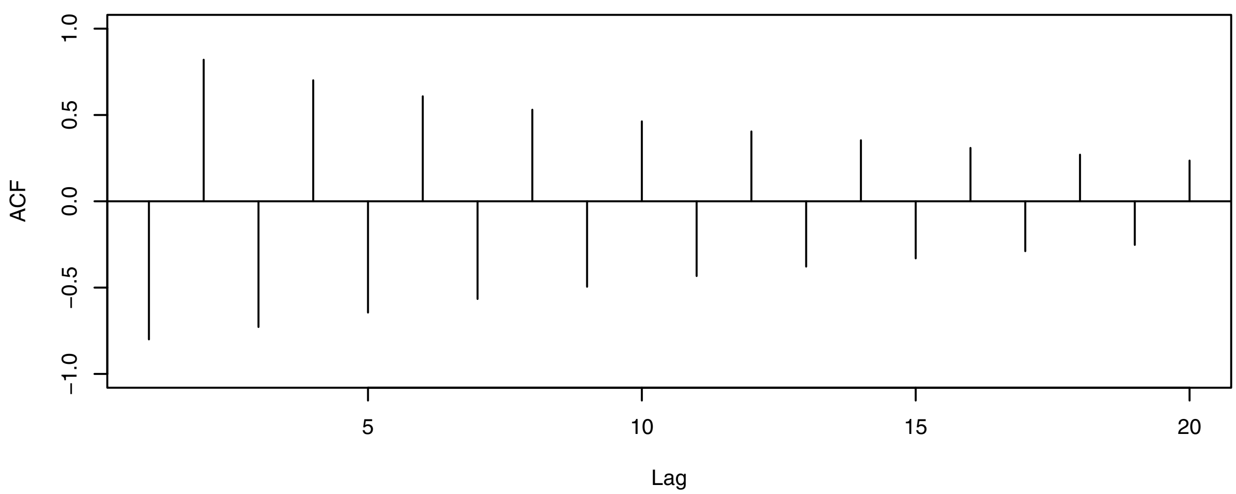


```
> polyroot(c(1,-phi1,-phi2)) # get the roots of the characteristic polynomial
```

```
[1]  1.081666+0i -3.081666+0i
```

Note that the roots are real and that one root is very close to the stationarity boundary (+1). This explains the slow decay of the autocorrelation function. You could also place the point (ϕ_1,ϕ_2) on the display in Exhibit (4.17), page 72, to show that the roots are real.

(b) $\phi_1 = -0.4$ and $\phi_2 = 0.5$

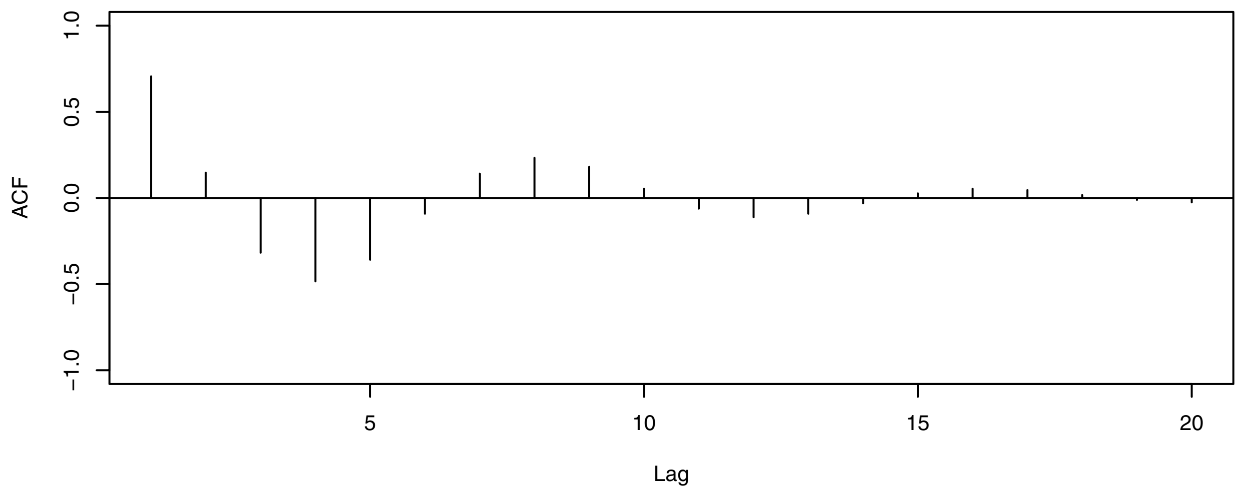


```
> polyroot(c(1,-phi1,-phi2))
```

```
> [1] -1.069694+0i  1.869694-0i
```

The roots are real in this case. If you look at the actual ACF values you will discover that the magnitude of the lag 2 value (0.82) is slightly larger than the magnitude of the lag one value (0.80). Of course, this could not happen with an AR(1) process.

(c) $\phi_1 = 1.2$ and $\phi_2 = -0.7$




```
> polyroot(c(1,-phi1,-phi2))
```

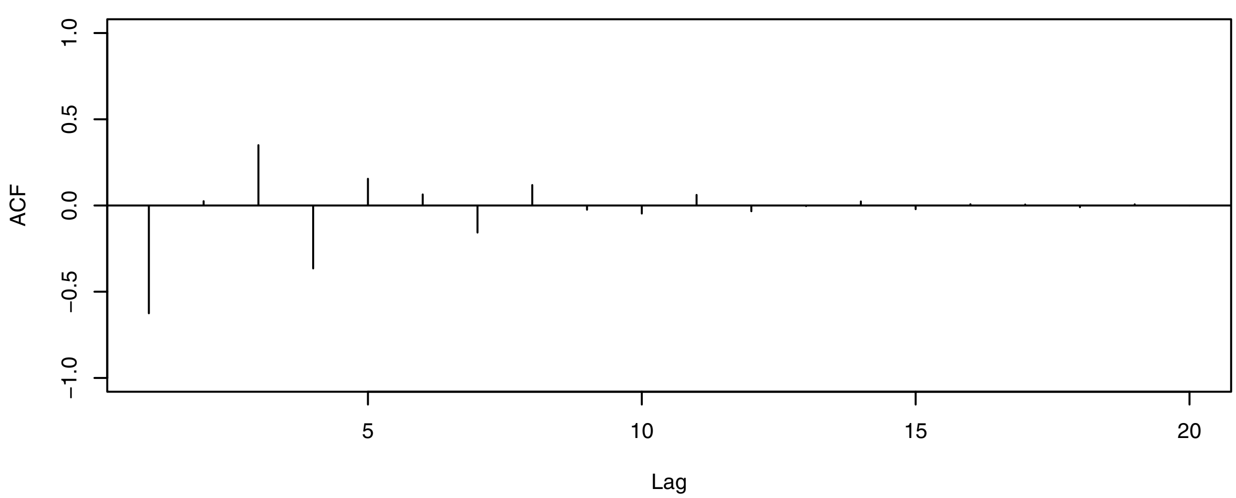
```
[1] 0.8571429+0.8329931i 0.8571429-0.8329931i
```

In this case the roots are complex.

```
> Damp = sqrt(-phi2) # damping factor R
> Freq = acos(phi1/(2*R)) # frequency Θ
> Phase = atan((1-phi2)/(1+phi2)) # phase Φ
> Damp; Freq; Phase # display the results
```

```
[1] 0.83666
[1] 0.7711105
[1] 1.396124
```

(d) $\phi_1 = -1$ and $\phi_2 = -0.6$




```
> > polyroot(c(1,-phi1,-phi2))
```

```
[1] -0.8333333+0.9860133i -0.8333333-0.9860133i
```

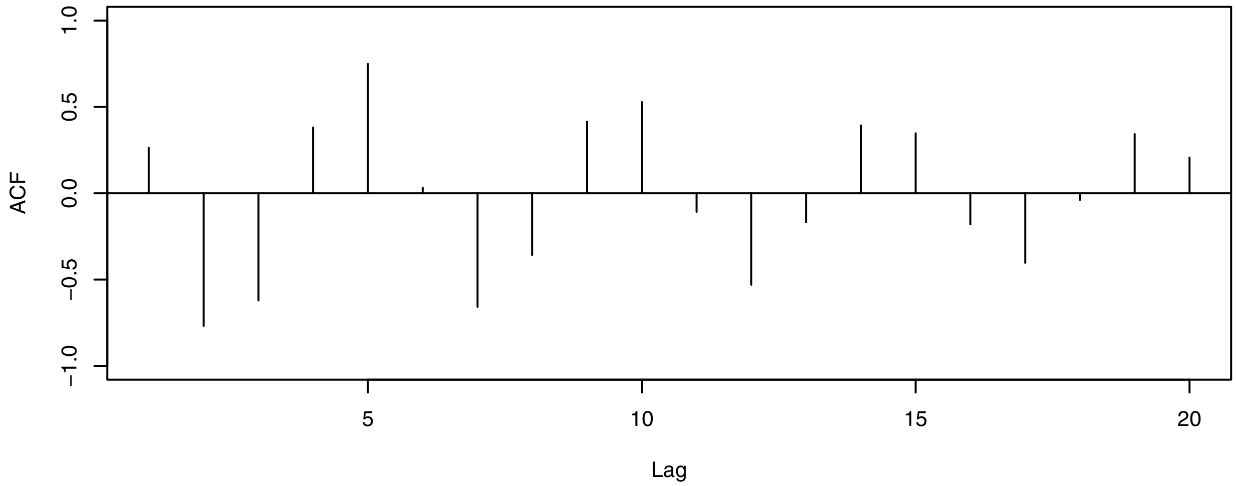
The roots are complex and we have

```
> Damp; Freq; Phase # display the results
```

```
[1] 0.7745967
```

```
[1] 2.211319
[1] 1.325818
```

(e) $\phi_1 = 0.5$ and $\phi_2 = -0.9$




```
> polyroot(c(1,-phi1,-phi2))
```

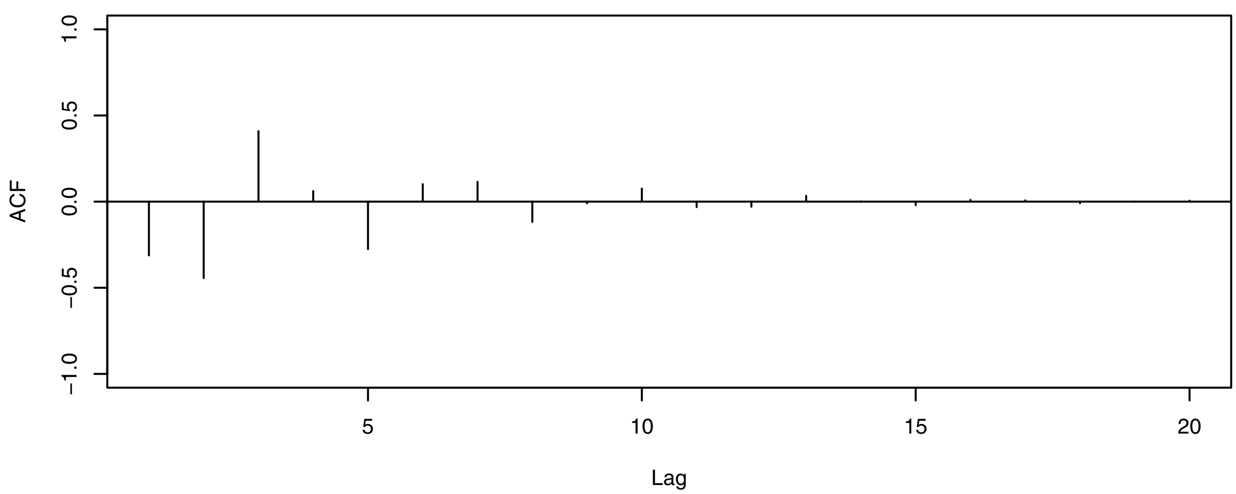
```
[1] 0.277778+1.016834i 0.277778-1.016834i
```

The roots are complex and

```
> Damp; Freq; Phase # display the results
```

```
[1] 0.9486833
[1] 1.267354
[1] 1.518213
```

(f) $\phi_1 = -0.5$ and $\phi_2 = -0.6$




```
> polyroot(c(1,-phi1,-phi2))
```

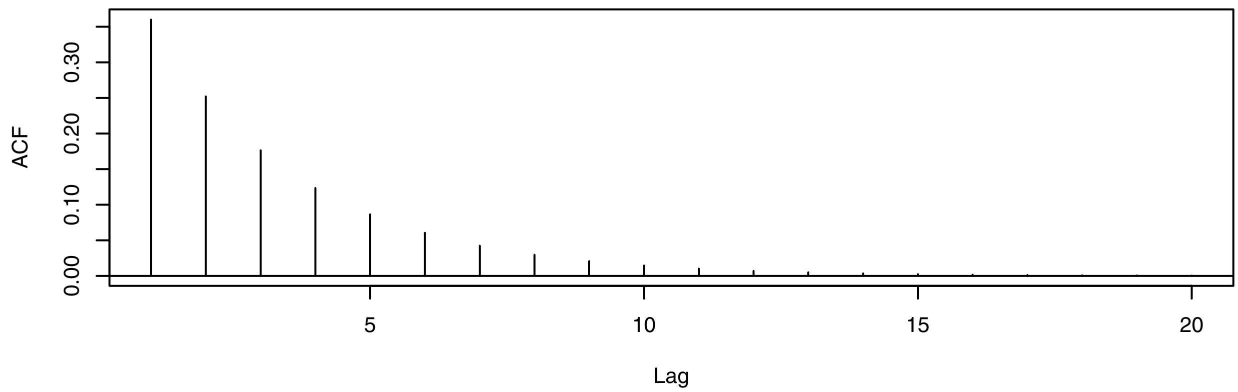
```
[1] -0.416667+1.221907i -0.416667-1.221907i
```

The roots are complex with

```
> Damp; Freq; Phase # display the results
[1] 0.7745967
[1] 1.874239
[1] 1.325818
```

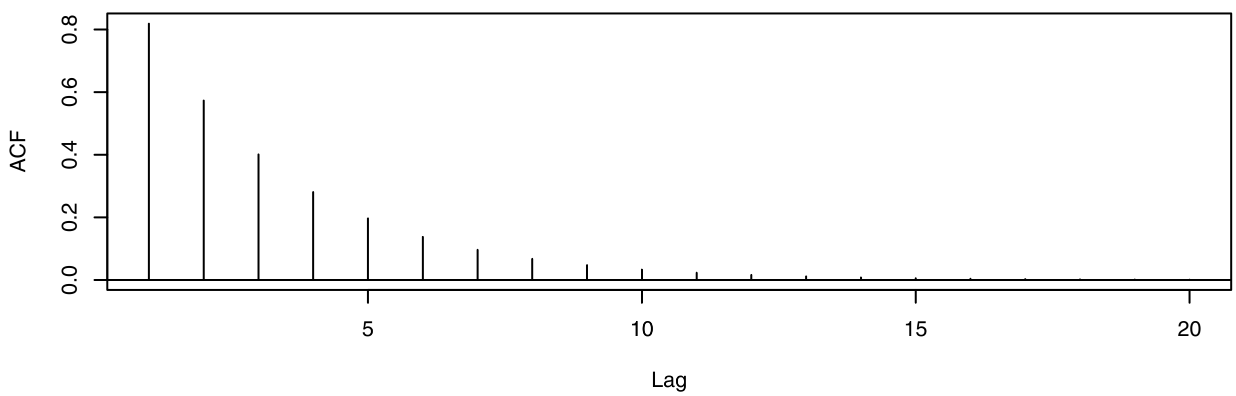
Exercise 4.10 Sketch the autocorrelation functions for each of the following ARMA models:

(a) ARMA(1,1) with $\phi = 0.7$ and $\theta = 0.4$. Could use Equation (4.4.5), page 78, and a calculator or use the R code which follows.




```
> ACF=ARMAacf(ar=0.7,ma=-0.4,lag.max=20)
> # Remember that R uses the negative of our theta values.
> plot(y=ACF[-1],x=1:20,xlab='Lag',ylab='ACF',type='h'); abline(h=0)
```

(b) ARMA(1,1) with $\phi = 0.7$ and $\theta = -0.4$.




```
> ACF=ARMAacf(ar=.7,ma=0.4,lag.max=20)
> plot(y=ACF[-1],x=1:20,xlab='Lag',ylab='ACF',type='h'); abline(h=0)
```

Exercise 4.11 For the ARMA(1,2) model $Y_t = 0.8Y_{t-1} + e_t + 0.7e_{t-1} + 0.6e_{t-2}$ show that

(a) $\rho_k = 0.8\rho_{k-1}$ for $k > 2$. With no loss of generality we assume the mean of the series is zero. Then

$$\begin{aligned} Cov(Y_t, Y_{t-k}) &= E[(0.8Y_{t-1} + e_t + 0.7e_{t-1} + 0.6e_{t-2})Y_{t-k}] \\ &= 0.8E(Y_{t-1}Y_{t-k}) + E(e_tY_{t-k}) + 0.7E(e_{t-1}Y_{t-k}) + 0.6E(e_{t-2}Y_{t-k}) \\ &= 0.8E(Y_{t-1}Y_{t-k}) + 0 + 0 + 0 \qquad \text{since } k > 2 \\ &= 0.8Cov(Y_t, Y_{t-(k-1)}) \end{aligned}$$

or $\gamma_k = 0.8\gamma_{k-1}$ and hence $\rho_k = 0.8\rho_{k-1}$ for $k > 2$.

(b) $\rho_2 = 0.8\rho_1 + 0.6\sigma_e^2/\gamma_0$.

$$\begin{aligned} Cov(Y_t, Y_{t-2}) &= E[(0.8Y_{t-1} + e_t + 0.7e_{t-1} + 0.6e_{t-2})Y_{t-2}] \\ &= E[(0.8Y_{t-1} + 0.6e_{t-2})Y_{t-2}] \\ &= 0.8Cov(Y_{t-1}, Y_{t-2}) + 0.6E(e_{t-2}Y_{t-2}) \end{aligned}$$

$$\text{or } \gamma_2 = 0.8\gamma_1 + 0.6E(e_{t-2}Y_{t-2})$$

Now $E(e_{t-2}Y_{t-2}) = E(e_tY_t) = E[e_t(0.8Y_{t-1} + e_t + 0.7e_{t-1} + 0.6e_{t-2})] = 0 + \sigma_e^2 + 0 + 0 = \sigma_e$ and the required result follows.

Exercise 4.12 Consider two MA(2) processes, one with $\theta_1 = \theta_2 = 1/6$ and another with $\theta_1 = -1$ and $\theta_2 = 6$.

(a) Show that these processes have the same autocorrelation function.

Plug these numbers into Equations (4.2.3), page 63, and you get the same results. Alternatively, use the following R code.

```
> ARMAacf(ma=c(1,-6))
> ARMAacf(ma=c(-1/6,-1/6))
```

(b) How do the roots of the corresponding characteristic polynomials compare?

Notice that $1 - \frac{1}{6}x - \frac{1}{6}x^2 = -\frac{1}{6}(x+3)(x-2)$ while $1 + x - 6x^2 = -6\left(x + \frac{1}{3}\right)\left(x - \frac{1}{2}\right)$. So the roots of the two polynomials are reciprocals of one another. Only the MA(2) model with $\theta_1 = \theta_2 = 1/6$ is invertible.

Exercise 4.13 Let $\{Y_t\}$ be a stationary process with $\rho_k = 0$ for $k > 1$. Show that we must have $|\rho_1| \le ½$. (Hint: Consider $Var(Y_{n+1} + Y_n + \cdots + Y_1)$ and then $Var(Y_{n+1} - Y_n + Y_{n-1} - \cdots \pm Y_1)$. Use the fact that both of these must be nonnegative for *all* n.)

$$Var(Y_{n+1} + Y_n + Y_{n-1} + \cdots + Y_1) = [(n+1) + 2n\rho_1]\gamma_0 = [1 + n(1 + 2\rho_1)]\gamma_0$$

$$Var(Y_{n+1} - Y_n + Y_{n-1} - \cdots \pm Y_1) = [(n+1) - 2n\rho_1]\gamma_0 = [1 + n(1 - 2\rho_1)]\gamma_0$$

So we must have both $1 + n(1 + 2\rho_1) \ge 0$ and $1 + n(1 - 2\rho_1) \ge 0$ for all n. The first of these inequalities is equivalent to $\rho_1 \ge -0.5((n+1)/n)$. Since this must hold for all n, we must have $\rho_1 \ge -0.5$. The inequality $\rho_1 \le 0.5$ follows similarly from the other inequality.

Exercise 4.14 Suppose that $\{Y_t\}$ is a zero mean, stationary process with $|\rho_1| < 0.5$ and $\rho_k = 0$ for $k > 1$. Show that $\{Y_t\}$ must be representable as an MA(1) process. That is, show that there is a white noise sequence $\{e_t\}$ such that $Y_t = e_t - \theta e_{t-1}$ where ρ_1 is correct and e_t is uncorrelated with Y_{t-k} for $k > 0$. (Hint: Choose θ such that $|\theta| < 1$ and $\rho_1 = -\theta/(1+\theta^2)$; then let $e_t = \sum_{j=0}^{\infty} \theta^j Y_{t-j}$. If we assume that $\{Y_t\}$ is a normal process, e_t will also be normal and zero correlation is equivalent to independence.)

From Exhibit (4.1), page 58, we know that there exists a unique θ in $(-1,+1)$ such that $\rho_1 = -\theta/(1+\theta^2)$. With that θ define e_t as $e_t = \sum_{j=0}^{\infty} \theta^j Y_{t-j}$. It is then straightforward to show that $Y_t = e_t - \theta e_{t-1}$ as required and that e_t is uncorrelated with Y_{t-k} for $k > 0$.

Exercise 4.15 Consider the AR(1) model $Y_t = \phi Y_{t-1} + e_t$. Show that if $|\phi| = 1$ the process cannot be stationary. (Hint: Take variances of both sides.)

Suppose $\{Y_t\}$ is stationary. Then $Var(Y_t) = \phi^2 Var(Y_{t-1}) + \sigma_e^2$ or $Var(Y_t) = \sigma_e^2/(1-\phi^2)$. If $|\phi| = 1$ this is impossible and we have a proof by contradiction.

Exercise 4.16 Consider the "nonstationary" AR(1) model $Y_t = 3Y_{t-1} + e_t$.

(a) Show that $Y_t = -\sum_{j=1}^{\infty} \left(\frac{1}{3}\right)^j e_{t+j}$ satisfies the AR(1) equation.

This follows by straightforward substitution.

(b) Show that the process defined in part (a) is stationary.

The calculations leading up to Equation (4.1.3), page 56, can be repeated with this situation to show stationarity.

(c) In what way is this solution unsatisfactory?

Since Y_t at time t depends on *future* error terms, this is an unsatisfactory model.

Exercise 4.17 Consider a process that satisfies the AR(1) equation $Y_t = ½Y_{t-1} + e_t$.

(a) Show that $Y_t = 10(½)^t + e_t + ½e_{t-1} + (½)^2e_{t-2} + \cdots$ is a solution of the AR(1) equation.

It is easy to see that $10(½)^t + e_t + ½e_{t-1} + (½)^2e_{t-2} + \cdots = ½[10(½)^{t-1} + e_{t-1} + ½e_{t-2} + (½)^2e_{t-3} + \cdots] + e_t$ so Y_t is indeed a solution of the AR(1) equation.

(b) Is the solution given in part (a) stationary?

Since $E(Y_t) = 10(½)^t$ is not constant in time, the solution is not stationary.

Exercise 4.18 Consider a process that satisfies the zero-mean, "stationary" AR(1) equation $Y_t = \phi Y_{t-1} + e_t$ with $-1 < \phi < +1$. Let c be any nonzero constant and define $W_t = Y_t + c\phi^t$.

(a) Show that $E(W_t) = c\phi^t$.

$E(W_t) = E(Y_t + c\phi^t) = 0 + c\phi^t = c\phi^t$.

(b) Show that $\{W_t\}$ satisfies the "stationary" AR(1) equation $W_t = \phi W_{t-1} + e_t$.

$Y_t + c\phi^t = \phi[Y_{t-1} + c\phi^{t-1}] + e_t$ is valid since the terms $c\phi^t$ on both sides cancel.

(c) Is $\{W_t\}$ stationary?

Since $E(W_t) = c\phi^t$ is not constant in time, the solution is not stationary.

Exercise 4.19 Consider an MA(6) model with $\theta_1 = 0.5$, $\theta_2 = -0.25$, $\theta_3 = 0.125$, $\theta_4 = -0.0625$, $\theta_5 = 0.03125$, and $\theta_6 = -0.015625$. Find a much simpler model that has nearly the same ψ-weights.

Notice that these coefficients decrease exponentially in magnitude at rate 0.5 while alternating in sign. Furthermore, the coefficients have nearly died out by θ_6. Thus, an AR(1) process with $\phi = -0.5$ would be nearly the same process.

Exercise 4.20 Consider an MA(7) model with $\theta_1 = 1$, $\theta_2 = -0.5$, $\theta_3 = 0.25$, $\theta_4 = -0.125$, $\theta_5 = 0.0625$, $\theta_6 = -0.03125$, and $\theta_7 = 0.015625$. Find a much simpler model that has nearly the same ψ-weights.

Notice that these coefficients decrease exponentially in magnitude at rate 0.5 while alternating in sign but starting at $\theta_1 = 1$. Furthermore, the coefficients have nearly died out by θ_7. Equation (4.4.6), page 78, shows this type of behaviour for an ARMA(1,1) process. Equating $\psi_1 = \phi - \theta = 1$ and $\psi_2 = (\phi - \theta)\phi = -0.5$ yields $\phi = -0.5$ and $\theta = 0.5$ in the ARMA(1,1) model that is nearly the same.

Exercise 4.21 Consider the model $Y_t = e_{t-1} - e_{t-2} + 0.5e_{t-3}$.

(a) Find the autocovariance function for this process.

$$Var(Y_t) = [1 + (-1)^2 + (0.5)^2]\sigma_e^2 = 2.25\sigma_e^2 \text{ and}$$

$$\begin{aligned} Cov(Y_t, Y_{t-1}) &= Cov(e_{t-1} - e_{t-2} + 0.5e_{t-3}, e_{t-2} - e_{t-3} + 0.5e_{t-4}) \\ &= Cov(-e_{t-2}, e_{t-2}) + Cov(0.5e_{t-3}, -e_{t-3}) \\ &= (-1 - 0.5)\sigma_e^2 \\ &= -1.5\sigma_e^2 \end{aligned} \qquad \text{So } \rho_1 = \frac{-1.5\sigma_e^2}{2.25\sigma_e^2} = -\frac{2}{3}$$

$$\begin{aligned} Cov(Y_t, Y_{t-2}) &= Cov(e_{t-1} - e_{t-2} + 0.5e_{t-3}, e_{t-3} - e_{t-4} + 0.5e_{t-5}) \\ &= Cov(0.5e_{t-3}, e_{t-3}) \\ &= 0.5\sigma_e^2 \end{aligned} \qquad \text{and so } \rho_1 = \frac{0.5\sigma_e^2}{2.25\sigma_e^2} = \frac{2}{9}.$$

All other autocorrelations are zero.

(b) Show that this is a certain ARMA(p,q) process in disguise. That is, identify values for p and q, and for the θ's and ϕ's such that the ARMA(p,q) process has the same statistical properties as $\{Y_t\}$.

This is really just the MA(2) process $Y_t = e_t - e_{t-1} + 0.5e_{t-2}$ in disguise. Since we do not observe the error terms, there is no way to tell the difference between the two sequences defined as e_t and $e_t' = e_{t-1}$.

Exercise 4.22 Show that the statement "The roots of $1 - \phi_1 x - \phi_2 x^2 - \cdots - \phi_p x^p = 0$ are greater than 1 in absolute value" is equivalent to the statement "The roots of $x^p - \phi_1 x^{p-1} - \phi_2 x^{p-2} - \ldots - \phi_p = 0$ are less than 1 in absolute value." (Hint: If G is a root of one equation, is $1/G$ a root of the other?)

Let G be a root of $1 - \phi_1 x - \phi_2 x^2 - \cdots - \phi_p x^p = 0$. Then

$0 = 1 - \phi_1 G - \phi_2 G^2 - \cdots - \phi_p G^p = G^p\left[\left(\frac{1}{G}\right)^p - \phi_1\left(\frac{1}{G}\right)^{p-1} - \phi_2\left(\frac{1}{G}\right)^{p-2} - \cdots - \phi_p\right]$. So $1/G$ is a root of

$x^p - \phi_1 x^{p-1} - \phi_2 x^{p-2} - \ldots - \phi_p = 0$.

Exercise 4.23 Suppose that $\{Y_t\}$ is an AR(1) process with $\rho_1 = \phi$. Define the sequence $\{b_t\}$ as $b_t = Y_t - \phi Y_{t+1}$.

(a) Show that $Cov(b_t, b_{t-k}) = 0$ for all t and $k \neq 0$. Without loss of generality, assume $Var(Y_t) = 1$. Then

$$\begin{aligned} Cov(b_t, b_{t-k}) &= Cov(Y_t - \phi Y_{t+1}, Y_{t-k} - \phi Y_{t-k+1}) \\ &= Cov(Y_t, Y_{t-k}) - \phi Cov(Y_t, Y_{t-k+1}) - \phi Cov(Y_{t+1}, Y_{t-k}) + \phi^2 Cov(Y_{t+1}, Y_{t-k+1}) \\ &= (\phi^k - \phi\phi^{k-1} - \phi\phi^{k+1} + \phi^2\phi^k) \\ &= 0 \end{aligned}$$

(b) Show that $Cov(b_t, Y_{t+k}) = 0$ for all t and $k > 0$.

$$\begin{aligned} Cov(b_t, Y_{t+k}) &= Cov(Y_t - \phi Y_{t+1}, Y_{t+k}) = Cov(Y_t, Y_{t+k}) - \phi Cov(Y_{t+1}, Y_{t+k}) \\ &= \phi^k - \phi\phi^{k-1} \\ &= 0 \end{aligned}$$

Exercise 4.24 Let $\{e_t\}$ be a zero mean, unit variance white noise process. Consider a process that begins at time $t = 0$ and is defined recursively as follows. Let $Y_0 = c_1 e_0$ and $Y_1 = c_2 Y_0 + e_1$. Then let $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t$ for $t > 1$ as in an AR(2) process.

(a) Show that the process mean is zero.

$E(Y_0) = cE(e_0) = 0$ and $E(Y_1) = c_2 E(Y_1) + E(e_0) = 0$ Now proceed by induction. Suppose $E(Y_t) = E(Y_{t-1}) = 0$. Then $E(Y_{t+1}) = c_1 E(Y_t) + c_2 E(Y_{t-1}) + E(e_{t+1}) = 0$ and the result is established by induction on t.

(b) For particular values of ϕ_1 and ϕ_2 within the stationarity region for an AR(2) model, show how to choose c_1 and c_2 so that both $Var(Y_0) = Var(Y_1)$ and the lag 1 autocorrelation between Y_1 and Y_0 matches that of a stationary AR(2) process with parameters ϕ_1 and ϕ_2.

With no loss of generality, assume $\sigma_e^2 = 1$. Then $Var(Y_0) = c_1^2 = Var(Y_1) = c_2^2 c_1^2 + 1$ or $c_1 = 1/\sqrt{1 - c_2^2}$.

Next $Cov(Y_0, Y_1) = Cov(c_1 e_0, c_2 c_1 e_0 + e_1) = Cov(c_1 e_0, c_2 c_1 e_0) = c_2 c_1^2$. So $\rho_1 = (c_2 c_1^2)/(c_1^2) = c_2$.

Finally, given parameters ϕ_1 and ϕ_2 within the stationarity region for an AR(2) model, set $c_2 = \phi_1/(1 - \phi_2)$ and $c_1 = 1/\sqrt{1 - c_2^2}$ and all of the requirements are met.

(c) Once the process $\{Y_t\}$ is generated, show how to transform it to a new process that has any desired mean and variance. (This exercise suggests a convenient method for simulating stationary AR(2) processes.)

The process $\sqrt{\gamma_0} Y_t / c_1 + \mu$ will have the desired properties.

Exercise 4.25 Consider an "AR(1)" process satisfying $Y_t = \phi Y_{t-1} + e_t$ where ϕ can be ***any*** number and $\{e_t\}$ is a white noise process such that e_t is independent of the past $\{Y_{t-1}, Y_{t-2}, \ldots\}$. Let Y_0 be a random variable with mean μ_0 and variance σ_0^2.

(a) Show that for $t > 0$ we can write

$$Y_t = e_t + \phi e_{t-1} + \phi^2 e_{t-2} + \phi^3 e_{t-3} + \cdots + \phi^{t-1} e_1 + \phi^t Y_0.$$

$$\begin{aligned} Y_t &= \phi Y_{t-1} + e_t = \phi(\phi Y_{t-2} + e_{t-1}) + e_t = \phi^2(\phi Y_{t-3} + e_{t-2}) + \phi e_{t-1} + e_t = \cdots \\ &= \phi^t Y_0 + e_t + \phi e_{t-1} + \cdots + \phi^{t-1} e_1 \qquad \text{as required.} \end{aligned}$$

(b) Show that for $t > 0$ we have $E(Y_t) = \phi^t\mu_0$.

$$E(Y_t) = E(\phi^t Y_0 + e_t + \phi e_{t-1} + \cdots + \phi^{t-1}e_1) = \phi^t E(Y_0) = \phi^t\mu_0$$

(c) Show that for $t > 0$

$$Var(Y_t) = \begin{cases} \dfrac{1-\phi^{2t}}{1-\phi^2}\sigma_e^2 + \phi^{2t}\sigma_0^2 & \text{for } \phi \neq 1 \\ t\sigma_e^2 + \sigma_0^2 & \text{for } \phi = 1 \end{cases}$$

$Var(Y_t) = Var(\phi^t Y_0 + e_t + \phi e_{t-1} + \cdots + \phi^{t-1}e_1) = \phi^{2t}\sigma_0^2 + \sigma_e^2 \sum_{k=0}^{t-1}\phi^{2k}$ Then sum the finite geometric series.

(d) Suppose now that $\mu_0 = 0$. Argue that, if $\{Y_t\}$ is stationary, we must have $\phi \neq 1$.

If $\mu_0 = 0$, the process has a zero mean. If $\phi = 1$ and the process is stationary, then $Var(Y_t) = Var(Y_{t-1}) + \sigma_e^2$ which is impossible.

(e) Continuing to suppose that $\mu_0 = 0$, show that, if $\{Y_t\}$ is stationary, then $Var(Y_t) = \sigma_e^2/(1-\phi^2)$ and so we must have $|\phi| <1$.

$Var(Y_t) = \phi^2 Var(Y_t) + \sigma_e^2$ and solve for $Var(Y_t) = \sigma_e^2/(1-\phi^2)$. Since this variance must be positive we must have $|\phi| <1$.

CHAPTER 5

Exercise 5.1 Identify as specific ARIMA models, that is, what are p, d, and q and what are the values of the parameters—the ϕ's and θ's?

(a) $Y_t = Y_{t-1} - 0.25Y_{t-2} + e_t - 0.1e_{t-1}$

This looks like an ARMA(2,1) model with $\phi_1 = 1$ and $\phi_2 = -0.25$. We need to check the stationarity conditions of Equation (4.3.11), page 72. Here $\phi_1 + \phi_2 = 0.75 < 1$, $\phi_2 - \phi_1 = -1.25 < 1$, and $|\phi_2| = 0.25 < 1$ so the process is a stationary and invertible ARMA(2,1) with $\phi_1 = 1$, $\phi_2 = -0.25$, and $\theta_1 = 0.1$.

(b) $Y_t = 2Y_{t-1} - Y_{t-2} + e_t$

Initially it looks like an AR(2) model but $2 + (-1) = 1$ which is not strictly less than 1. Rewriting as $Y_t - Y_{t-1} = (Y_{t-1} - Y_{t-2}) + e_t$ suggests an AR(1) model in the differences $Y_t - Y_{t-1}$ but the AR coefficient would be equal to 1. Actually, the second difference $Y_t - 2Y_{t-1} + Y_{t-2} = e_t$ is white noise, so that $\{Y_t\}$is an IMA(2,0) model.

(c) $Y_t = 0.5Y_{t-1} - 0.5Y_{t-2} + e_t - 0.5e_{t-1} + 0.25e_{t-2}$

The AR part is stationary since the inequalities of Equation (4.3.11), page 72, are satisfied. Applying the same equations to the MA part of the model, we see that the MA part is invertible. Therefore, the model is a stationary, and invertible ARMA(2,2) model with $\phi_1 = 0.5$, $\phi_2 = -0.5$, $\theta_1 = 0.5$, and $\theta_2 = -0.25$.

Exercise 5.2 For each of the ARIMA models below, give the values for $E(\nabla Y_t)$ and $Var(\nabla Y_t)$.

(a) $Y_t = 3 + Y_{t-1} + e_t - 0.75e_{t-1}$

Here $\nabla Y_t = Y_t - Y_{t-1} = 3 + e_t - 0.75e_{t-1}$ so that $E(\nabla Y_t) = 3$ and $Var(\nabla Y_t) = [1 + (0.75)^2]\sigma_e^2 = \left[\frac{25}{16}\right]\sigma_e^2$.

(b) $Y_t = 10 + 1.25Y_{t-1} - 0.25Y_{t-2} + e_t - 0.1e_{t-1}$

In this case $\nabla Y_t = Y_t - Y_{t-1} = 10 + 0.25(Y_{t-1} - Y_{t-2}) + e_t - 0.1e_{t-1}$. So the model is a stationary, invertible, ARIMA(1,1,1) model with $\phi = 0.25$, $\theta = 0.1$, and $\theta_0 = 10$. Hence $E(\nabla Y_t) = \dfrac{\theta_0}{1-\phi} = \dfrac{10}{1-0.25} = \dfrac{10}{0.75} = \dfrac{40}{3}$.

Also, from Equation (4.4.4), page 78,

$$Var(\nabla Y_t) = \frac{(1-2\phi\theta+\theta^2)}{1-\phi^2}\sigma_e^2 = \frac{(1-2(0.25)(0.1)+(0.1)^2)}{1-(0.25)^2}\sigma_e^2 = 1.024\sigma_e^2$$

(c) $Y_t = 5 + 2Y_{t-1} - 1.7Y_{t-2} + 0.7Y_{t-3} + e_t - 0.5e_{t-1} + 0.25e_{t-2}$

Factoring the AR characteristic polynomial we have $1 - 2x + 1.7x^2 - 0.7x^3 = (1 - x)(1 - x + 0.7x^2)$. This shows that a first difference is needed after which a stationary AR(2) obtains. Thus the model may be rewritten as

$\nabla Y_t = 5 + \nabla Y_{t-1} - 0.7\nabla Y_{t-2} + e_t - 0.5e_{t-1} + 0.25e_{t-2}$. So the model is an ARIMA(2,1,2) with $\phi_1 = 1$, $\phi_2 = -0.75$, $\theta_1 = 0.5$, $\theta_2 = -0.25$, and $\theta_0 = 5$.

Exercise 5.3 Suppose that $\{Y_t\}$ is generated according to $Y_t = e_t + ce_{t-1} + ce_{t-2} + ce_{t-3} + \cdots + ce_0$ for $t > 0$.

(a) Find the mean and covariance functions for $\{Y_t\}$. Is $\{Y_t\}$ stationary?

$E(Y_t) = 0$ and $Var(Y_t) = Var(e_t + ce_{t-1} + ce_{t-2} + \cdots + ce_0) = (1 + tc^2)\sigma_e^2$ which, in general, varies with t.

Assume that $t < s$. Then

$$\begin{aligned} Cov(Y_t, Y_s) &= Cov(e_t + ce_{t-1} + ce_{t-2} + \cdots + ce_0, e_s + ce_{s-1} + \cdots + ce_t + ce_{t-1} + \cdots + ce_0) \\ &= (c + c^2 t)\sigma_e^2 = c(1 + ct)\sigma_e^2 \end{aligned}$$

(b) Find the mean and covariance functions for $\{\nabla Y_t\}$. Is $\{\nabla Y_t\}$ stationary?

$\nabla Y_t = (e_t + ce_{t-1} + ce_{t-2} + \cdots + ce_0) - (e_{t-1} + ce_{t-2} + ce_{t-3} + \cdots + ce_0) = e_t - (1 - c)e_{t-1}$ so this is stationary for any value of c. See part (c).

(c) Identify $\{Y_t\}$ as a specific ARIMA process.

The process $\{\nabla Y_t\}$ is an MA(1) process so that $\{Y_t\}$ is IMA(1,1) or ARIMA(0,1,1) with $\theta = 1 - c$. The $\{\nabla Y_t\}$ process is invertible if $|c| < 1$.

Exercise 5.4 Suppose that $Y_t = A + Bt + X_t$ where $\{X_t\}$ is a random walk. First suppose that A and B are constants.

(a) Is $\{Y_t\}$ stationary?

Since $E(Y_t) = A + Bt$, in general, varies with t, the process $\{Y_t\}$ is not stationary.

(b) Is $\{\nabla Y_t\}$ stationary? $\nabla Y_t = (A + Bt + X_t) - [A + B(t - 1) + X_{t-1}] = B + X_t - X_{t-1} = B + \nabla X_t$. So $E(\nabla Y_t) = B$.

$Cov(\nabla Y_t, \nabla Y_{t-k}) = Cov(B + \nabla X_t, B + \nabla X_{t-k}) = 0$ for $k > 0$ since ∇X_t is white noise and B is a constant.

Now suppose that A and B are random variables that are independent of the random walk $\{X_t\}$.

(c) Is $\{Y_t\}$ stationary?

No, since $E(Y_t) = E(A) + E(B)t$, in general, varies with t, the process $\{Y_t\}$ is not stationary.

(d) Is $\{\nabla Y_t\}$ stationary? We still have $\nabla Y_t = B + \nabla X_t$. So $E(\nabla Y_t) = E(B)$ which is constant in t.

$Cov(\nabla Y_t, \nabla Y_{t-k}) = Cov(B + \nabla X_t, B + \nabla X_{t-k}) = Var(B)$ for all k. So we do have stationarity.

Exercise 5.5 Using the simulated white noise values in Exhibit (5.2), page 88, verify the values shown for the explosive process Y_t.

Direct calculation.

Exercise 5.6 Consider a stationary process $\{Y_t\}$. Show that if $\rho_1 < ½$, ∇Y_t has a larger variance than does Y_t.

$Var(\nabla Y_t) = Var(Y_t - Y_{t-1}) = Var(Y_t) + Var(Y_{t-1}) - 2Cov(Y_t, Y_{t-1}) = 2(1 - \rho_1)Var(Y_t)$. So if $\rho_1 < ½$, $2(1 - \rho_1)$ is larger than 1 and the result follows.

Exercise 5.7 Consider two models:

A: $Y_t = 0.9Y_{t-1} + 0.09Y_{t-2} + e_t$

Since $\phi_1 + \phi_2 < 1$, $\phi_2 - \phi_1 < 1$, and $|\phi_2| < 1$, the process is a stationary AR(2) process. with $\phi_1 = 0.9$ and $\phi_2 = 0.09$.

B: $Y_t = Y_{t-1} + e_t - 0.1e_{t-1}$

Since $Y_t - Y_{t-1} = e_t - 0.1e_{t-1}$, this is an IMA(1,1) process with $\theta = 0.1$.

(a) Identify each as a specific ARIMA model. That is, what are p, d, and q and what are the values of the parameters, ϕ's and θ's?

(b) In what ways are the two models different?

One is stationary while the other is nonstationary.

(c) In what ways are the two models similar? (Compare ψ-weights and π-weights.)

Using Equations (4.3.21) on page 75 we can calculate the ψ-weights for the AR(2) model. This could be done with a calculator or the following R code:

```
> psi=NULL
> phi1=0.9 ;phi2=0.09; max.lag=20; psi[1]=1;psi[2]=phi1
> for (k in 3:max.lag) psi[k]=phi1*psi[k-1]+phi2*psi[k-2]
> psi # The indexing here is off by 1
```

```
 [1] 1.0000000 0.9000000 0.9000000 0.8910000 0.8829000 0.8748000 0.8667810
 [8] 0.8588349 0.8509617 0.8431607 0.8354312 0.8277725 0.8201841 0.8126652
[15] 0.8052152 0.7978336 0.7905196 0.7832726 0.7760921 0.7689775
```

Alternatively, you can use the ARMAtoMA function:

```
> ARMAtoMA(ar=c(phi1,phi2),lag.max=20)
```

```
 [1] 0.9000000 0.9000000 0.8910000 0.8829000 0.8748000 0.8667810 0.8588349
 [8] 0.8509617 0.8431607 0.8354312 0.8277725 0.8201841 0.8126652 0.805215
[15] 0.7978336 0.7905196 0.7832726 0.7760921 0.7689775 0.7619280
```

From Equation (5.2.6), page 93, the ψ-weights for the IMA(1,1) model are 1, $1 - 0.1 = 0.9$, $1 - 0.1 = 0.9$, $1 - 0.1 = 0.9$,… . So the ψ-weights for the two models are very similar for many lags.

The π-weights for the IMA(1,1) model are obtained from Equation (4.5.5), page 80. We find that $\pi_k = (1-\theta)\theta^{k-1}$ for $k = 1, 2,\ldots$. So $\pi_1 = (1 - 0.1) = 0.9$, $\pi_2 = (1 - 0.1)(0.1) = 0.09$, $\pi_3 = (1 - 0.1)(0.1)^2 = 0.009$, and so on. The first two π-weights for the two models are identical and the remaining π-weights are nearly the same. These two models would be essentially impossible to distinguish in practice.

Exercise 5.8 Consider a nonstationary "AR(1)" process defined as a solution to Equation (5.1.2), page 88, with $|\phi| > 1$.

(a) Derive an equation similar to Equation (5.1.3), page 88, for this more general case. Use $Y_0 = 0$ as an initial condition. We find that

$$\begin{aligned}
Y_t &= \phi Y_{t-1} + e_t = e_t + \phi(\phi Y_{t-2} + e_{t-1}) = e_t + \phi e_{t-1} + \phi^2 Y_{t-2} = \cdots \\
&= e_t + \phi e_{t-1} + \phi^2 e_{t-2} + \cdots + \phi^{t-1} e_1 + \phi^t Y_0 \\
&= e_t + \phi e_{t-1} + \phi^2 e_{t-2} + \cdots + \phi^{t-1} e_1
\end{aligned}$$

(b) Derive an equation similar to Equation (5.1.4), page 89, for this more general case.

$$Var(Y_t) = Var(e_t + \phi e_{t-1} + \phi^2 e_{t-2} + \cdots + \phi^{t-1} e_1) = \sigma_e^2 \sum_{k=0}^{t-1} (\phi^2)^k = \sigma_e^2 \left(\frac{\phi^{2t} - 1}{\phi^2 - 1} \right)$$

(c) Derive an equation similar to Equation (5.1.5), page 89, for this more general case.

$$\begin{aligned}
Cov(Y_t, Y_{t-k}) &= Cov(e_t + \phi e_{t-1} + \cdots + \phi^k e_{t-k} + \phi^{k+1} e_{t-k-1} + \cdots + \phi^{t-1} e_1, e_{t-k} + \phi e_{t-k-1} + \cdots + \phi^{t-k-1} e_1) \\
&= (\phi^k + \phi^{k+2} + \cdots + \phi^{2t-k-2})\sigma_e^2 \\
&= \phi^k (1 + \phi^2 + \phi^4 + \cdots + \phi^{2(t-k-1)})\sigma_e^2 \\
&= \phi^k \left(\frac{\phi^{2(t-k)} - 1}{\phi^2 - 1} \right) \sigma_e^2
\end{aligned}$$

(d) Is it true that for any $|\phi| > 1$, $Corr(Y_t, Y_{t-k}) \approx 1$ for large t and moderate k?

$$\text{Yes! } Corr(Y_t, Y_{t-k}) = \frac{\phi^k\left(\frac{\phi^{2(t-k)}-1}{\phi^2-1}\right)\sigma_e^2}{\sqrt{\left(\left(\frac{\phi^{2t}-1}{\phi^2-1}\right)\sigma_e^2\right)\left(\left(\frac{\phi^{2(t-k)}-1}{\phi^2-1}\right)\sigma_e^2\right)}} = \phi^k\sqrt{\frac{\phi^{2(t-k)}-1}{\phi^{2t}-1}} = \sqrt{\frac{1-\phi^{2(k-t)}}{1-\phi^{-2t}}} \approx 1$$

for large t and moderate k since $|\phi| > 1$.

Exercise 5.9 Verify Equation (5.1.10), page 90.

$\nabla Y_t = \varepsilon_t + e_t - e_{t-1}$ where $\{e_t\}$ and $\{\varepsilon_t\}$ are independent white noise series. So

$$Var(\nabla Y_t) = Var(\varepsilon_t + e_t - e_{t-1}) = \sigma_\varepsilon^2 + 2\sigma_e^2$$

$$Cov(\nabla Y_t, \nabla Y_{t-1}) = Cov(\varepsilon_t + e_t - e_{t-1}, \varepsilon_{t-1} + e_{t-1} - e_{t-2}) = -\sigma_e^2$$

$$Corr(\nabla Y_t, \nabla Y_{t-1}) = \frac{-\sigma_e^2}{\sigma_\varepsilon^2 + 2\sigma_e^2} = -\frac{1}{[2 + \sigma_\varepsilon^2/\sigma_e^2]}$$

Exercise 5.10 Nonstationary ARIMA series can be simulated by first simulating the corresponding stationary ARMA series and then “integrating” it (really partial summing it). Use statistical software to simulate a variety of IMA(1,1) and IMA(2,2) series with a variety of parameter values. Note any stochastic “trends” in the simulated series.

R code can do it all. Remember that R uses the negative of our θ values.

```
> plot(arima.sim(model=list(order=c(0,1,1),ma=-0.5),n=200),type='o',ylab='Series')
> # Do this several times with the same parameters to see the possible variation then
> # change the various parameters, ma and n.
```

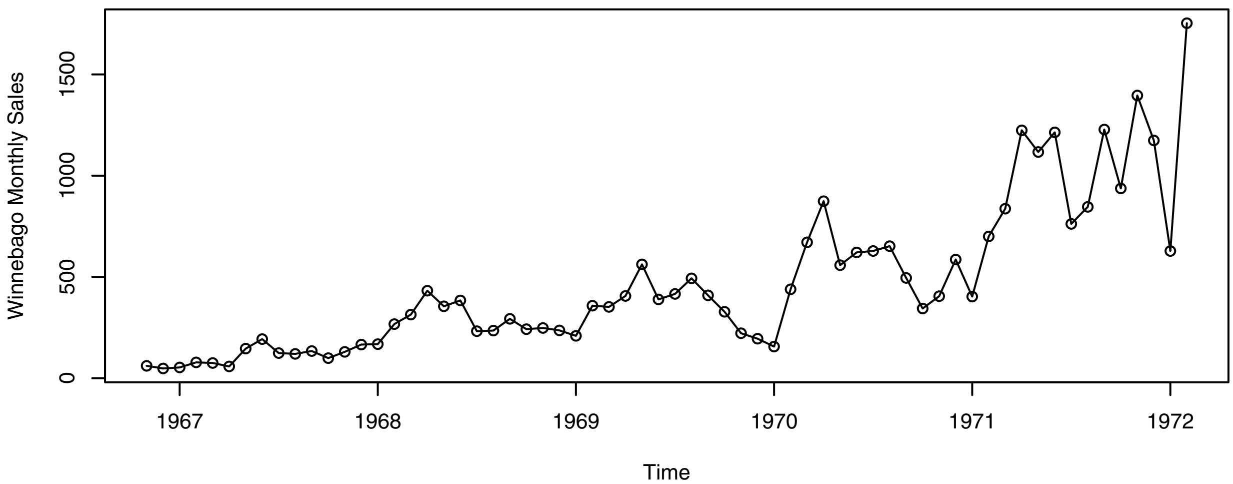
For the IMA(2,2) use, for example,

```
> plot(arima.sim(model=list(order=c(0,2,2),ma=c(-0.7,-0.1)),n=50),type='o',
  ylab='Series')
```

Again, it is instructive to repeat the simulation several times with the same parameters and several times with different parameters.

Exercise 5.11 The data file winnebago contains monthly unit sales of recreational vehicles (RVs) from Winnebago, Inc. from November 1966 through February 1972.

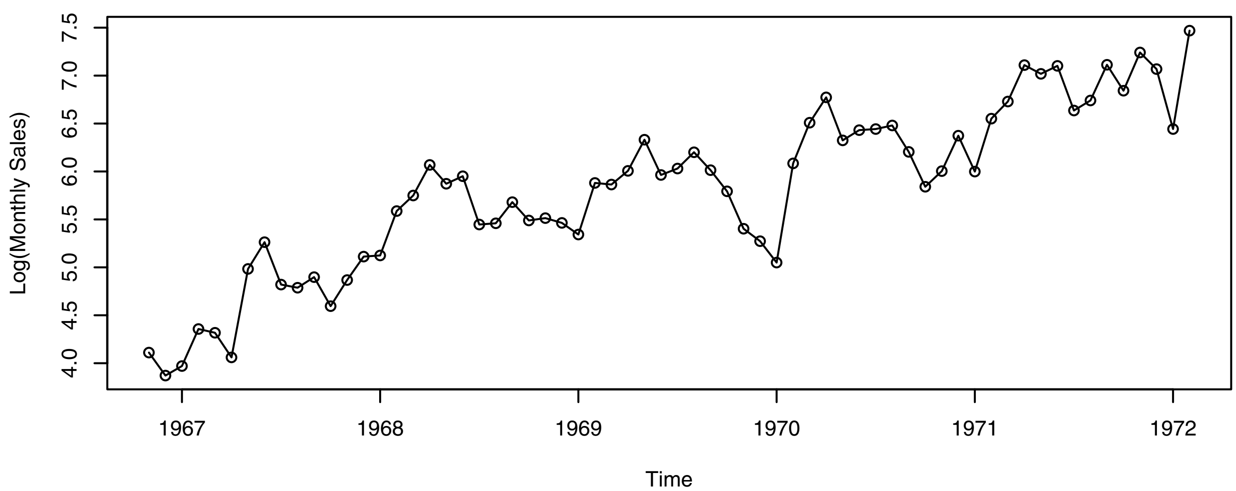
(a) Display and interpret the time series plot for these data.




```
> data(winnebago); win.graph(width=6.5,height=3,pointsize=8)
> plot(winnebago,type='o',ylab='Winnebago Monthly Sales')
```

The series increases over time and the variation is larger as the series level gets higher—a series begging us to take logarithms.

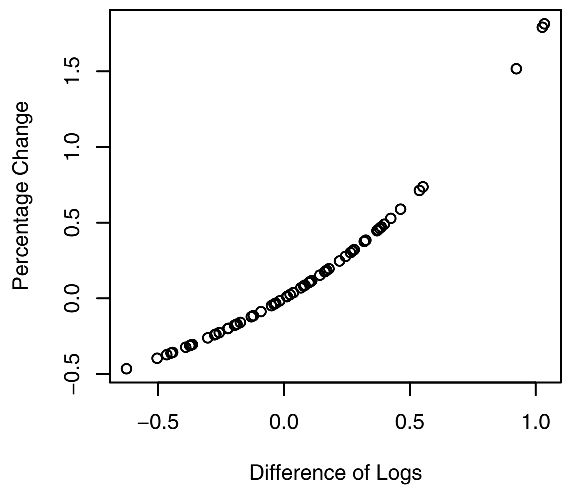
(b) Now take natural logarithms of the monthly sales figures and display the time series plot of the transformed values. Describe the effect of the logarithms on the behavior of the series.




```
> plot(log(winnebago),type='o',ylab='Log(Monthly Sales)')
```

The series still increases over time, but now the variation around the general level is quite similar at all levels of the series.

(c) Calculate the fractional relative changes, $(Y_t - Y_{t-1})/Y_{t-1}$, and compare them to the differences of (natural) logarithms, $\nabla\log(Y_t) = \log(Y_t) - \log(Y_{t-1})$. How do they compare for smaller values and for larger values?

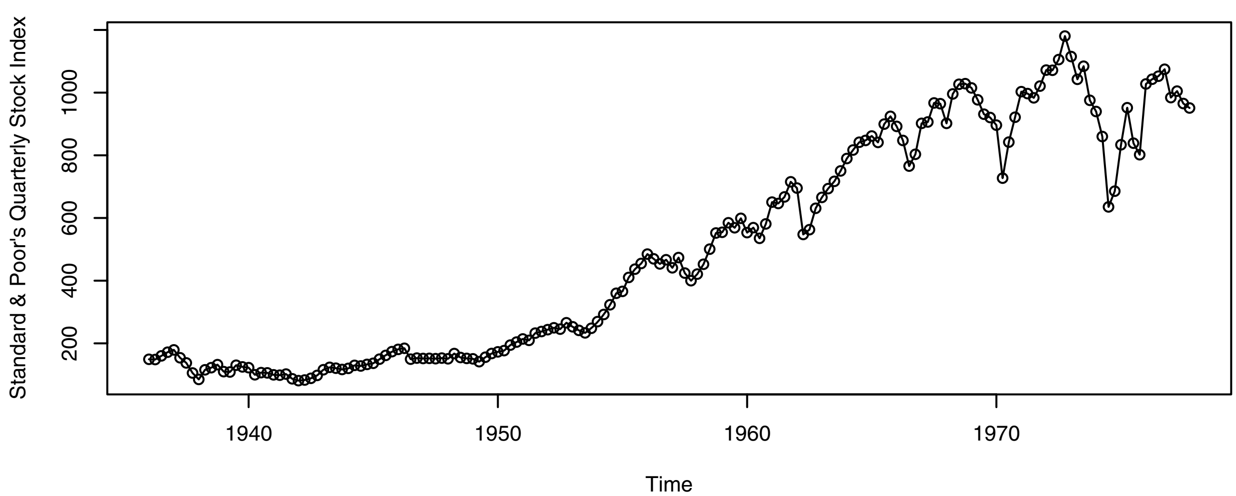



```
> percentage=na.omit((winnebago-zlag(winnebago))/zlag(winnebago))
> win.graph(width=3,height=3,pointsize=8)
> plot(x=diff(log(winnebago))[-1],y=percentage[-1], ylab='Percentage Change',
   xlab='Difference of Logs')
> cor(diff(log(winnebago))[-1],percentage[-1])
```

If there were a perfect relationship, the above plot would be a straight line. Clearly, the relationship is good but not perfect. The correlation coefficient in this plot is 0.96 so the agreement is quite good. Of course, there is seasonality in this series that has not been modeled.

Exercise 5.12 The data file SP contains quarterly Standard & Poor's Composite Index of stock price values from the first quarter of 1936 through the fourth quarter of 1977.

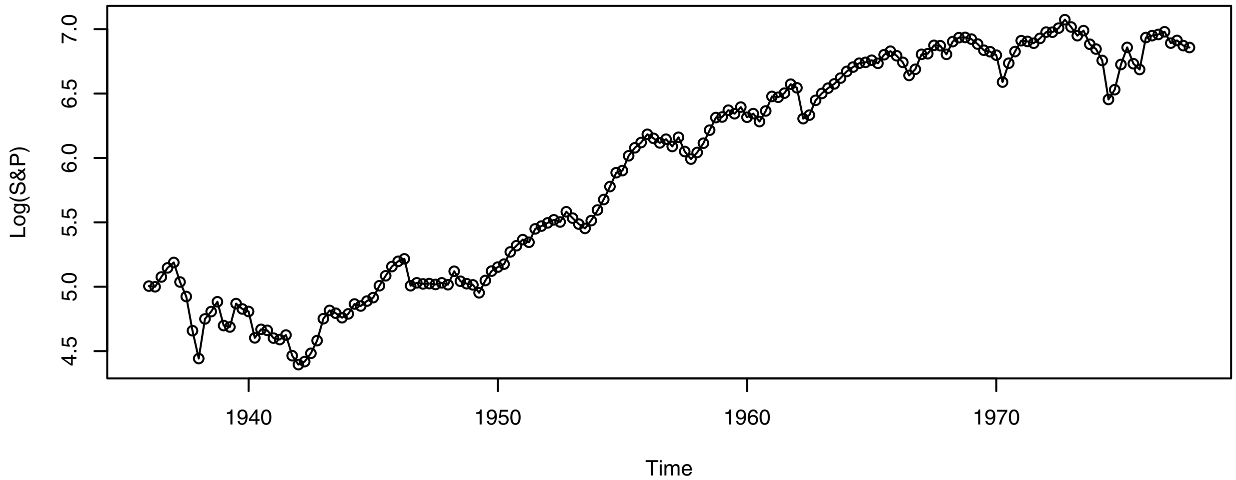
(a) Display and interpret the time series plot for these data.




```
> data(SP); plot(SP,type='o',ylab='Standard & Poor\'s Quarterly Stock Index')
```

Another general upward "trend" but with increased variation at the higher levels.

(b) Now take natural logarithms of the quarterly values and display and the time series plot of the transformed values. Describe the effect of the logarithms on the behavior of the series.




```
> plot(log(SP),type='o',ylab='Log(S&P)')
```

Now the variance is stabilized but, of course, the upward trend is still there.

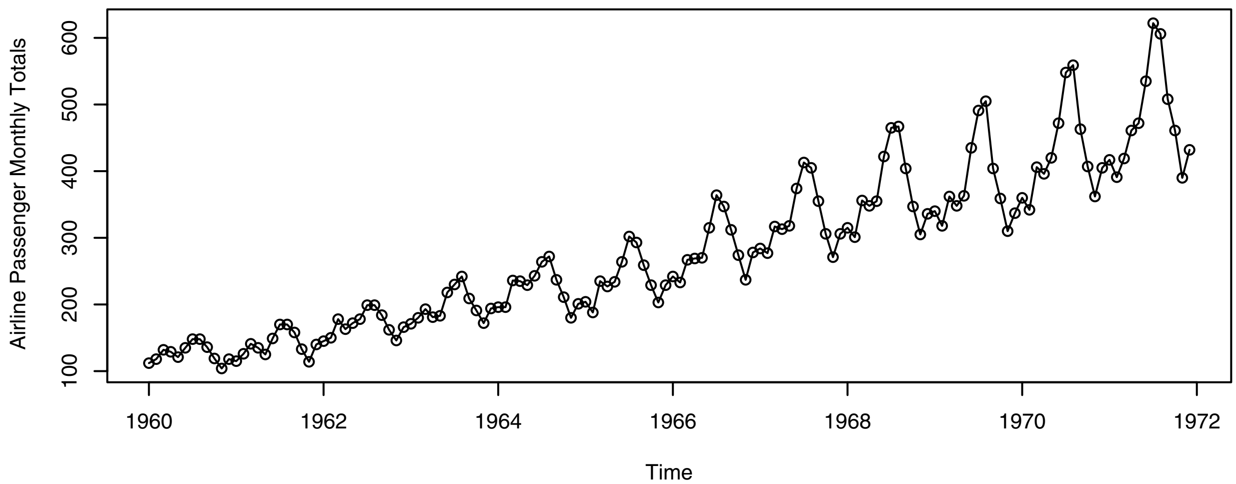
(c) Calculate the (fractional) relative changes, $(Y_t - Y_{t-1})/Y_{t-1}$, and compare them to the differences of (natural) logarithms, $\nabla\log(Y_t)$. How do they compare for smaller values and for larger values?


```
> percentage=na.omit((SP-zlag(SP))/zlag(SP))
> win.graph(width=3,height=3,pointsize=8)
> plot(x=diff(log(SP))[-1],y=percentage[-1], ylab='Percentage Change',
   xlab='Difference of Logs')
> cor(diff(log(SP))[-1],percentage[-1])
```

Here the agreement between the two is very good and the correlation coefficient is 0.996.

Exercise 5.13 The data file airpass contains international airline passenger monthly totals (in thousands) from January 1949 through the December of 1960. This is a classic time series analyzed in Box and Jenkins (1976).

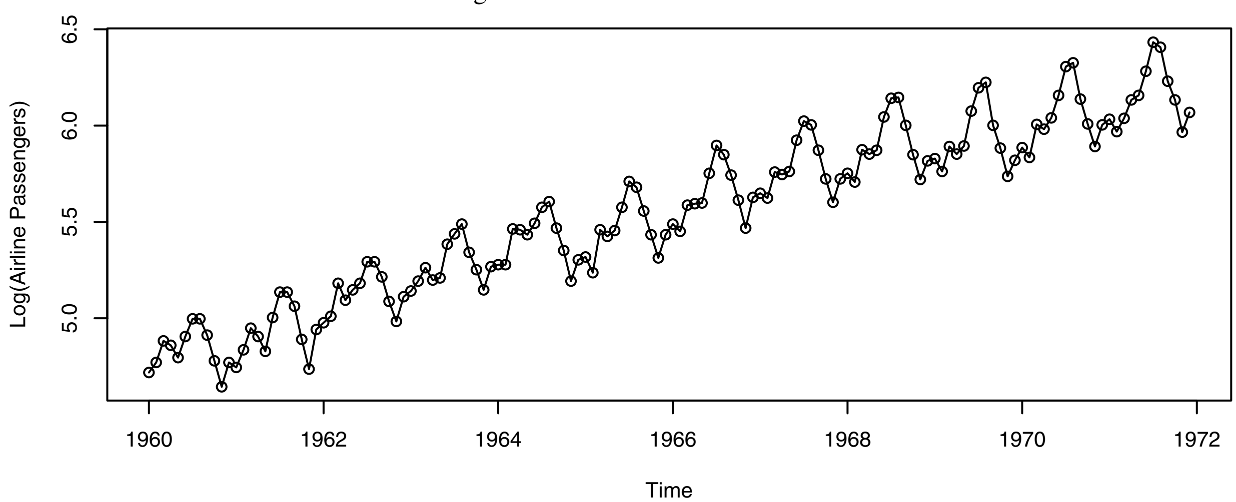
(a) Display and interpret the time series plot for these data.




```
> win.graph(width=6.5,height=3,pointsize=8)
> data(airpass); plot(airpass,type='o',ylab='Airline Passenger Monthly Totals')
```

There is a general upward "trend" with increased variation at the higher levels. There is also evidence of seasonality.

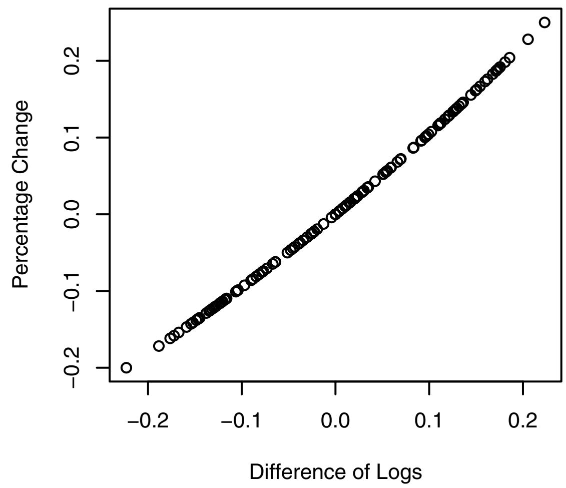
(b) Now take natural logarithms of the monthly values and display and the time series plot of the transformed values. Describe the effect of the logarithms on the behavior of the series




```
> plot(log(airpass),type='o',ylab='Log(Airline Passengers)')
```

Now the variation is similar at both high, low, and middle levels of the series.

(c) Calculate the (fractional) relative changes, $(Y_t - Y_{t-1})/Y_{t-1}$, and compare them to the differences of (natural) logarithms, $\nabla\log(Y_t)$. How do they compare for smaller values and for larger values?

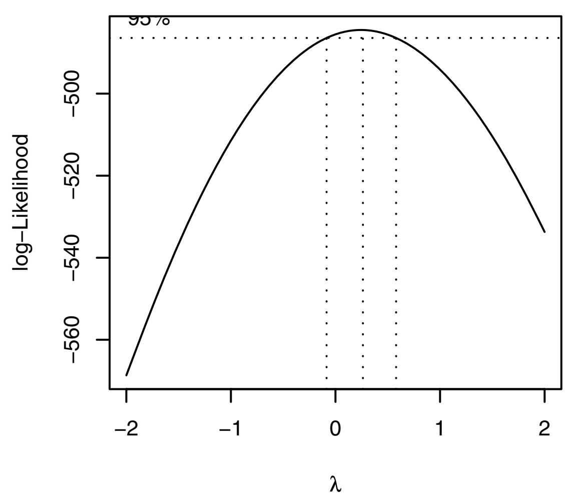



```
> percentage=na.omit((airpass-zlag(airpass))/zlag(airpass))
> win.graph(width=3,height=3,pointsize=8)
> plot(x=diff(log(airpass))[-1],y=percentage[-1],
   ylab='Percentage Change',xlab='Difference of Logs')
> cor(diff(log(airpass))[-1],percentage[-1])
```

The is excellent agreement between the two transformed series in this case. The correlation coefficient in this plot is 0.999. Either transformation would be extremely helpful in modeling this series further.

Exercise 5.14 Consider the annual rainfall data for Los Angeles shown in Exhibit (1.1), page 2. The quantile-quantile normal plot of these data, shown in Exhibit (3.17), page 50, convinced us that the data were not normal. The data are in the file larain.

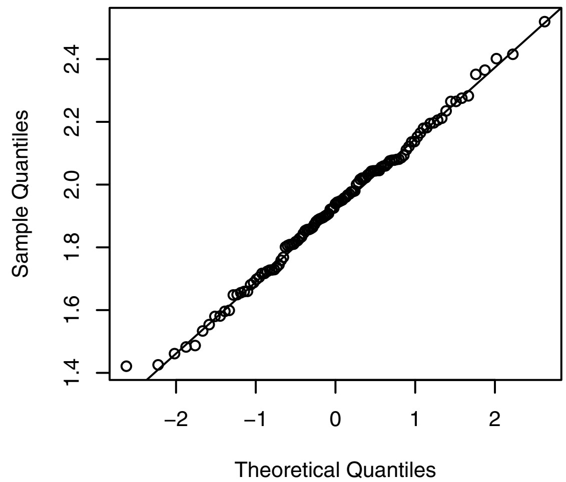
(a) Use software to produce a plot similar to Exhibit (5.11), page 102, and determine the "best" value of λ for a power transformation of the data.




```
> win.graph(width=3,height=3,pointsize=8)
> data(larain); BoxCox.ar(larain, method='ols')
```

The maximum likelihood value for lambda is about 0.26 but the 95% confidence interval includes the logarithm transformation (lambda = 0) and square root transformation (lambda = 0.5). We choose lambda = 0.25 or fourth root for the remaining sections of this exercise.

(b) Display a quantile-quantile plot of the transformed data. Are they more normal?

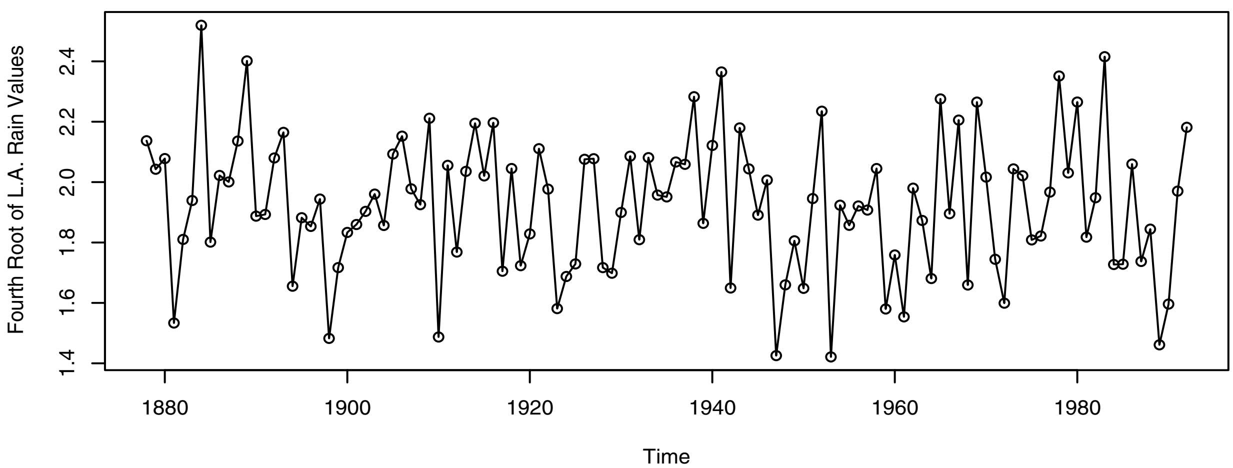



```
> win.graph(width=3,height=3,pointsize=8)
> qqnorm((larain)^.25,main='')
> qqline((larain)^.25)
> shapiro.test((larain)^.25)
```

```
        Shapiro-Wilk normality test
        data:  (larain)^0.25
        W = 0.9941, p-value = 0.9096
```

The values transformed by the fourth root look quite normal.

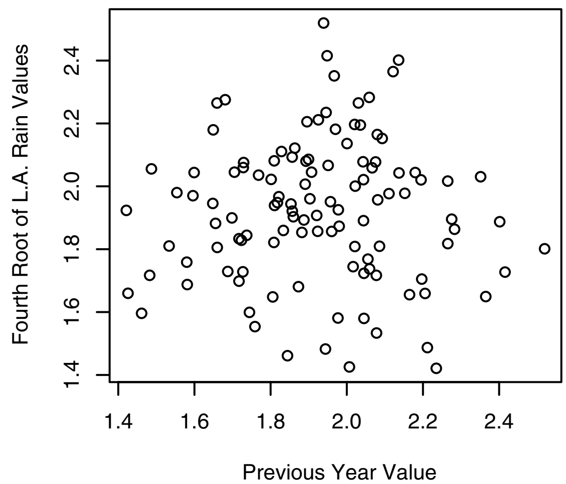
(c) Produce a time series plot of the transformed values.




```
> win.graph(width=6.5,height=3,pointsize=8)
> plot(larain^.25,type='o',ylab='Fourth Root of L.A. Rain Values')
```

This transformed series could now be considered as normal white noise with a nonzero mean.

(d) Use the transformed values to display a plot of Y_t versus Y_{t-1} as in Exhibit (1.2), page 2. Should we expect the transformation to change the dependence or lack of dependence in the series?

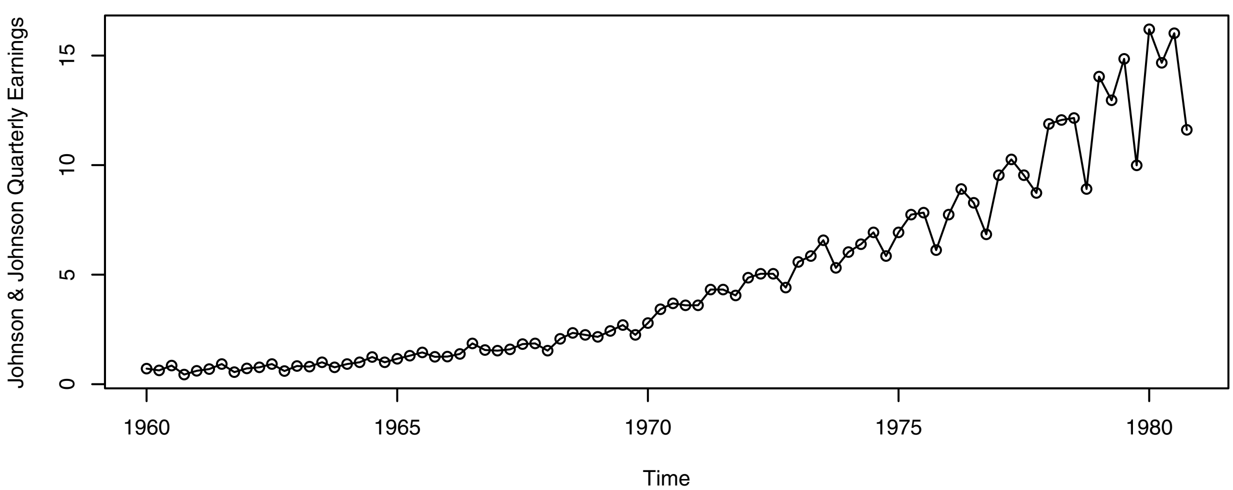



```
> win.graph(width=3,height=3,pointsize=8)
> plot(y=(larain)^.25,x=zlag((larain)^.25),
   ylab='Fourth Root of L.A. Rain Values',xlab='Previous Year Value')
```

The lack of correlation or any other kind of dependency between year values is clear from this plot. Instantaneous transformations cannot induce correlation where none was present.

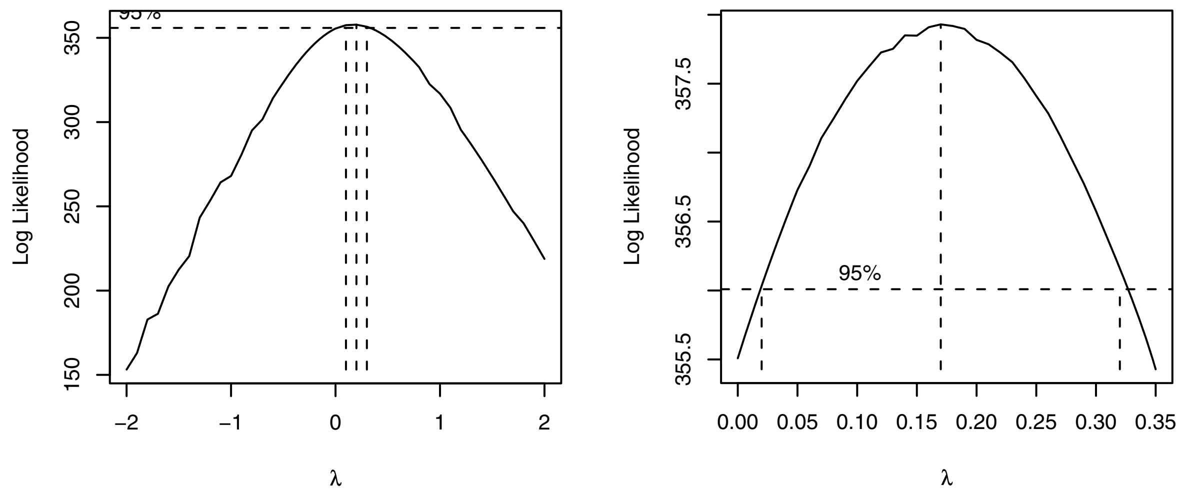
Exercise 5.15 Quarterly earnings per share for the Johnson & Johnson company are given in the data file named JJ. The data cover the years from 1960 through 1980.

(a) Display a time series plot of the data. Interpret the interesting features in the plot.




```
> win.graph(width=6.5,height=3,pointsize=8)
> data(JJ); plot(JJ,type='o',ylab='Johnson & Johnson Quarterly Earnings')
```

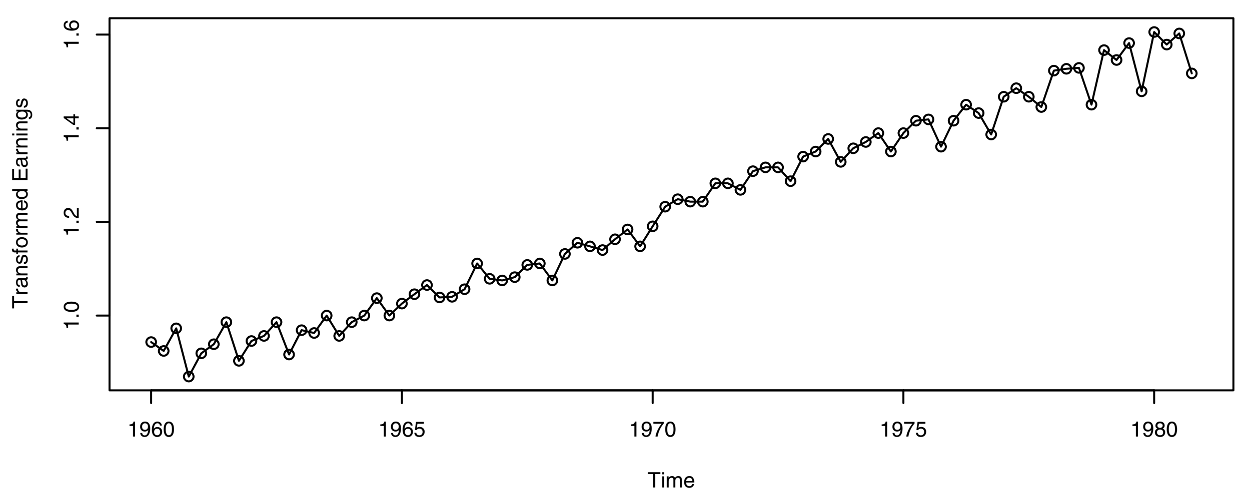
(a) Use software to produce a plot similar to Exhibit (5.11), page 102, and determine the “best” value of λ for a power transformation of these data..




```
> win.graph(width=3,height=3,pointsize=8)
> data(JJ); BC=BoxCox.ar(JJ)
> BC=BoxCox.ar(JJ,lambda=seq(0.0,0.35,0.01)) # New range for lambda to see more detail
```

The plot on the left shows the initial default Box-Cox analysis. The plot on the right shows more detail as the range for the lambda parameter has been restricted to 0.0 to 0.35. The maximum likelihood estimate of lambda is 0.17 and the 95% confidence interval runs from 0.02 to 0.32. We use 0.17 as the lambda parameter in the remaining analysis.

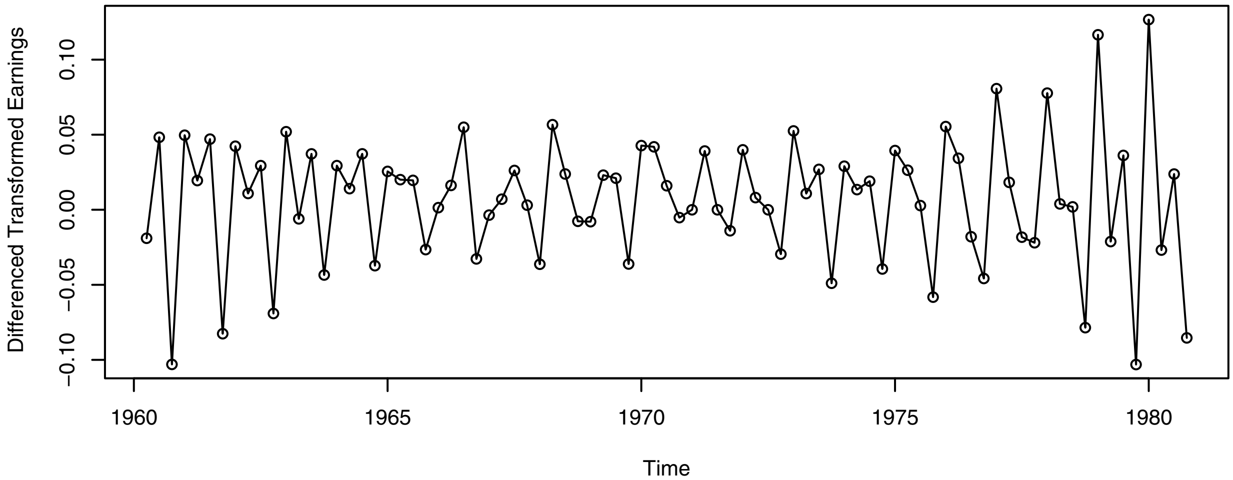
(b) Display a time series plot of the transformed values. Does this plot suggest that a stationary model might be appropriate?




```
> win.graph(width=6.5,height=3,pointsize=8)
> plot((JJ)^0.17,type='o',ylab='Transformed Earnings')
```

The variance has been stabilized but the strong trend must be accounted for before we can entertain a stationary model.

(c) Display a time series plot of the differences of the transformed values. Does this plot suggest that a stationary model might be appropriate for the differences?

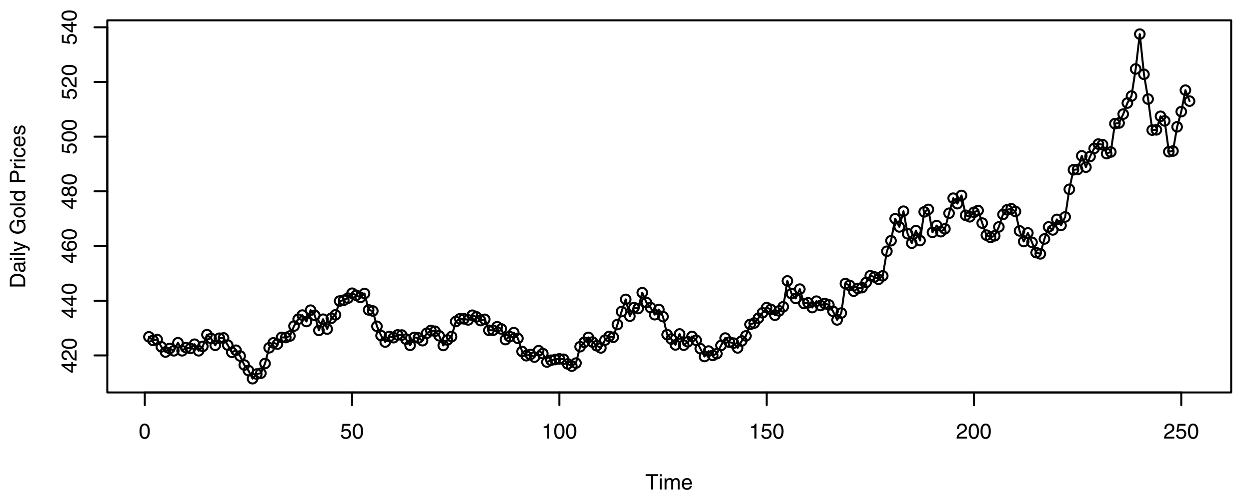



```
> plot(diff((JJ)^0.17),type='o',ylab='Differenced Transformed Earnings')
```

The trend is now gone but the variation does not appear to be constant across time and there may be quarterly seasonality to deal with.

Exercise 5.16 The file named gold contains the daily price of gold (in dollars per troy ounce) for the 252 trading days of year 2005.

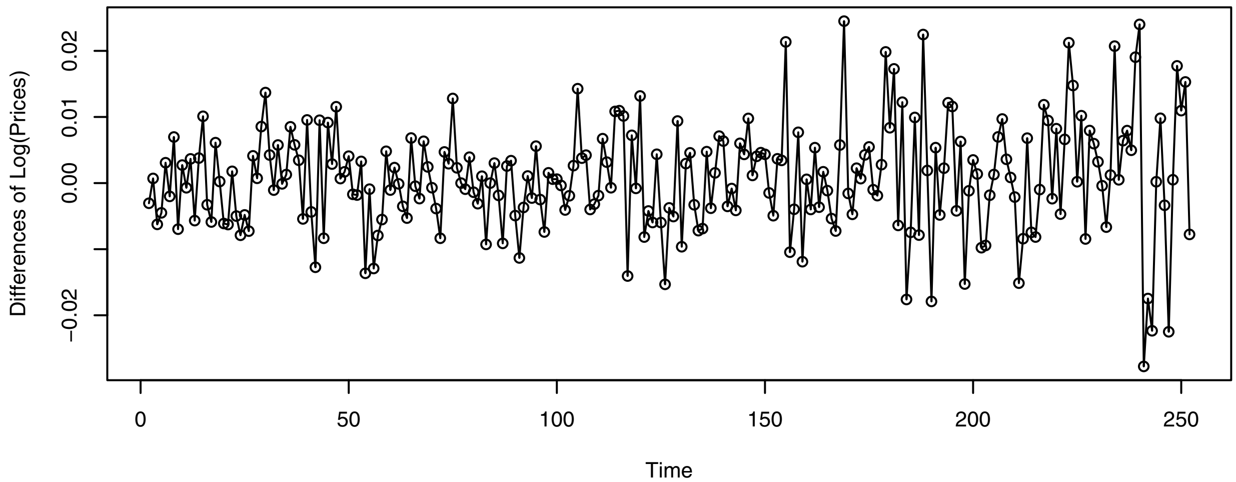
(a) Display the time series plot of these data. Interpret the plot.




```
> data(gold); plot(gold,type='o',ylab='Daily Gold Prices')
```

After a period of generally flat prices, the last half of the year shows substantial increases. (Look up today's gold prices at www.lbma.org.uk to see how the series is doing more recently.)

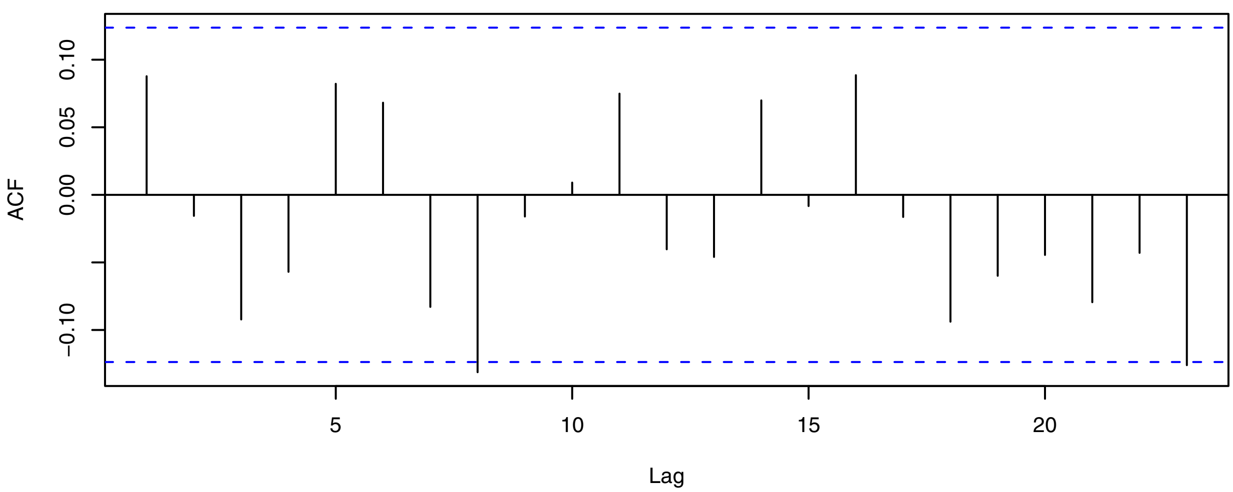
(b) Display the time series plot of the differences of the logarithms of these data. Interpret this plot.




```
> plot(diff(log(gold)),type='o',ylab='Differences of Log(Prices)')
```

The “trend” has been accounted for but there may be increased variability in the last half of the series.

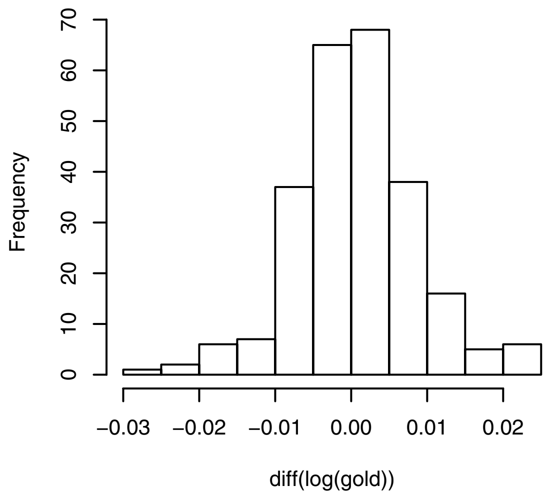
(c) Calculate and display the sample ACF for the differences of the logarithms of these data and argue that the logarithms appear to follow a random walk model.




```
> acf(diff(log(gold)),main='')
```

The differences of the logarithms of gold prices display the autocorrelation structure of while noise. Therefore, the logarithms of gold prices could be consided as a random walk.

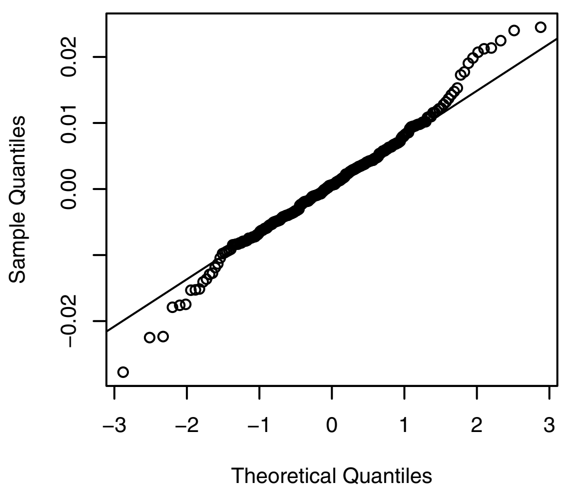
(d) Display the differences of logs in a histogram and interpret.




```
> win.graph(width=3,height=3,pointsize=8)
> hist(diff(log(gold)))
```

The distribution looks reasonably “bell-shaped,” but see part (e) below.

(e) Display the differences of logs in a quantile-quantile normal plot and interpret.




```
> qqnorm(diff(log(gold))); qqline(diff(log(gold)))
> shapiro.test(diff(log(gold)))
```

```
Shapiro-Wilk normality test
data:  diff(log(gold))
W = 0.9861, p-value = 0.01519
```

The Q-Q plot indicates that the distribution deviates from normality. In particular, the tails are lighter than a normal distribution. The Shapiro-Wilk test confirms this with a p-value of 0.015.

Exercise 5.17 Use calculus to show that, for any fixed $x > 0$, as $\lambda \to 0$, $(x^{\lambda} - 1)/\lambda \to \log x$.

First rewrite $x^{\lambda} = e^{\lambda \log(x)}$. Then $\dfrac{dx^{\lambda}}{d\lambda} = \dfrac{de^{\lambda \log(x)}}{d\lambda} = \log(x)e^{\lambda}$. So by l’Hospital’s rule

$$\lim_{\lambda \to 0} \frac{(x^\lambda - 1)}{\lambda} = \lim_{\lambda \to 0} \left[\frac{\frac{d}{d\lambda}(x^\lambda - 1)}{\frac{d\lambda}{d\lambda}} \right] = \lim_{\lambda \to 0} \left[\frac{\log(x)e^\lambda}{1} \right] = \log(x)$$

Chapter 6

Exercise 6.1 Verify Equation (6.1.3), page 110 for the white noise process.

For a white noise process $c_{ij} = \sum_{k=-\infty}^{\infty} (\rho_{k+i}\rho_{k+j} + \rho_{k-i}\rho_{k+j} - 2\rho_i\rho_k\rho_{k+j} - 2\rho_j\rho_k\rho_{k+i} + 2\rho_i\rho_j\rho_k^2)$ reduces to $c_{ii} = \rho_0^2 = 1$ and $c_{ij} = 0$ for $i \neq j$ and hence the result.

Exercise 6.2 Verify Equation (6.1.4), page 110 for the AR(1) process.

Without loss of generality, let $1 \le j$. Then

$$\begin{aligned} c_{jj} &= \sum_{k=-\infty}^{\infty} (\rho_{k+j}\rho_{k+j} + \rho_{k-j}\rho_{k+j} - 2\rho_j\rho_k\rho_{k+j} - 2\rho_j\rho_k\rho_{k+j} + 2\rho_j\rho_j\rho_k^2) \\ &= \sum_{k=-\infty}^{\infty} (\rho_{k+j}^2 + \rho_{k-j}\rho_{k+j} - 4\rho_j\rho_k\rho_{k+j} + 2\rho_j^2\rho_k^2) \\ &= (1 + 2\rho_j^2) \sum_{k=-\infty}^{\infty} \rho_k^2 + \sum_{k=-\infty}^{\infty} \rho_{k-j}\rho_{k+j} - 4\rho_j \sum_{k=-\infty}^{\infty} \rho_k\rho_{k+j} \end{aligned}$$

Now, with $\rho_k = \phi^k$, for $0 \le k$, we deal with the three sums separately.

$$\sum_{k=-\infty}^{\infty} \rho_k^2 = 1 + 2\sum_{k=1}^{\infty} \phi^{2k} = 1 + 2\frac{\phi^2}{1-\phi^2} = \frac{1+\phi^2}{1-\phi^2}.$$

Next

$$\begin{aligned} \sum_{k=-\infty}^{\infty} \rho_{k-j}\rho_{k+j} &= \sum_{k=-\infty}^{-j-1} \phi^{j-k-k-j} + \sum_{k=-j}^{j-1} \phi^{j-k+k+j} + \sum_{k=j}^{\infty} \phi^{k-j+k+j} \\ &= \sum_{k=-\infty}^{-j-1} \phi^{-2k} + \sum_{k=-j}^{j-1} \phi^{2i} + \sum_{k=j}^{\infty} \phi^{2k} \\ &= \sum_{k=-\infty}^{-j-1} \phi^{-2k} + \sum_{k=-j}^{j-1} \phi^{2j} + \sum_{k=j}^{\infty} \phi^{2k} \\ &= \frac{\phi^{2(j+1)}}{1-\phi^2} + 2j\phi^{2j} + \frac{\phi^{2j}}{1-\phi^2} \\ &= \phi^{2j}\left(\frac{1+\phi^2}{1-\phi^2}\right) + 2j\phi^{2j} \end{aligned}$$

For the third sum

$$\begin{aligned} \sum_{k=-\infty}^{\infty} \rho_k\rho_{k+j} &= \sum_{k=-\infty}^{-j} \phi^{-k}\phi^{-k-j} + \sum_{k=-j+1}^{0} \phi^{-k}\phi^{k+j} + \sum_{k=1}^{\infty} \phi^k\phi^{k+j} \\ &= \phi^{-j}\sum_{k=j}^{\infty} \phi^{2k} + j\phi^j + \phi^j\sum_{k=1}^{\infty} \phi^{2k} \\ &= \frac{\phi^j}{1-\phi^2} + j\phi^j + \phi^j\left(\frac{\phi^2}{1-\phi^2}\right) \\ &= \left(\frac{1+\phi^2}{1-\phi^2}\right)\phi^j + j\phi^j \end{aligned}$$

Substituting these three results into $c_{jj} = (1+2\rho_j^2)\sum_{k=-\infty}^{\infty}\rho_k^2 + \sum_{k=-\infty}^{\infty}\rho_{k-j}\rho_{k+j} - 4\rho_j\sum_{k=-\infty}^{\infty}\rho_k\rho_{k+j}$ and doing the final simplfication gives the desired result.

Exercise 6.3 Verify the line in Exhibit (6.1), page 111, for the values $\phi = \pm 0.9$.

Use $Var(r_1) \approx \frac{1-\phi^2}{n}$, $Var(r_k) \approx \frac{1}{n}\left[\frac{(1+\phi^2)(1-\phi^{2k})}{1-\phi^2} - 2k\phi^{2k}\right]$, and $Corr(r_1, r_2) \approx 2\phi\sqrt{\frac{1-\phi^2}{1+2\phi^2-3\phi^4}}$ and the R code

```
> phi=0.9;k=10;c11=sqrt((1-phi^2));c22=sqrt((1+phi^2)^2-4*phi^4)
> c12=2*phi*sqrt((1-phi^2)/(1+2*phi^2-3*phi^4))
> ckk=sqrt((1+phi^2)*(1-phi^(2*k))/(1-phi^2)-2*k*phi^(2*k))
> c11;c22;c12;ckk
```

The R code with $\phi = \pm 0.9$ gives 0.4358899, 0.8072794, ±0.9719086, and 2.436515 as rounded in the Exhibit.

Exercise 6.4 Add new entries to Exhibit (6.1), page 111, for the following values:

(a) $\phi = \pm 0.99$.

The R code shown in the previous exercise with $\phi = \pm 0.99$ gives 0.1410674, 0.2800214, ±0.9974716, and 1.326868.

(b) $\phi = \pm 0.5$.

Results for $\phi = \pm 0.5$ are 0.8660254, 1.145644, ±0.755929, and 1.290986.

(c) $\phi = \pm 0.1$.

Results for $\phi = \pm 0.1$ are 0.9949874, 1.009802, ±0.1970659, and 1.010051. These values are quite close to what we would get with a white noise process.

Exercise 6.5 Verify Equation (6.1.9), page 111 and Equation (6.1.10) for the MA(1) process.

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

$$
\begin{aligned}
c_{11} &= \sum_{k=-\infty}^{\infty}(\rho_{k+1}\rho_{k+1} + \rho_{k-1}\rho_{k+1} - 2\rho_1\rho_k\rho_{k+1} - 2\rho_1\rho_k\rho_{k+1} + 2\rho_1\rho_1\rho_k^2) \\
&= [(\rho_{-1}^2 + \rho_0^2 + \rho_1^2) + \rho_{-1}\rho_1 - 4(\rho_{-1}\rho_0\rho_1 + \rho_1\rho_0\rho_1) + 2\rho_1^2(\rho_{-1}^2 + \rho_0^2 + \rho_1^2)] \\
&= 1 - 3\rho_1^2 + 4\rho_1^4
\end{aligned}
$$

Also, for $j > 1$,

$$
\begin{aligned}
c_{jj} &= \sum_{k=-\infty}^{\infty}(\rho_{k+j}\rho_{k+j} + \rho_{k-j}\rho_{k+j} - 2\rho_j\rho_k\rho_{k+j} - 2\rho_j\rho_k\rho_{k+j} + 2\rho_j\rho_j\rho_k^2) \\
&= (\rho_{-1}^2 + \rho_0^2 + \rho_1^2) = 1 + 2\rho_1^2
\end{aligned}
$$

Finally,

$$
\begin{aligned}
c_{12} &= \sum_{k=-\infty}^{\infty}(\rho_{k+1}\rho_{k+2} + \rho_{k-1}\rho_{k+2} - 2\rho_1\rho_k\rho_{k+2} - 2\rho_2\rho_k\rho_{k+1} + 2\rho_1\rho_2\rho_k^2) \\
&= [(\rho_{-2+1}\rho_{-2+2} + \rho_{-1+1}\rho_{-1+2}) - 2\rho_1\rho_{-1}\rho_{-1+2}] \\
&= 2\rho_1(1-\rho_1^2)
\end{aligned}
$$

Exercise 6.6 Verify the line in Exhibit (6.2), page 112, for the values $\theta = \pm 0.9$.

```
> theta=0.9; rho1=-theta/(1+theta^2)
> c11=sqrt(1-3*rho1^2+4*rho1^4); c22=sqrt(1+2*rho1^2)
```

```
> c12=2*rho1*(1-rho1^2)/(c11*c22); c11; c22; c12
```

Here are the results from the R code: 0.7090734, 1.222494, and –0.8635941.

Exercise 6.7 Add new entries to Exhibit (6.2), page 112, for the following values:

(a) $\theta = \pm 0.99$.

Use the R code shown in the previous exercise with $\theta = \pm 0.99$ to get: 0.7071246, 1.224724, and ∓ 0.8660035.

(b) $\theta = \pm 0.8$.

Results: 0.7159797, 1.214869, and ∓ 0.8547268.

(c) $\theta = \pm 0.2$.

Results: 0.9457928, 1.036323, and ∓ 0.377894.

Exercise 6.8 Verify Equation (6.1.11), page 112, for the general MA(q) process.

In general, $c_{jj} = \sum_{k=-\infty}^{\infty} (\rho_{k+j}^2 + \rho_{k-j}\rho_{k+j} - 2\rho_j\rho_k\rho_{k+j} - 2\rho_j\rho_k\rho_{k+j} + 2\rho_j^2\rho_k^2)$

In the MA(q) case, only $\rho_{-q}, \rho_{-q+1}, \ldots, \rho_0, \rho_1, \ldots, \rho_q$ are non-zero. So $c_{jj} = 1 + 2\sum_{k=1}^{q}\rho_k^2$ as required.

Exercise 6.9 Use Equation (6.2.3), page 113, to verify the value for the lag 2 partial autocorrelation function for the MA(1) process given in Equation (6.2.5), page 114.

In general, $\phi_{22} = \dfrac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$. So for an MA(1) process

$$\phi_{22} = \frac{0 - (-\theta/(1+\theta^2))^2}{1 - (-\theta/(1+\theta^2))^2} = \frac{\theta^2}{(1+\theta^2)^2 - \theta^2} = \frac{\theta^2}{1+\theta^2+\theta^4}.$$

Exercise 6.10 Show that the general expression for the partial autocorrelation function of an MA(1) process given in Equation (6.2.6), page 114, satisfies the Yule-Walker recursion given in Equation (6.2.7).

We need to show that Equation (6.2.6) $\phi_{kk} = -\dfrac{\theta^k(1-\theta^2)}{1-\theta^{2(k+1)}}$ for $k \geq 1$ satisfies the Yule-walker equations.

The $k \times k$ Yule-Walker Equations (6.2.7) written in matrix form are

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{bmatrix}$$

In particular, these become:

$$\begin{bmatrix} \rho \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & \rho & 0 & \ldots & 0 & 0 \\ \rho & 1 & \rho & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & \rho \\ 0 & 0 & 0 & \ldots & \rho & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk-1} \\ \phi_{kk} \end{bmatrix}$$

for an MA(1) process where, for simplicity, we write ρ for ρ_1. For a given k, write this matrix equation as $\boldsymbol{b} = A_k\boldsymbol{x}$. We will use Cramer's Rule to solve for x_k (= ϕ_{kk}). By Cramer's Rule $x_k = |B|/|A_k|$ where B is the matrix A_k with the k^{th} column replaced by the column vector $\boldsymbol{x}$ andthe vertical bars (| |) indicate determinants. In particular,

$$B = \begin{bmatrix} 1 & \rho & 0 & \dots & 0 & \rho \\ \rho & 1 & \rho & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & \rho & 0 \end{bmatrix}$$

and expanding the determinant around the last column using the special form of B, we obtain

$$|B| = (-1)^{k+1}\rho\left|\begin{bmatrix} \rho & 1 & \rho & \dots & 0 & 0 \\ 0 & \rho & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & 0 & 0 & \dots & \rho & 1 \\ 0 & 0 & 0 & \dots & 0 & \rho \end{bmatrix}\right| = (-1)^{k+1}\rho\rho^{k-1} = (-1)^{k+1}\rho^k$$

We now develop a second-order recursion for the determinant $|A_k|$. Expanding the determinant $|A_k|$ around the first row, we can write $|A_k| = 1|A_{k-1}| - \rho\rho|A_{k-2}|$ or $|A_k| = |A_{k-1}| - \rho^2|A_{k-2}|$ which is valid for $k > 2$.

To start the recursion, we have $|A_1| = 1$ and $|A_2| = 1 - \rho^2$. Then $|A_3| = |A_2| - \rho^2|A_1| = 1 - \rho^2 - \rho^2 1 = 1 - 2\rho^2$. So

$$\phi_{33} = x_3 = \frac{|B|}{|A_3|} = \frac{(-1)^4\rho^3}{1-2\rho^2} = \frac{\left(-\frac{\theta}{1+\theta^2}\right)^3}{1-2\left(-\frac{\theta}{1+\theta^2}\right)^2} = -\frac{\theta^3}{(1+\theta^2)^3 - 2\theta^2(1+\theta^2)} = -\frac{\theta^3}{(1+\theta^2)(1+\theta^4)}$$

$$= -\frac{\theta^3(1-\theta^2)}{[1+\theta^{2(3+1)}]} \qquad \text{as required by Equation (6.2.6).}$$

To use induction on k we next proceed to $k = 4$. $|A_4| = |A_3| - \rho^2|A_2| = (1-2\rho^2) - \rho^2(1-\rho^2) = 1 - 3\rho^2 + 4\rho^4$. So

$$\phi_{44} = x_4 = \frac{|B|}{|A_4|} = \frac{(-1)^5\rho^4}{1-3\rho^2+4\rho^4} = \frac{(-1)\left(-\frac{\theta}{1+\theta^2}\right)^4}{1-3\left(-\frac{\theta}{1+\theta^2}\right)^2 + 4\left(-\frac{\theta}{1+\theta^2}\right)^4} = -\frac{\theta^4}{1+\theta^2+\theta^4+\theta^6+\theta^8}$$

$$= -\frac{\theta^4(1-\theta^2)}{[1-\theta^{2(4+1)}]} \qquad \text{as required by Equation (6.2.6).}$$

Now we proceed by induction on k. Suppose that Equation 6.2.6) is satisfied by for k–1 and k–2. Then

$$\phi_{kk} = x_k = \frac{|B|}{|A_k|} = \frac{(-1)^{k+1}\rho^k}{|A_{k-1}| - \rho^2|A_{k-2}|} = \frac{(-1)^{k+1}\rho^k}{\frac{(-1)^k\rho^{k-1}}{\phi_{k-1,k-1}} - \rho^2\frac{(-1)^{k-1}\rho^{k-2}}{\phi_{k-2,k-2}}} = -\frac{\rho}{\frac{1}{\phi_{k-1,k-1}} + \frac{1}{\phi_{k-2,k-2}}}$$

$$\vdots$$

$$= -\frac{\theta^k(1-\theta^2)}{1-\theta^{2k+2}} \qquad \text{as required by Equation (6.2.6) after (tedious) but straight forward algebra!}$$

Hence, by induction, the required result is established.

Exercise 6.11 Use Equation (6.2.8), page 114, to find the (theoretical) partial autocorrelation function for an AR(2) model in terms of ϕ_1 and ϕ_2 and lag $k = 1, 2, 3, \dots$.

Trivially $\rho_1 = \phi_{11}$. Now, for $k = 2$,

$$\left.\begin{array}{l} \rho_1 = \phi_{21}\rho_0 + \phi_{22}\rho_1 \\ \rho_2 = \phi_{21}\rho_1 + \phi_{22}\rho_0 \end{array}\right\} \text{ or } \left.\begin{array}{l} \rho_1 = \phi_{21} + \phi_{22}\rho_1 \\ \rho_2 = \phi_{21}\rho_1 + \phi_{22} \end{array}\right\}\text{. But we know that an AR(2) process satisfies } \left.\begin{array}{l} \rho_1 = \phi_1 + \phi_2\rho_1 \\ \rho_2 = \phi_1\rho_1 + \phi_2 \end{array}\right\} \text{ so}$$

it must be that $\phi_{22} = \phi_2$. For $k = 3$, we have

$$\left.\begin{aligned}\rho_1 &= \phi_{31}\rho_0 + \phi_{32}\rho_1 + \phi_{33}\rho_2 \\ \rho_2 &= \phi_{31}\rho_1 + \phi_{32}\rho_0 + \phi_{33}\rho_1 \\ \rho_3 &= \phi_{31}\rho_2 + \phi_{32}\rho_1 + \phi_{33}\rho_0\end{aligned}\right\} \text{ or } \left.\begin{aligned}\rho_1 &= \phi_{31} + \phi_{32}\rho_1 + \phi_{33}\rho_2 \\ \rho_2 &= \phi_{31}\rho_1 + \phi_{32} + \phi_{33}\rho_1 \\ \rho_3 &= \phi_{31}\rho_2 + \phi_{32}\rho_1 + \phi_{33}\end{aligned}\right\}. \text{ But we know that } \left.\begin{aligned}\rho_1 &= \phi_1 + \phi_2\rho_1 + 0\rho_2 \\ \rho_2 &= \phi_1\rho_1 + \phi_2 + 0\rho_1 \\ \rho_3 &= \phi_1\rho_2 + \phi_2\rho_1 + 0\end{aligned}\right\} \text{ so we}$$

must have $\phi_{33} = 0$. Similarly, $\phi_{kk} = 0$ for $k > 2$.

Exercise 6.12 From a time series of 100 observations, we calculate $r_1 = -0.49$, $r_2 = 0.31$, $r_3 = -0.21$, $r_4 = 0.11$, and $|r_k| < 0.09$ for $k > 4$. On this basis alone, what ARIMA model would we tentatively specify for the series?

Using $2/\sqrt{100} = 0.2$ as a guide, we might consider MA(2) or MA(3) as possibilities. If MA(2) is tentatively assumed, then Equation (6.1.11), page 112, gives $Var(r_3) \approx (1 + 3[(-0.49)^2 + (0.31)^2])/100 = 0.016724$ so that $r_3/(\sqrt{Var(r_3)}) \approx -0.21/(\sqrt{0.016724}) = -1.62$ and MA(2) is not rejected.

Exercise 6.13 A stationary time series of length 121 produced sample partial autocorrelation of $\hat{\phi}_{11} = 0.8$, $\hat{\phi}_{22} = -0.6$, $\hat{\phi}_{33} = 0.08$, and $\hat{\phi}_{44} = 0.00$. Based on this information alone, what model would we tentatively specify for the series?

$2/(\sqrt{121}) = 0.181$ so an AR(2) model should be entertained.

Exercise 6.14 For a series of length 169, we find that $r_1 = 0.41$, $r_2 = 0.32$, $r_3 = 0.26$, $r_4 = 0.21$, and $r_5 = 0.16$. What ARIMA model fits this pattern of autocorrelations?

Note that $r_2/r_1 = 0.78$, $r_3/r_2 = 0.81$, $r_4/r_3 = 0.81$, and $r_5/r_4 = 0.76$ and we do not have $r_k \approx r_1^k$. This would seem to rule out an AR(1) model but support an ARMA(1,1) with $\phi \approx 0.8$.

Exercise 6.15 The sample ACF for a series and its first difference are given in the following table. Here $n = 100$.

lag	**1**	**2**	**3**	**4**	**5**	**6**
ACF for Y_t	0.97	0.97	0.93	0.85	0.80	0.71
ACF for ∇Y_t	–0.42	0.18	–0.02	0.07	–0.10	–0.09

Based on this information alone, which ARIMA model(s) would we consider for the series?

The lack of decay in the sample acf suggests nonstationarity After differencing the correlations seem much more reasonable. In particular, $(1 + 2(-0.42)^2)/100 = 0.0135$ and $0.18/(\sqrt{0.0135}) = 1.55$. Therefore, an IMA(1,1) model warrants further consideration.

Exercise 6.16 For a series of length 64, the sample partial autocorrelations are given as:

Lag	**1**	**2**	**3**	**4**	**5**
PACF	0.47	–0.34	0.20	0.02	–0.06

Which models should we consider in this case?

Notice that $2/(\sqrt{64}) = 0.25$ and that all partial autocorrelations from lag 3 on are smaller in magnitude than 0.25. This suggests an AR(2) model for the series.

Exercise 6.17 Consider an AR(1) series of length 100 with $\phi = 0.7$.

(a) Would you be surprised if $r_1 = 0.6$?

For an AR(1) with $\phi = 0.7$ and $n = 100$, Exhibit (6.1), page 111, shows $\sqrt{Var(r_1)} \approx 0.71/(\sqrt{100}) = 0.071$ and $r_1 = 0.6$ is less than two standard deviations from $\rho_1 = \phi = 0.7$. We should not be surrprised at all.

(b) Would $r_{10} = -0.15$ be unusual?

For an AR(1) with $\phi = 0.7$ and $n = 100$, Exhibit (6.1), page 111, shows $\sqrt{Var(r_{10})} \approx 1.70/(\sqrt{100}) = 0.17$. Thus $r_{10} = -0.15$ is less than one standard deviation away from its approximate mean of $\rho_{10} = (0.7)^{10} = 0.028$.

Exercise 6.18 Suppose the $\{X_t\}$ is a stationary AR(1) process with parameter ϕ but that we can only observe $Y_t = X_t + N_t$ where $\{N_t\}$ is the white noise measurement error independent of $\{X_t\}$.

(a) Find the autocorrelation function for the observed process in terms of ϕ, σ_X^2, and σ_N^2.

From the solution to Exercise (2.24), page 23, we know that $Corr(Y_t, Y_{t-k}) = \dfrac{Corr(X_t, X_{t-k})}{1 + \sigma_N^2/\sigma_X^2} = c\phi^k$ for $k \ge 1$.

(b) Which ARIMA model might we specify for $\{Y_t\}$?

This is the pattern of an ARMA(1,1) model.

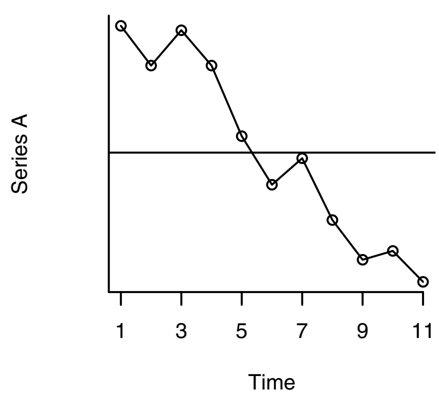
Exercise 6.19 The time plots of two series are shown below.

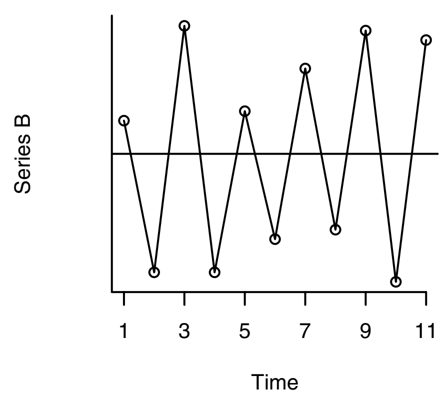
(a) For each of the series, describe r_1 using the terms strongly positive, moderately positive, near zero, moderately negative, or strongly negative. Do you need to know the scale of measurement for the series to answer this?

The lag one autocorrelation for Series A will be strongly positive since neighboring points in time are almost universally on the same side of the mean. The scale is not relevant. The lag one autocorrelation for Series B, on the other hand, will be strongly negative since neighboring points in time are almost universally on opposite sides of the mean.

(b) Repeat part (a) for r_2.

The lag two autocorrelation for Series A will also be positive again since points two apart in in time are almost universally on the same side of the mean. The lag two autocorrelation for Series B will be (strongly) positive since points two apart in time are almost uniersally on the same side of the mean.





Exercise 6.20 Simulate an AR(1) time series with $n = 48$ and with $\phi = 0.7$.

```
> set.seed(241357); series=arima.sim(n=48,list(ar=0.7))
```

(a) Calculate the theoretical autocorrelations at lag 1 and lag 5 for this model.

$\rho_1 = 0.7$ and $\rho_5 = (0.7)^5 = 0.16807$.

(b) Calculate the sample autocorrelations at lag 1 and lag 5 and compare the values with their theoretical values. Use Equations (6.1.5) and (6.1.6), page 111, to quantify the comparisons.

```
> acf(series,lag.max=5)[1:5]
```

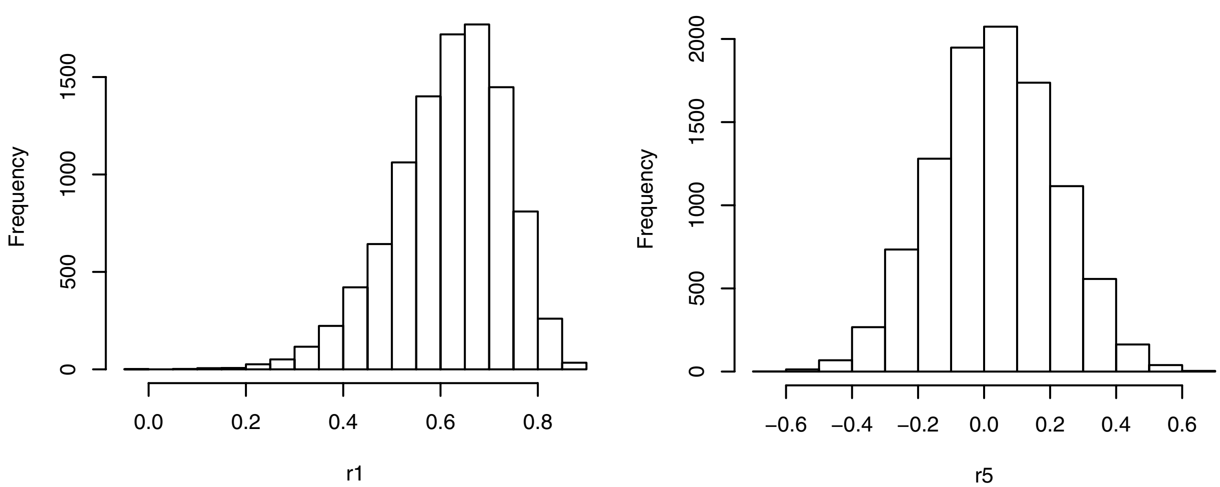
```
       Autocorrelations of series 'series', by lag

           1     2     3     4     5
       0.768 0.626 0.436 0.318 0.14
```

The standard error of r_1 is $\sqrt{(1-\phi^2)/n} = \sqrt{(1-(0.7)^2)/48} = \sqrt{0.010625} = \sqrt{0.010625} \approx 0.10$ and of r_5 is $\sqrt{\frac{1}{n}\left[\frac{1+\phi^2}{1-\phi^2}\right]} = \sqrt{\frac{1}{48}\left[\frac{1+(0.7)^2}{1-(0.7)^2}\right]} \approx 0.25$. With these standard errors in mind, the estimates of 0.768 and 0.14 are excellent estimates of 0.7 and 0.16807, respectively.

(c) Repeat part (b) with a new simulation. Describe how the precision of the estimate varies with different samples selected under identical conditions.

(d) If software permits, repeat the simulation of the series and calculation of r_1 and r_5 many times and form the sampling distributions of r_1 and r_5. Describe how the precision of the estimate varies with different samples selected under identical conditions. How well does the large-sample variance given in Equation (6.1.5), page 111, approximate the variance in your sampling distribution?

```
> set.seed(132435); r1=rep(NA,10000); r5=r1 # We are doing 10,000 replications.
> for (k in 1:10000) {series=arima.sim(n=48, list(ar=0.7))
> ;r1[k]=acf(series,lag.max=1,plot=F)$acf[1]
> ;r5[k]=acf(series,lag.max=5,plot=F)$acf[5]}
> hist(r1); mean(r1); sd(r1); median(r1)
> hist(r5); mean(r5); sd(r5); median(r5)
```



For the sampling distribution of r_1, the mean is 0.618 ($\rho_1 = 0.7$) and the median is 0.631. This agrees with the observed skewness toward the lower values. The standard deviation in this distribution is 0.11 which agrees well with asymptotic theory (0.10).

The sampling distribution of r_5 has a mean of 0.033 ($\rho_5 = 0.168$) and a median of 0.032 and this agrees with the near symmetry of this distribution. The standard deviation in this distribution is 0.18 which agrees reasonably well with asymptotic theory (0.25).

This exercise illustrates the difficulty of estimating the autocorrelation function of a simple AR(1) series with a sample size of $n = 48$. You might want to repeat this exercise with a larger n, say, $n = 96$ or larger. You could also try different values for ϕ.

Exercise 6.21 Simulate an MA(1) time series with $n = 60$ and with $\theta = 0.5$.

```
> set.seed(6453421); series=arima.sim(n=60,list(ma=-0.5))
```

(a) Calculate the theoretical autocorrelation at lag 1 for this model.

$\rho_1 = -\theta/(1+\theta^2) = -(0.5)/(1+(0.5)^2) = -0.4$.

(b) Calculate the sample autocorrelation at lag 1, and compare the value with its theoretical value. Use Exhibit (6.2), page 112, to quantify the comparisons.

```
> acf(series,lag.max=1)[1]
```

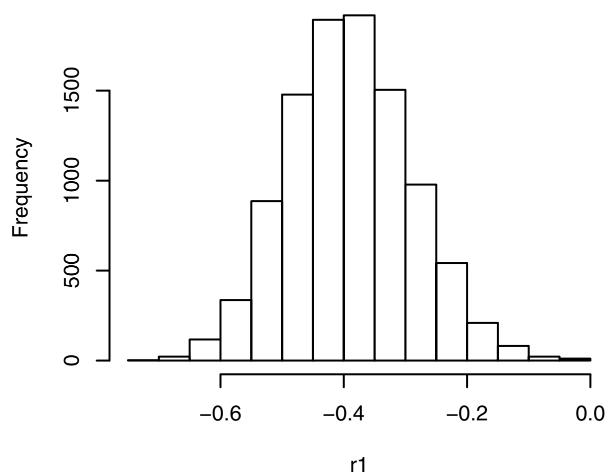
```
Autocorrelations of series ‘series’, by lag

     1
-0.362
```

The standard error of r_1 is $\sqrt{c_{11}/n} = \sqrt{1-3\rho_1^2+4\rho_1^4/n} = \sqrt{1-3(-0.4)^2+4(-0.4)^4/60} \approx 0.10$. The estimate of –0.362 is well within two standrad errors of the true value of –0.4.

(c) Repeat part (b) with a new simulation. Describe how the precision of the estimate varies with different samples selected under identical conditions.

(d) If software permits, repeat the simulation of the series and calculation of r_1 many times and form the sampling distribution of r_1. Describe how the precision of the estimate varies with different samples selected under identical conditions. How well does the large-sample variance given in Exhibit (6.2), page 112, approximate the variance in your sampling distribution?

```
> set.seed(534261); r1=rep(NA,10000); r5=r1 # We are doing 10,000 replications.
> for (k in 1:10000) {series=arima.sim(n=60, list(ma=-0.5))
> ;r1[k]=acf(series,lag.max=1,plot=F)$acf[1]}
> hist(r1); mean(r1); sd(r1); median(r1)
```



Remember that $\rho_1 = -0.4$. Here the mean of the sampling distribution is –0.390 (median = –0.393) and the standard deviation 0.100. The large-sample standard deviation given in Exhibit (6.2), page 112, is $0.79/\sqrt{60} = 0.102$ so the large-sample value is an excellent approximation to the one obtained in the sampling distribution.

Exercise 6.22 Simulate an AR(1) time series with $n = 48$, with

```
> set.seed(5342310; series=arima.sim(n=48,list(ar=0.9))
```

(a) $\phi = 0.9$, and calculate the theoretical autocorrelations at lag 1 and lag 5;

$\rho_1 = 0.9$ and $\rho_5 = (0.7)^5 = 0.16807$.

(b) $\phi = 0.6$, and calculate the theoretical autocorrelations at lag 1 and lag 5;

$\rho_1 = 0.6$ and $\rho_5 = (0.6)^5 = 0.07776$.

(c) $\phi = 0.3$, and calculate the theoretical autocorrelations at lag 1 and lag 5.

$\rho_1 = 0.3$ and $\rho_5 = (0.3)^5 = 0.00243$.

(d) For each of the series in parts (a), (b), and (c), calculate the sample autocorrelations at lag 1 and lag 5 and compare the values with their theoretical values. Use Equations (6.1.5) and (6.1.6), page 111, to quantify the comparisons. In general, describe how the precision of the estimate varies with the value of ϕ.

Case (a) $\phi = 0.9$: Recall that for an AR(1), $\sqrt{Var(r_1)} \approx \sqrt{\frac{(1-\phi^2)}{n}}$ and $\sqrt{Var(r_k)} \approx \sqrt{\frac{1}{n}\left[\frac{1+\phi^2}{1-\phi^2}\right]}$. For $\phi = 0.9$ we have

$\sqrt{Var(r_1)} \approx 0.06$ and $\sqrt{Var(r_5)} \approx 0.40$. From the simulated series we obtain

```
> set.seed(5342310); series=arima.sim(n=48,list(ar=0.9)); acf(series)[1:5]
```

```
      Autocorrelations of series 'series', by lag

          1     2     3     4     5
      0.862 0.739 0.569 0.420 0.232
```

The estimate of 0.862 compares well with the true value of $\rho_1 = 0.9$ when the standard error of 0.06 kept in mind. Similarly, the estimate of 0.232 compares well with the true value of $\rho_5 = 0.16807$ when the standard error is 0.40.

Case (b) $\phi = 0.6$: Now $\sqrt{Var(r_1)} \approx 0.12$ and $\sqrt{Var(r_5)} \approx 0.21$. From the simulation

```
> set.seed(5342310); series=arima.sim(n=48,list(ar=0.6)); acf(series)[1:5]
```

```
      Autocorrelations of series 'series', by lag

          1     2     3     4     5
      0.617 0.388 0.392 0.228 0.191
```

The estimate of 0.617 compares well with the true value of $\rho_1 = 0.6$ when the standard error is 0.12. The estimate of 0.191 compares well with the true value of $\rho_5 = 0.07776$ when the standard error is 0.21.

Case (c) $\phi = 0.3$: In this case $\sqrt{Var(r_1)} \approx 0.14$ and $\sqrt{Var(r_5)} \approx 0.16$. From the simulation

```
> set.seed(5342310); series=arima.sim(n=48,list(ar=0.3)); acf(series)[1:5]
```

```
      Autocorrelations of series 'series', by lag

           1      2      3      4      5
       0.188  0.032  0.294 -0.081  0.048
```

The estimate of 0.188 compares well with the true value of $\rho_1 = 0.3$ when the standard error is 0.14. The estimate of 0.048 compares well with the true value of $\rho_5 = 0.00243$ when the standard error is 0.16.

Exercise 6.23 Simulate an AR(1) time series with $\phi = 0.6$, with

(a) $n = 24$, and estimate $\rho_1 = \phi = 0.6$ with r_1;

```
> set.seed(162534)
> series=arima.sim(model=list(order=c(1,0,0),ar=0.6),n=24); acf(series)[1]
```

```
      Autocorrelations of series 'series', by lag

          1
      0.459
```

(b) $n = 60$, and estimate $\rho_1 = \phi = 0.6$ with r_1;

```
> series=arima.sim(model=list(order=c(1,0,0),ar=0.6),n=60); acf(series)[1]
```

```
      Autocorrelations of series 'series', by lag

          1
      0.385
```

(c) $n = 120$, and estimate $\rho_1 = \phi = 0.6$ with r_1.

```
> series=arima.sim(model=list(order=c(1,0,0),ar=0.6),n=120); acf(series)[1]
```

```
      Autocorrelations of series 'series', by lag

          1
```

```
0.627
```

(d) For each of the series in parts (a), (b), and (c), compare the estimated values with the theoretical value. Use Equation (6.1.5), page 111, to quantify the comparisons. In general, describe how the precision of the estimate varies with the sample size.

Case (a) $n = 24$: $\sqrt{Var(r_1)} \approx \sqrt{\frac{1-(0.6)^2}{24}} \approx 0.16$. The estimate of 0.459 is well within two standard errors of the true value of 0.6.

Case (b) $n = 60$: $\sqrt{Var(r_1)} \approx \sqrt{\frac{1-(0.6)^2}{60}} \approx 0.10$. Notice that even though the sample size is larger, this series gave a less accurate estimate 0.385 of $\rho_1 = \phi = 0.6$ than the one in part (a). In fact this estimate is more than two standard errors away from the true value of $\rho_1 = \phi = 0.6$. However, in general, estimates with larger sample sizes give better estimates as the standard errors are smaller.

Case (c) $n = 120$: $\sqrt{Var(r_1)} \approx \sqrt{\frac{1-(0.6)^2}{120}} \approx 0.07$. This estimate 0.627 is the best one of the three and has the smallest standard error.

Exercise 6.24 Simulate an MA(1) time series with $\theta = 0.7$, with

(a) $n = 24$, and estimate ρ_1 with r_1;

First recall that $\rho_1 = -\theta/(1+\theta^2) = -(0.7)/(1+(0.7)^2) = -0.4$.

```
> set.seed(172534); series=arima.sim(n=24,list(ma=-0.7)); acf(series)[1]
```

```
Autocorrelations of series 'series', by lag

     1
-0.595
```

(b) $n = 60$, and estimate ρ_1 with r_1;

```
> set.seed(172534); series=arima.sim(n=60,list(ma=-0.7)); acf(series)[1]
```

```
Autocorrelations of series 'series', by lag

     1
-0.527
```

(c) $n = 120$, and estimate ρ_1 with r_1.

```
> set.seed(172534); series=arima.sim(n=120,list(ma=-0.7)); acf(series)[1]
```

```
Autocorrelations of series 'series', by lag

     1
-0.458
```

(d) For each of the series in parts (a), (b), and (c), compare the estimated values of ρ_1 with the theoretical value. Use Exhibit (6.2), page 112, to quantify the comparisons. In general, describe how the precision of the estimate varies with the sample size.

The standard errors are $0.73/\sqrt{n} = 0.73/\sqrt{24} \approx 0.15$, $0.73/\sqrt{60} = 0.09$, and $0.73/\sqrt{120} = 0.07$, respectively. With these particular simulations, the estimates get better with increasing sample size.

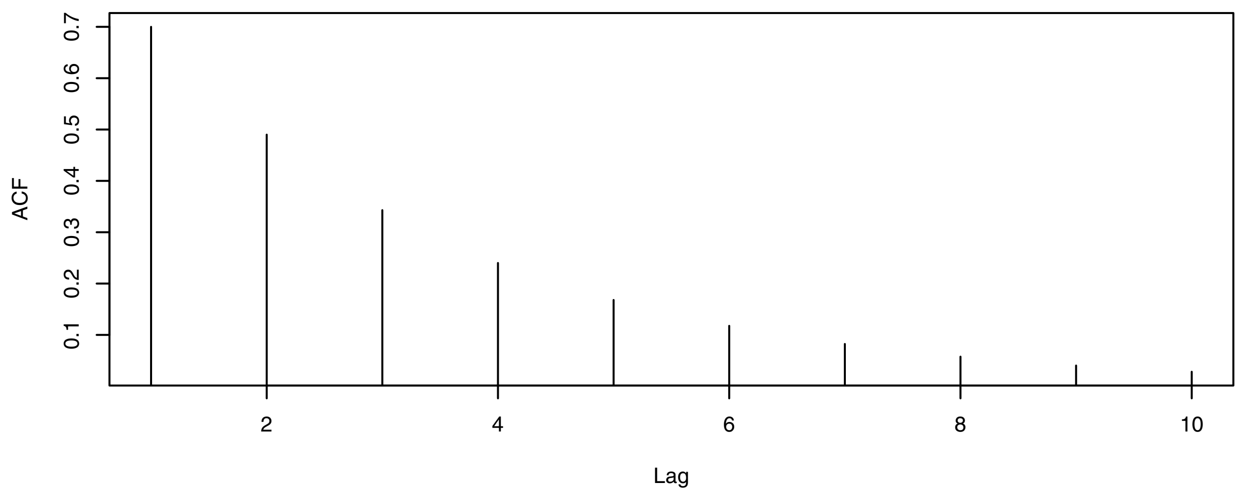
Exercise 6.25 Simulate an AR(1) time series of length $n = 36$ with $\phi = 0.7$.

(a) Calculate and plot the theoretical autocorrelation function for this model. Plot sufficient lags until the correlations are negligible.

```
> round(ARMAacf(ar=0.7,lag.max=10),digits=3)
```

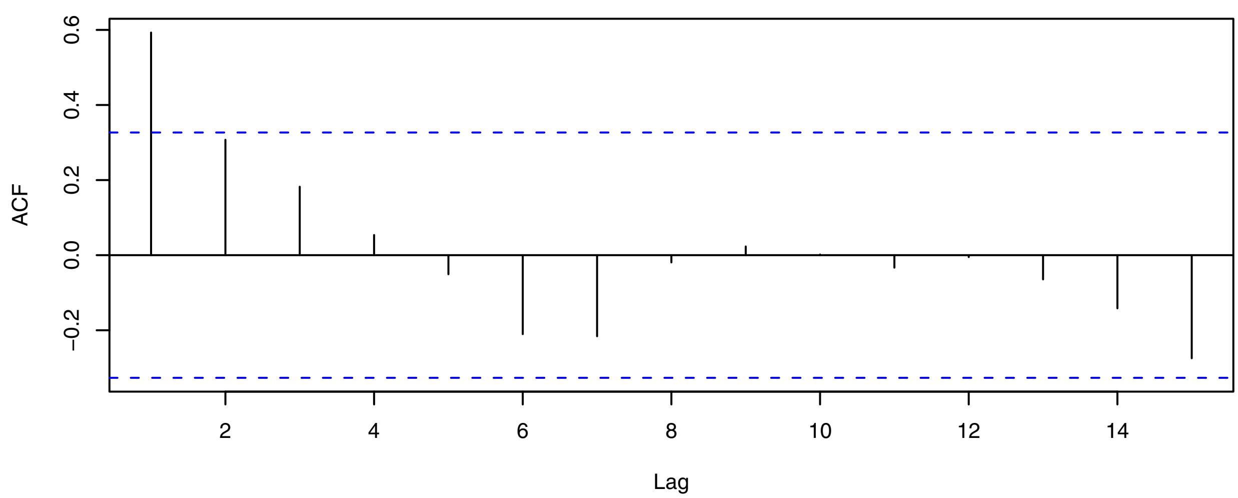
```
     0     1     2     3     4     5     6     7     8     9    10
 1.000 0.700 0.490 0.343 0.240 0.168 0.118 0.082 0.058 0.040 0.028
```

```
> win.graph(width=6.5,height=3,pointsize=8)
> ACF=ARMAacf(ar=0.7,lag.max=10)
> plot(y=ACF[-1],x=1:10,xlab='Lag',ylab='ACF',type='h'); abline(h=0)
```



(b) Calculate and plot the sample ACF for your simulated series. How well do the values and patterns match the theoretical ACF from part (a)?

```
> set.seed(162534); series=arima.sim(n=36,list(ar=0.7)); acf(series)
```



The pattern match is not that good but remember that $n = 36$.

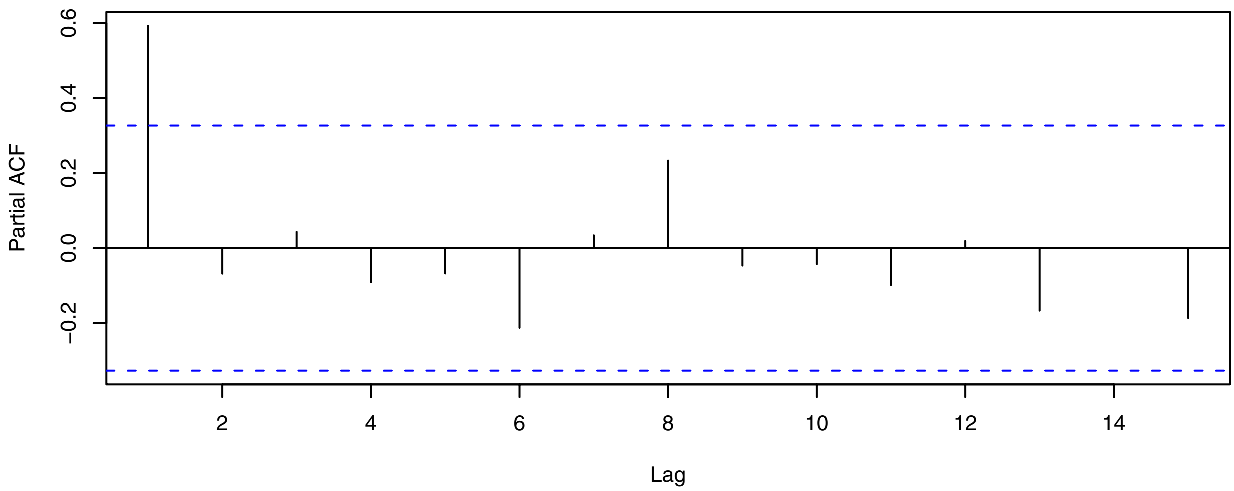
(c) What are the theoretical partial autocorrelations for this model?

$\phi_{11} = 0.7$ and $\phi_{kk} = 0$ otherwise.

(d) Calculate and plot the sample ACF for your simulated series. How well do the values and patterns match the theoretical ACF from part (a)? Use the large-sample standard errors reported in Exhibit (6.1), page 111, to quantify your answer.

See answer for part (b). Here $\sqrt{Var(r_1)} \approx \sqrt{(1-\phi^2)/n} = \sqrt{(1-0.7^2)/36} = 0.12$ so the observed r_1 is well within two standard errors of the true value. Similarly, for higher order lags.

(e) Calculate and plot the sample PACF for your simulated series. How well do the values and patterns match the theoretical PACF from part (c)? Use the large-sample standard errors reported on page 115 to quantify your answer.




```
> pacf(series)
```

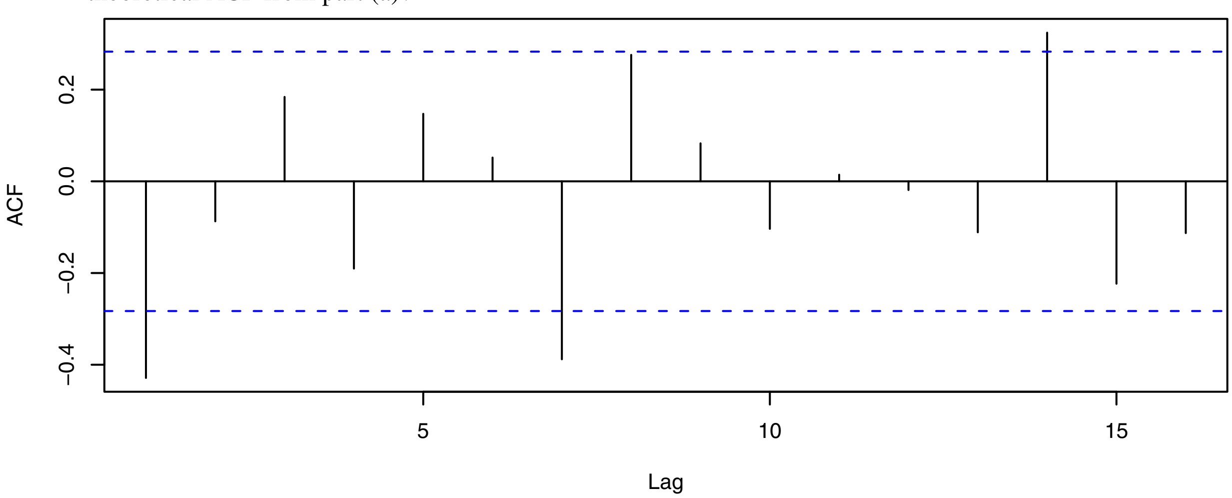
Using the approximate standard errors of $1/\sqrt{n} = 1/\sqrt{36} = 0.167$, the sample pacf matches the theoretical pacf quite well.

Exercise 6.26 Simulate an MA(1) time series of length $n = 48$ with $\theta = 0.5$.

(a) What are the theoretical autocorrelations for this model?

$\rho_1 = -\theta/(1+\theta^2) = -(0.5)/(1+(0.5)^2) = -0.4$ is the only nonzero autocorrelation for this model.

(b) Calculate and plot the sample ACF for your simulated series. How well do the values and patterns match the theoretical ACF from part (a)?



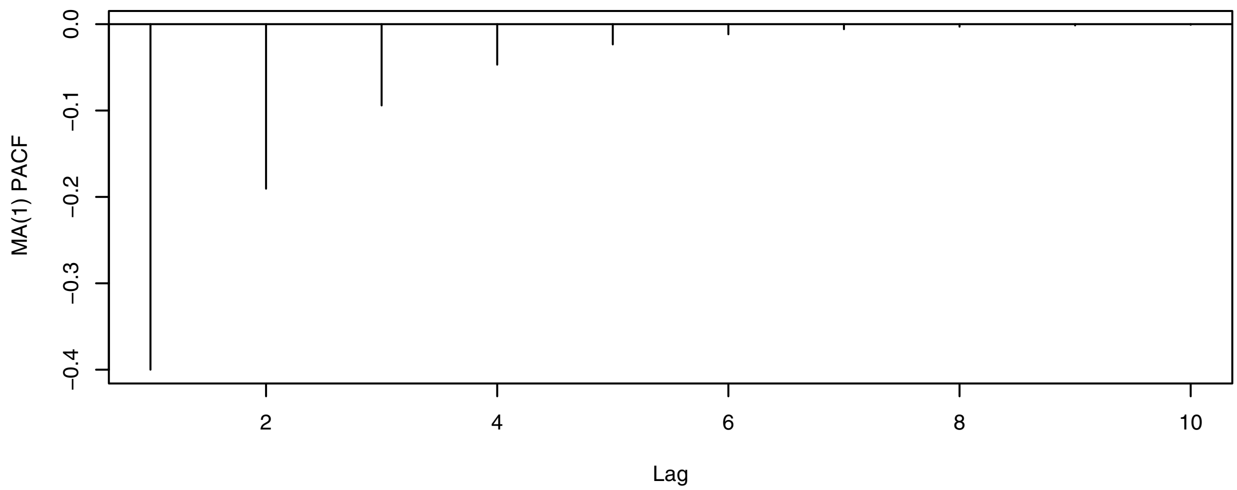

```
> set.seed(162534); series=arima.sim(n=48,list(ma=-0.5)); acf(series)
```

The acf at lag 1 looks reasonable, but there are “significant” but spurious correlations at lags 7 and 14 also.

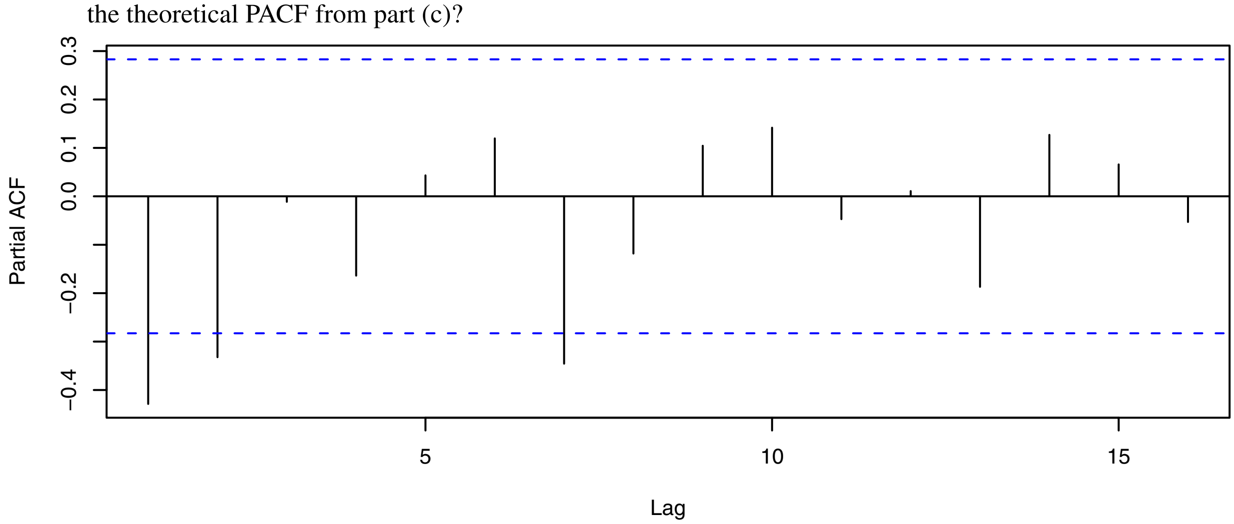
(c) Calculate and plot the theoretical partial autocorrelation function for this model. Plot sufficient lags until the correlations are negligible.

(Hint: See Equation (6.2.6), page 114.) $\phi_{kk} = -\dfrac{\theta^k(1-\theta^2)}{1-\theta^{2(k+1)}}$ for $k \geq 1$

```
> theta=0.5; phikk=rep(NA,10)
> for (k in 1:10) {phikk[k]=-(theta^k)*(1-theta^2)/(1-theta^(2*(k+1)))}
> plot(phikk,type='h',ylab='MA(1) PACF',xlab='Lag'); abline(h=0)
```



(d) Calculate and plot the sample PACF for your simulated series. How well do the values and patterns match the theoretical PACF from part (c)?



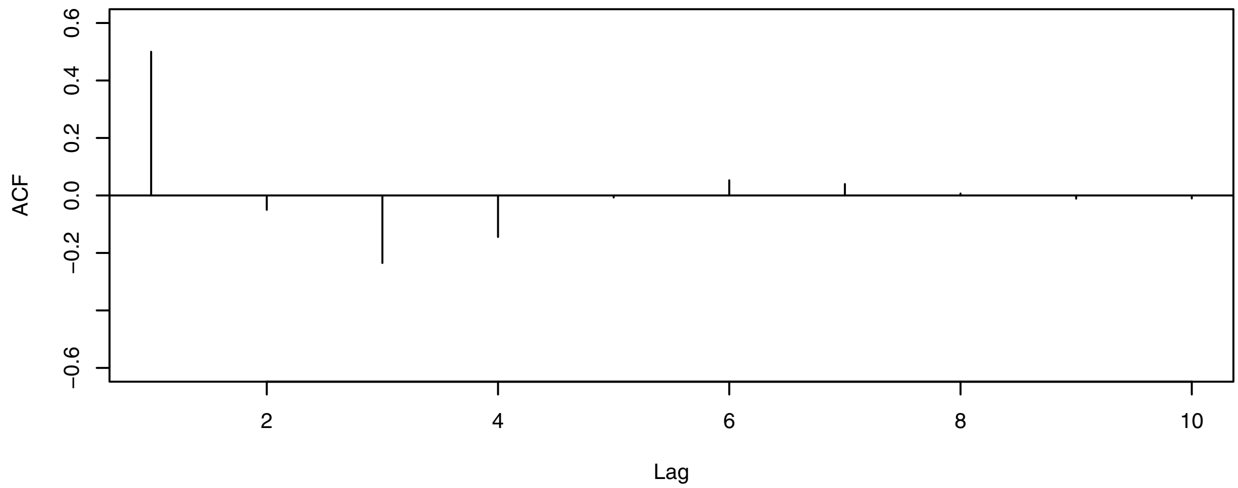

```
> pacf(series)
```

Only the first two lags match well. However, the approximate standard errors of $1/\sqrt{n} = 1/\sqrt{48} = 0.14$ indicate that only the sample pacf at lag 7 is unexpected.

Exercise 6.27 Simulate an AR(2) time series of length $n = 72$ with $\phi_1 = 0.7$ and $\phi_2 = -0.4$.

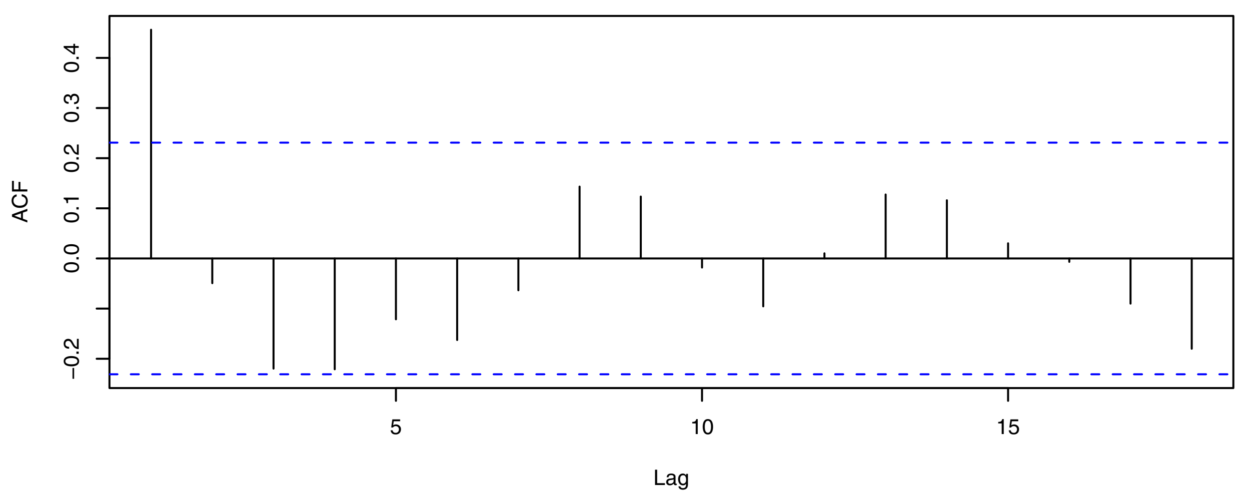
```
> set.seed(162534); series=arima.sim(n=72,list(ar=c(0.7,-0.4)))
```

(a) Calculate and plot the theoretical autocorrelation function for this model. Plot sufficient lags until the correlations are negligible.




```
> phi1=0.7; phi2=-0.4; ACF=ARMAacf(ar=c(phi1,phi2),lag.max=10)
> plot(y=ACF[-1],x=1:10,xlab='Lag',ylab='ACF',type='h',ylim=c(-0.6,0.6)); abline(h=0)
```

(b) Calculate and plot the sample ACF for your simulated series. How well do the values and patterns match the theoretical ACF from part (a)?




```
> acf(series)
```

The lag 1 sample ACF matches well and the "damped sine wave" is somewhat apparent but the values at large lags do not die out like the theoretical ACF.

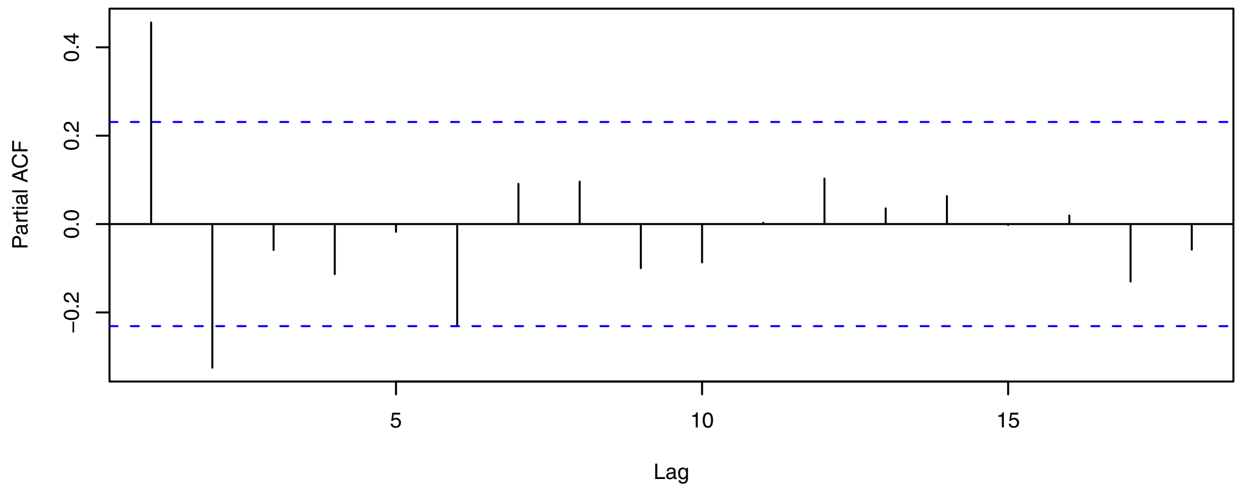
(c) What are the theoretical partial autocorrelations for this model?

$\phi_{11} = 0.5$, $\phi_{22} = 0.7$, and $\phi_{kk} = 0$ otherwise.

(d) Calculate and plot the sample ACF for your simulated series. How well do the values and patterns match the theoretical ACF from part (a)?

This question repeats part (b)

(e) Calculate and plot the sample PACF for your simulated series. How well do the values and patterns match the theoretical PACF from part (c)?



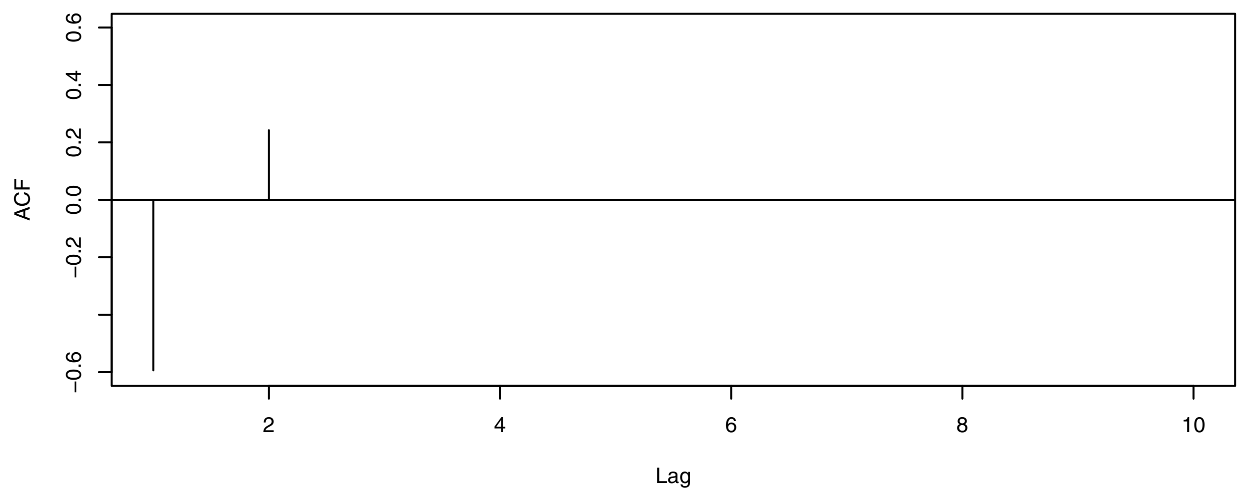

```
> pacf(series)
```

This sample pacf matches the theoretical pacf quite well.

Exercise 6.28 Simulate an MA(2) time series of length $n = 36$ with $\theta_1 = 0.7$ and $\theta_2 = -0.4$.

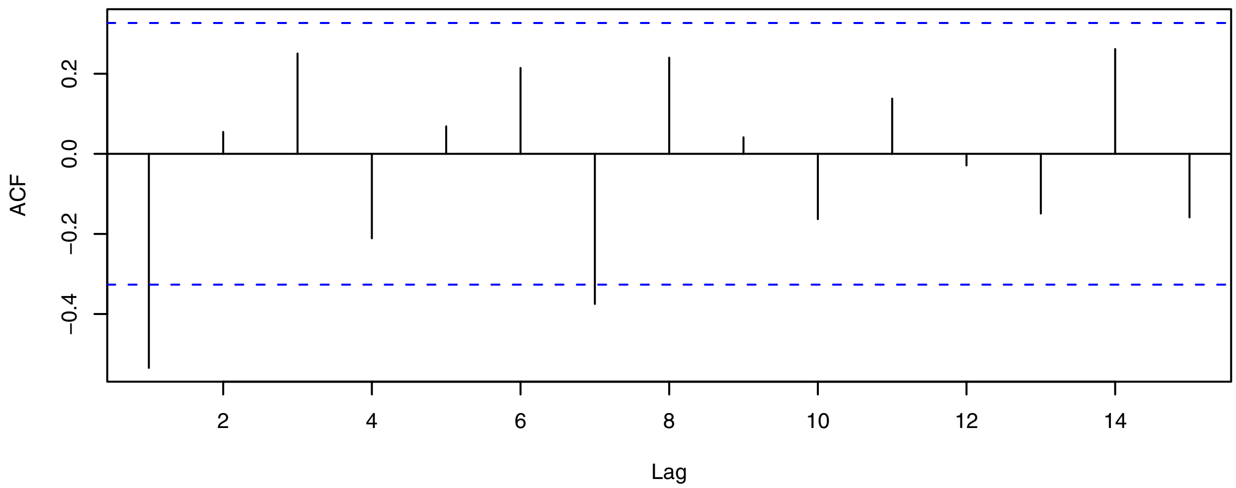
```
> set.seed(162534); series=arima.sim(n=36,list(ma=c(-0.7,0.4)))
```

(a) What are the theoretical autocorrelations for this model?




```
> theta1=0.7; theta2=-0.4; ACF=ARMAacf(ma=c(-theta1,-theta2),lag.max=10)
> plot(y=ACF[-1],x=1:10,xlab='Lag',ylab='ACF',type='h',ylim=c(-.6,.6)); abline(h=0)
```

(b) Calculate and plot the sample ACF for your simulated series. How well do the values and patterns match the theoretical ACF from part (a)?

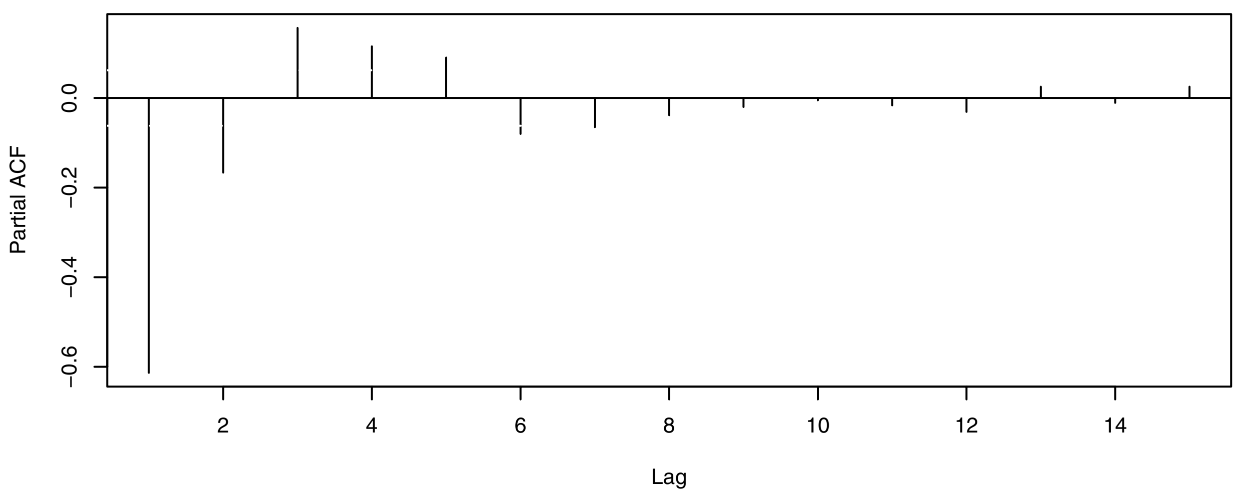



```
> acf(series)
```

With this small sample size we only get a reasonably good match at lag 1.

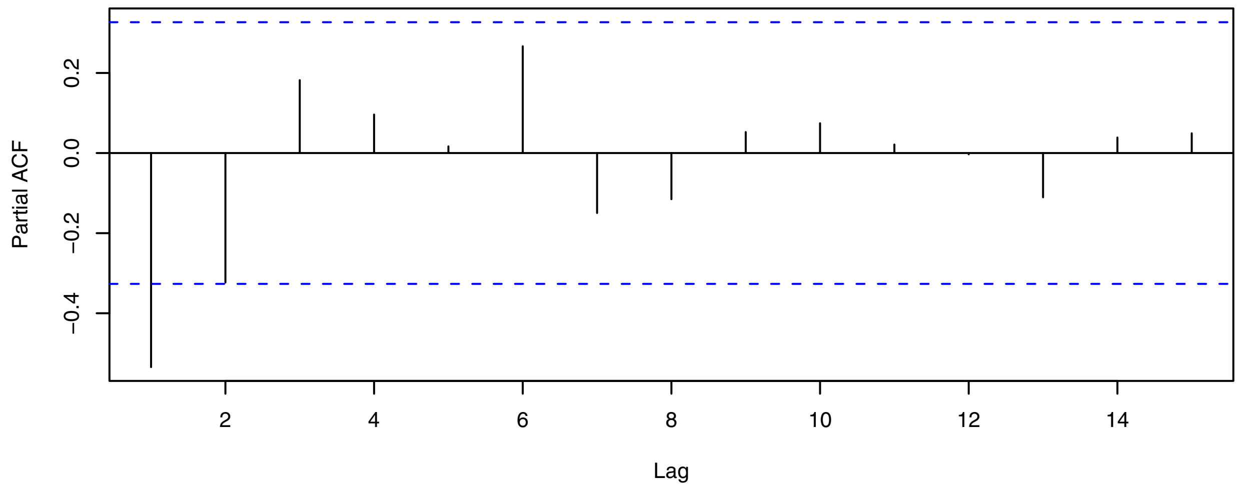
(c) Calculate and plot the theoretical partial autocorrelation function for this model. Plot sufficient lags until the correlations are negligible. (Hint: See Equation (6.2.6), page 114.)

Unfortunately, Equation (6.2.6), page 114, does not apply to an MA(2) model. We can do a very large sample simulation and display the sample pacf to get a good idea of the theoretical pacf for this model.




```
> series2=arima.sim(n=1000,list(ma=c(-0.7,0.4))); pacf(series2,ci.col=NULL)
```

(d) Calculate and plot the sample PACF for your simulated series. How well do the values and patterns match the theoretical PACF from part (c)?



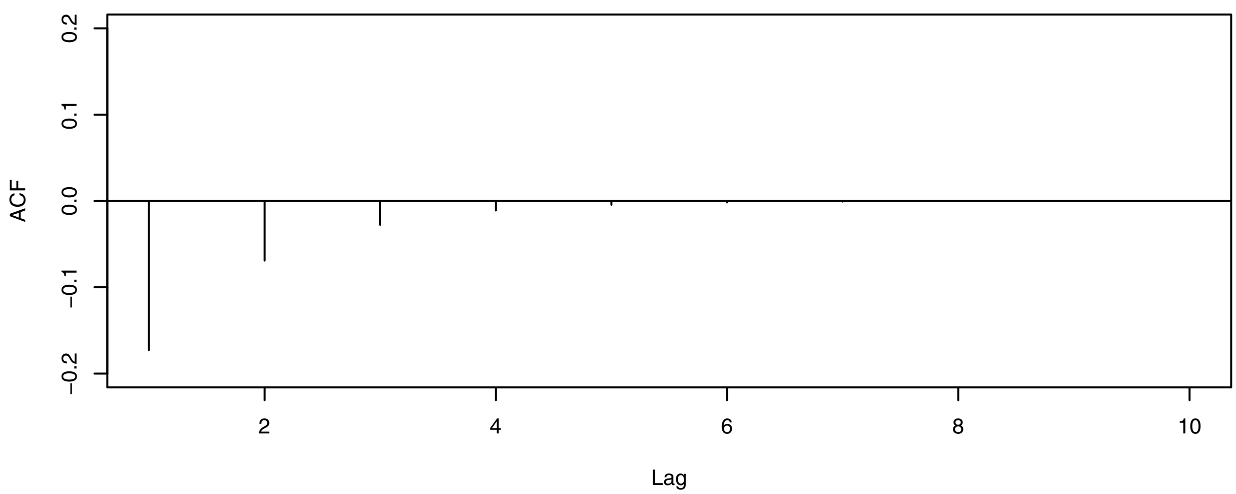

```
> pacf(series2)
```

The first four partials match the theoretical pattern remarkably well—especially given the small sample size of 36.

Exercise 6.29 Simulate a mixed ARMA(1,1) model of length $n = 60$ with $\phi = 0.4$ and $\theta = 0.6$.

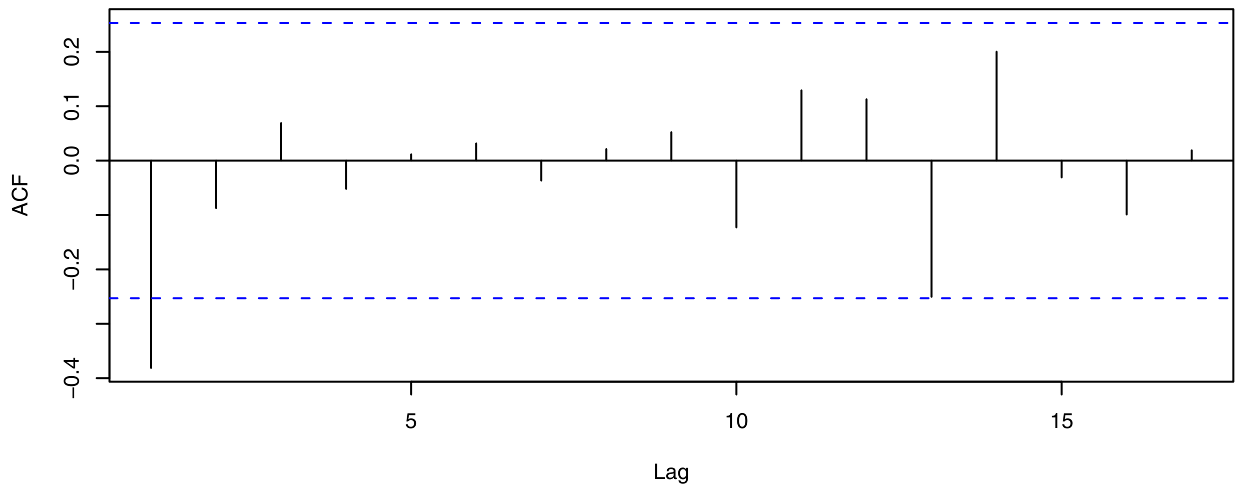
```
> set.seed(762534); series=arima.sim(n=60,list(ar=0.4,ma=-0.6))
```

(a) Calculate and plot the theoretical autocorrelation function for this model. Plot sufficient lags until the correlations are negligible.




```
> phi=0.4; theta=0.6; ACF=ARMAacf(ar=phi,ma=-theta,lag.max=10)
> plot(y=ACF[-1],x=1:10,xlab='Lag',ylab='ACF',type='h',ylim=c(-.2,.2)); abline(h=0)
```

(b) Calculate and plot the sample ACF for your simulated series. How well do the values and patterns match the theoretical ACF from part (a)?




```
> acf(series)
```

The pattern matches somewhat at the first few lags but there is a lot of spurious autocorrelation at higher lags.

(c) Calculate and interpret the sample EACF for this series. Does the EACF help you specify the correct orders for the model?

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	o	o	o	o	o	o	o	o	o	o	o	o	o
1	x	x	o	o	o	o	o	o	o	o	o	o	o	o
2	x	x	o	o	o	o	o	o	o	o	o	o	o	o
3	x	x	x	o	o	o	o	o	o	o	o	o	o	o
4	x	o	x	o	o	o	o	o	o	o	o	o	o	o
5	o	o	o	o	o	o	o	o	o	o	o	o	o	o
6	x	x	o	o	o	o	o	o	o	o	o	o	o	o
7	o	x	o	o	o	o	o	o	o	o	o	o	o	o

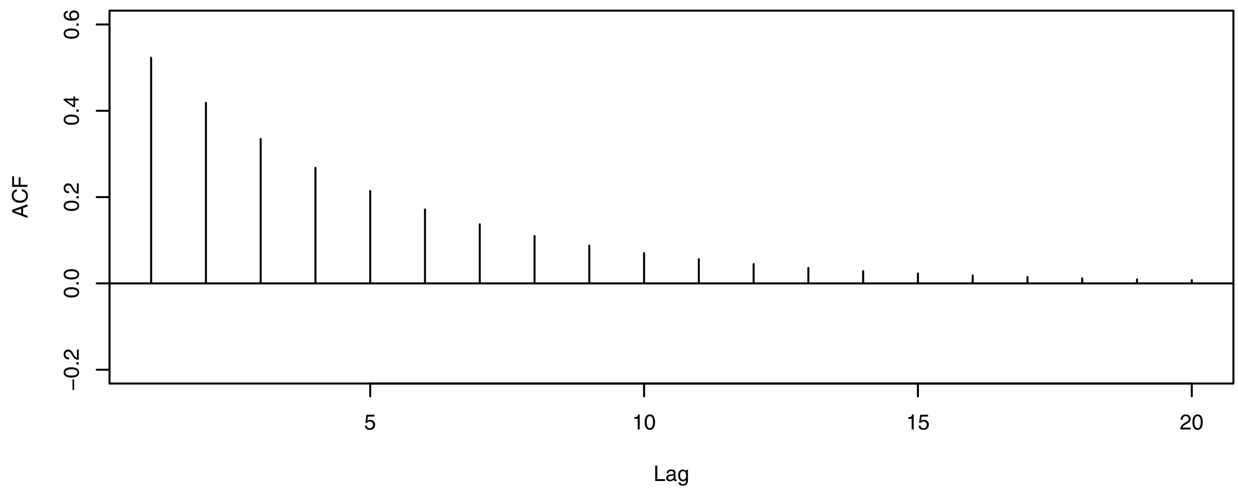
This sample EACF seems to point to an MA(1) model rather than the mixed ARMA(1,1).

(d) Repeat parts (b) and (c) with a new simulation using the same parameter values and sample size.
(e) Repeat parts (b) and (c) with a new simulation using the same parameter values but sample size $n = 36$.
(f) Repeat parts (b) and (c) with a new simulation using the same parameter values but sample size $n = 120$.

Exercise 6.30 Simulate a mixed ARMA(1,1) model of length $n = 100$ with $\phi = 0.8$ and $\theta = 0.4$.

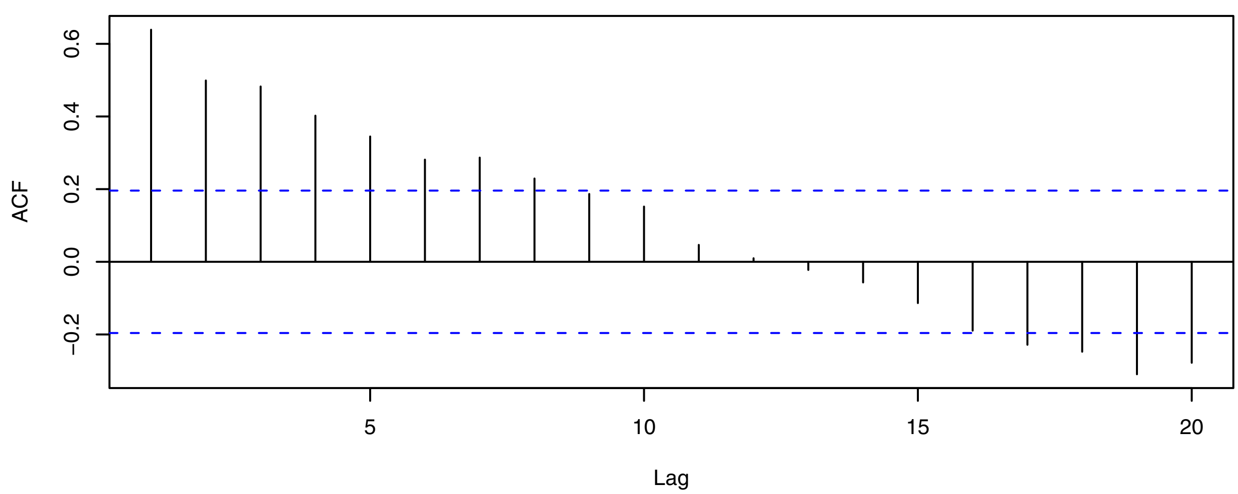
```
> set.seed(325346); series=arima.sim(n=100,list(ar=0.8,ma=-0.4))
```

(a) Calculate and plot the theoretical autocorrelation function for this model. Plot sufficient lags until the correlations are negligible.




```
> phi=0.8; theta=0.4; ACF=ARMAacf(ar=phi,ma=-theta,lag.max=20)
> plot(y=ACF[-1],x=1:20,xlab='Lag',ylab='ACF',type='h',ylim=c(-.2,.6)); abline(h=0)
```

(b) Calculate and plot the sample ACF for your simulated series. How well do the values and patterns match the theoretical ACF from part (a)?




```
> acf(series)
```

The sample acf generally matches the pattern of the theoretical acf for the first 10 or so lags but, as is quite typical, it displays spurious autocorrelation at higher lags.

(c) Calculate and interpret the sample EACF for this series. Does the EACF help you specify the correct orders for the model?

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	x	x	x	x	x	x	x	o	o	o	o	o	o
1	x	o	o	o	o	o	o	o	o	o	o	o	o	o
2	x	x	o	o	o	o	o	o	o	o	o	o	o	o
3	o	o	o	o	o	o	o	o	o	o	o	o	o	o
4	x	o	o	o	o	o	o	o	o	o	o	o	o	o
5	x	x	o	x	o	o	o	o	o	o	o	o	o	o
6	o	x	x	x	o	o	o	o	o	o	o	o	o	o
7	x	o	x	x	o	o	o	o	o	o	o	o	o	o

This sample EACF points to the mixed ARMA(1,1) quite well.

(d) Repeat parts (b) and (c) with a new simulation using the same parameter values and sample size.
(e) Repeat parts (b) and (c) with a new simulation using the same parameter values but sample size $n = 48$.
(f) Repeat parts (b) and (c) with a new simulation using the same parameter values but sample size $n = 200$.

Exercise 6.31 Simulate a nonstationary time series with $n = 60$ according to the model ARIMA(0,1,1) with $\theta = 0.8$. **(Note: This is a better exercise if you use $\theta = -0.8$.)**

```
> set.seed(15243); series=arima.sim(n=60,list(order=c(0,1,1),ma=-0.8))[-1]
```

(a) Perform the (augmented) Dickey-Fuller test on the series with $k = 0$ in Equation (6.4.1), page 128. (With $k = 0$, this is the Dickey-Fuller test and is **not** augmented.) Comment on the results.

```
> library(uroot); ADF.test(series,selectlags=list(Pmax=0),itsd=c(1,0,0))
```

```
--------- ------ - ------ ----
  Augmented Dickey & Fuller test
  --------- ------ - ------ ----

  Null hypothesis: Unit root.
  Alternative hypothesis: Stationarity.

----
  ADF statistic:

        Estimate Std. Error t value Pr(>|t|)
adf.reg   -0.782      0.129   -6.04     0.01

  Lag orders: 0
  Number of available observations: 60
Warning message:
In interpolpval(code = code, stat = adfreg[, 3], N = N) :
  p-value is smaller than printed p-value
```

The Dickey-Fuller test rejects nonstationarity (unit root). This is a Type I error but the series does not look very nonstationary!

(b) Perform the augmented Dickey-Fuller test on the series with k chosen by the software—that is, the “best” value for k. Comment on the results.

```
> ar(diff(series))
```

```
Call:
ar(x = diff(series))

Coefficients:
      1        2        3
-0.7299  -0.5195  -0.3835

Order selected 3  sigma^2 estimated as  0.8227
```

The selected order is 3. This will be used next in the Augmented Dickey-Fuller test.

```
> ADF.test(series,selectlags=list(mode=c(1,2,3)),itsd=c(1,0,0))
```

```
    --------- ------ - ------ ----
      Augmented Dickey & Fuller test
      --------- ------ - ------ ----

      Null hypothesis: Unit root.
      Alternative hypothesis: Stationarity.

    ----
      ADF statistic:

            Estimate Std. Error t value Pr(>|t|)
    adf.reg   -0.414      0.208   -1.99      0.1

      Lag orders: 1 2 3
      Number of available observations: 57
    Warning message:
    In interpolpval(code = code, stat = adfreg[, 3], N = N) :
      p-value is greater than printed p-value
```

Now the Augmented Dickey-Fuller test does not reject nonstationarity (a unit root).

(c) Repeat parts (a) and (b) but use the differences of the simulated series. Comment on the results. (Here, of course, you should reject the unit root hypothesis.)

```
> ADF.test(diff(series),selectlags=list(Pmax=0),itsd=c(1,0,0))
```

```
    --------- ------ - ------ ----
      Augmented Dickey & Fuller test
      --------- ------ - ------ ----

      Null hypothesis: Unit root.
      Alternative hypothesis: Stationarity.

    ----
      ADF statistic:

            Estimate Std. Error t value Pr(>|t|)
    adf.reg   -1.489      0.116 -12.866     0.01

      Lag orders: 0
      Number of available observations: 59
    Warning message:
    In interpolpval(code = code, stat = adfreg[, 3], N = N) :
      p-value is smaller than printed p-value
```

We (correctly) reject the unit root hypothesis for the differenced series.

```
> ar(diff(diff(series))) # order 5 selected
```

```
    Call:
    ar(x = diff(diff(series)))

    Coefficients:
          1        2        3        4        5
    -1.4026  -1.4281  -1.2405  -0.7083  -0.2737

    Order selected 5  sigma^2 estimated as  1.205
```

Order 5 is selected.

```
> ADF.test(diff(series),selectlags=list(mode=c(1,2,3,4,5)),itsd=c(1,0,0))
```

```
    --------- ------ - ------ ----
      Augmented Dickey & Fuller test
      --------- ------ - ------ ----

      Null hypothesis: Unit root.
      Alternative hypothesis: Stationarity.
```

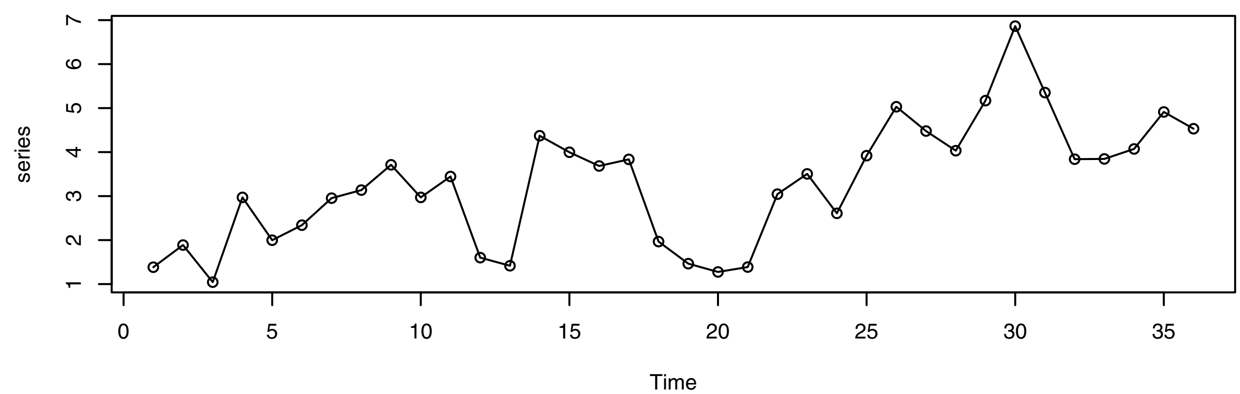
```
----
  ADF statistic:

        Estimate Std. Error t value Pr(>|t|)
adf.reg   -2.678      0.869  -3.083    0.036

  Lag orders: 1 2 3 4 5
  Number of available observations: 54
```

We (correctly) reject the unit root at the usual significance levels.

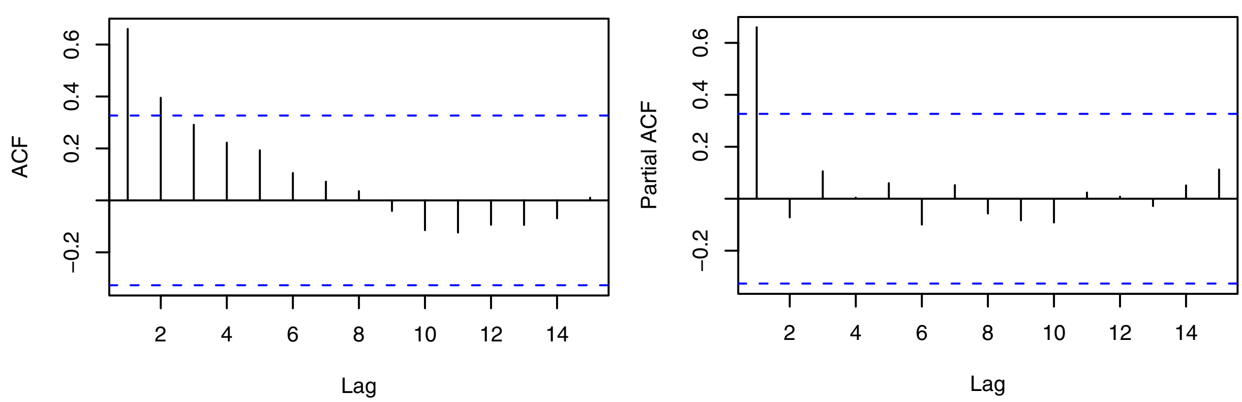
Exercise 6.32 Simulate a stationary time series of length $n = 36$ according to an AR(1) model with $\phi = 0.95$. This model is stationary, but just barely so. With such a series and a short history, it will be difficult if not impossible to distinguish between stationary and nonstationary with a unit root.

```
> set.seed(274135); series=arima.sim(n=36,list(ar=0.95))
```

(a) Plot the series and calculate the sample ACF and PACF and describe what you see.




```
> win.graph(width=6.5, height=2.5,pointsize=8)
> plot(series,type='o')
```




```
> win.graph(width=3.25, height=2.5,pointsize=8); acf(series); pacf(series)
```

The ACF and PACF graphs would lead us to at least entertain an AR(1) model. However, the "upward trend" in the time series plot suggests nonstationarity of some kind.

(b) Perform the (augmented) Dickey-Fuller test on the series with $k = 0$ in Equation (6.4.1), page 128. (With $k = 0$ this is the Dickey-Fuller test and is not augmented.) Comment on the results.

```
> library(uroot); ADF.test(series, selectlags=list(Pmax=0), itsd=c(1,0,0))
```

```
  --------- ------ - ------ ----
  Augmented Dickey & Fuller test
```

```
        --------- ------ - ------ ----

        Null hypothesis: Unit root.
        Alternative hypothesis: Stationarity.

      ----
        ADF statistic:

              Estimate Std. Error t value Pr(>|t|)
      adf.reg   -0.323      0.124  -2.597      0.1

        Lag orders: 0
        Number of available observations: 35
```

The Dickey-Fuller test results suggest that we consider a nonstationary model for these data.

(c) Perform the augmented Dickey-Fuller test on the series with k chosen by the software—that is, the "best" value for k. Comment on the results.

```
> ar(diff(series)) # order 2 is selected as "best" for the differenced series
> ADF.test(series, selectlags=list(mode=c(1,2)), itsd=c(1,0,0))
```

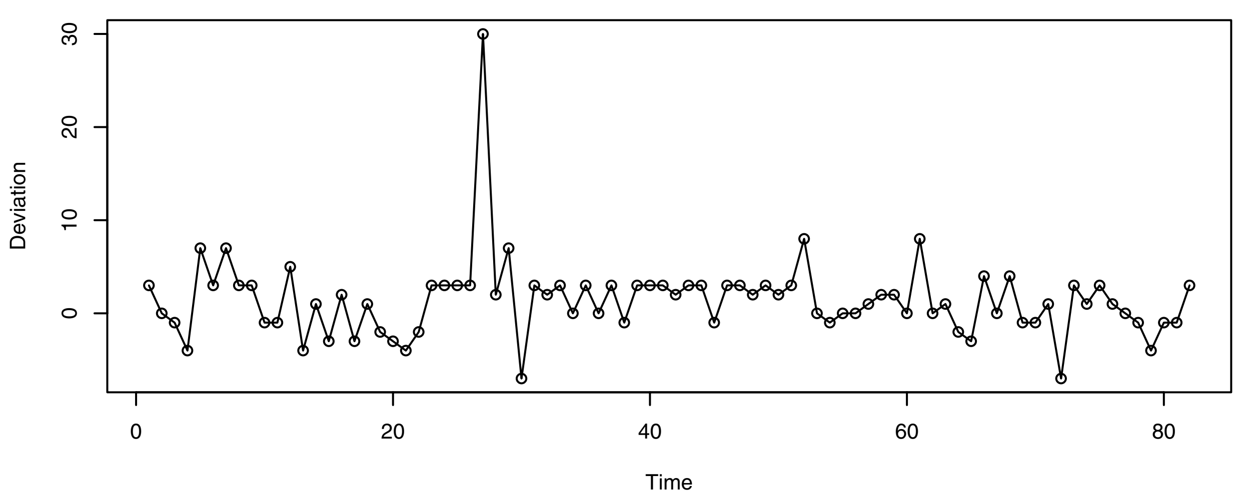
```
        --------- ------ - ------ ----
        Augmented Dickey & Fuller test
        --------- ------ - ------ ----

        Null hypothesis: Unit root.
        Alternative hypothesis: Stationarity.

      ----
        ADF statistic:

              Estimate Std. Error t value Pr(>|t|)
      adf.reg   -0.335      0.155  -2.156      0.1

        Lag orders: 1 2
        Number of available observations: 33
```

The augmented Dickey-Fuller test, also suggests a unit root for this series.

(d) Repeat parts (a), (b), and (c) but with a new simulation with n = 100.

Exercise 6.33 The data file named deere1 contains 82 consecutive values for the amount of deviation (in 0.000025 inch units) from a specified target value that an industrial machining process at Deere & Co. produced under certain specified operating conditions.

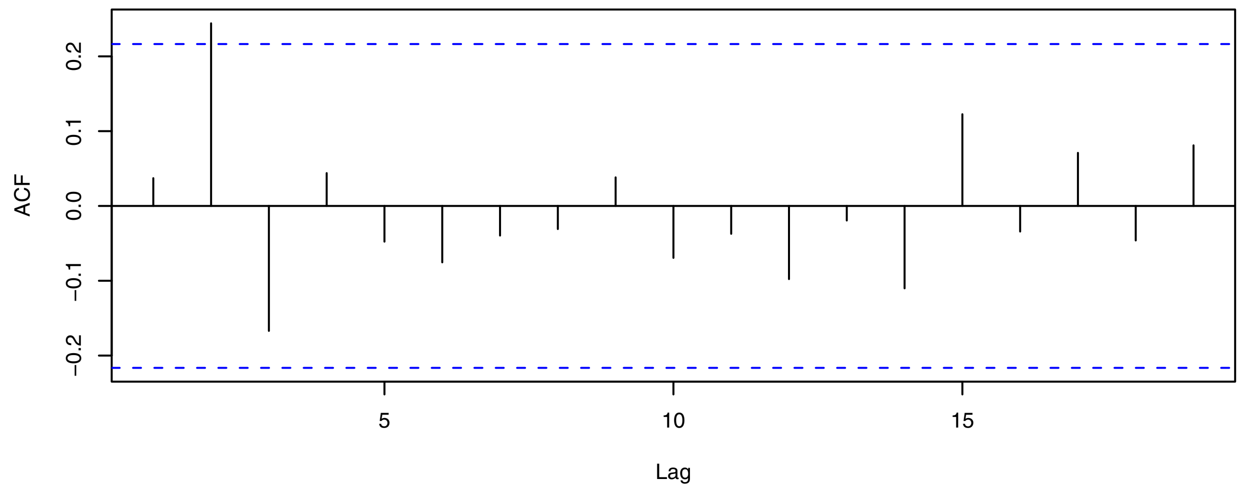
(a) Display the time series plot of this series and comment on any unusual points.




```
> data(deere1); plot(deere1,type='o',ylab='Deviation')
```

Except for one point of 30 at $t = 27$ the process seems relatively stable and stationary.

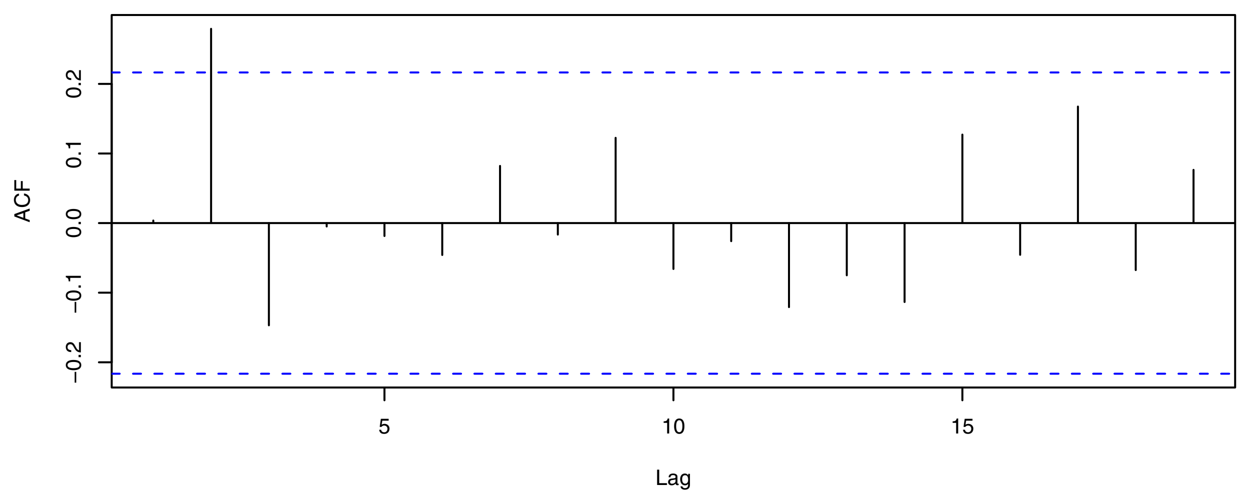
(b) Calculate the sample ACF for this series and comment on the results.




```
> acf(deere1)
```

The graph indicates a statistically significant autocorrelation at lag 2.

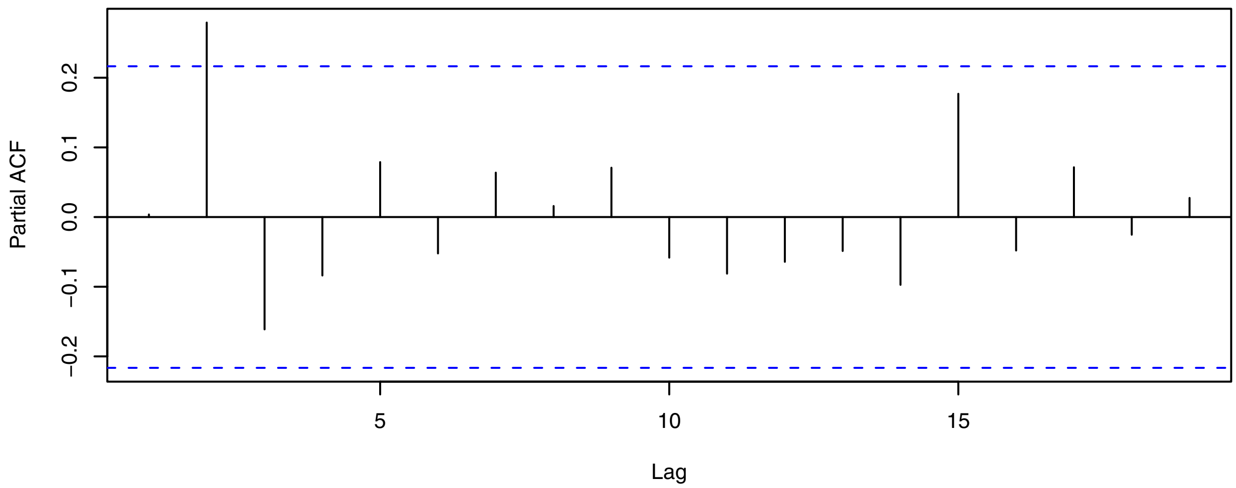
(c) Now replace the unusual value by a much more typical value and recalculate the sample ACF. Comment on the change from what you saw in part (b).




```
> deere1[27]=8; acf(deere1)
```

We replaced the unusual value of 30 at time 27 with the next largest value of 8. This had only a small effect on the sample autocorrelation function.

(d) Calculate the sample PACF based on the revised series that you used in part (c). What model would you specify for the revised series? (Later we will investigate other ways to handle outliers in time series modeling.)

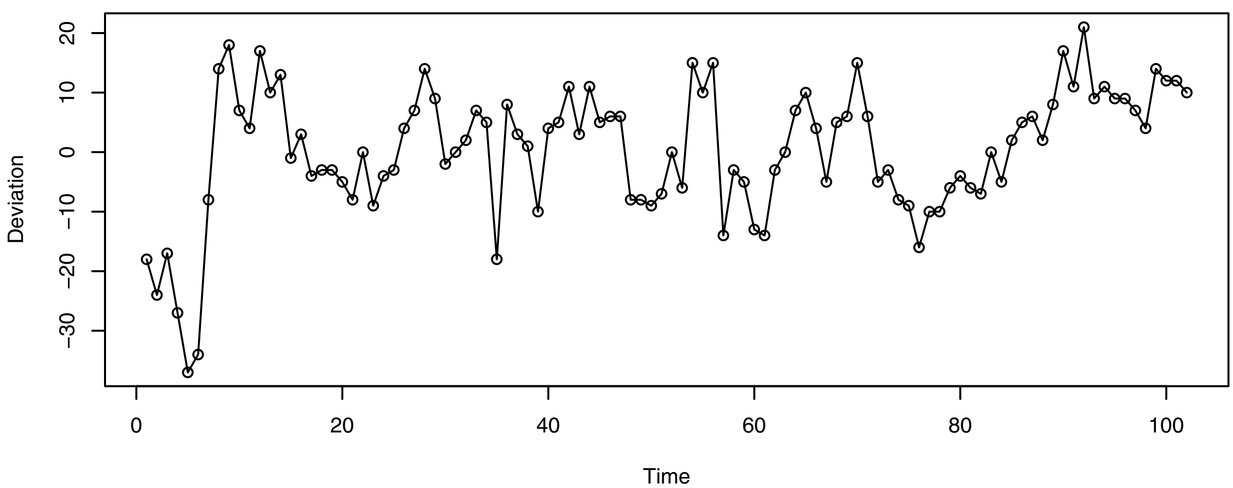



```
> pacf(deere1)
```

This pacf suggests an AR(2) model for the series.

Exercise 6.34 The data file named deere2 contains 102 consecutive values for the amount of deviation (in 0.0000025 inch units) from a specified target value that another industrial machining process produced at Deere & Co.

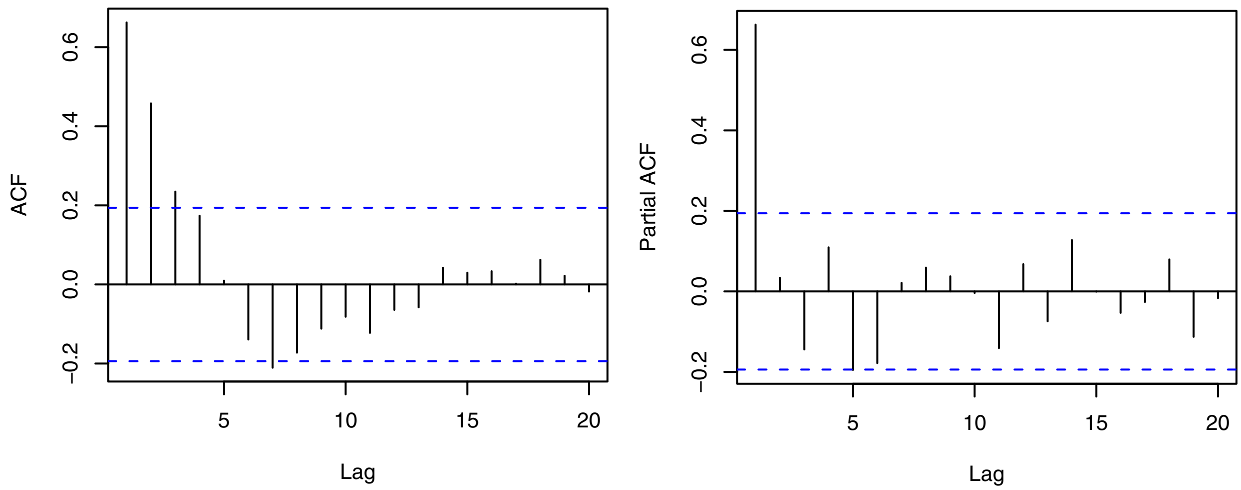
(a) Display the time series plot of this series and comment on its appearance. Would a stationary model seem to be appropriate?




```
> data(deere2); plot(deere2,type='o',ylab='Deviation')
```

There are some unusual observations at the beginning of this series. These are possibly some startup effects. After the startup, the series might well be considered stationary.

(b) Display the sample ACF and PACF for this series and select tentative orders for an ARMA model for the series.

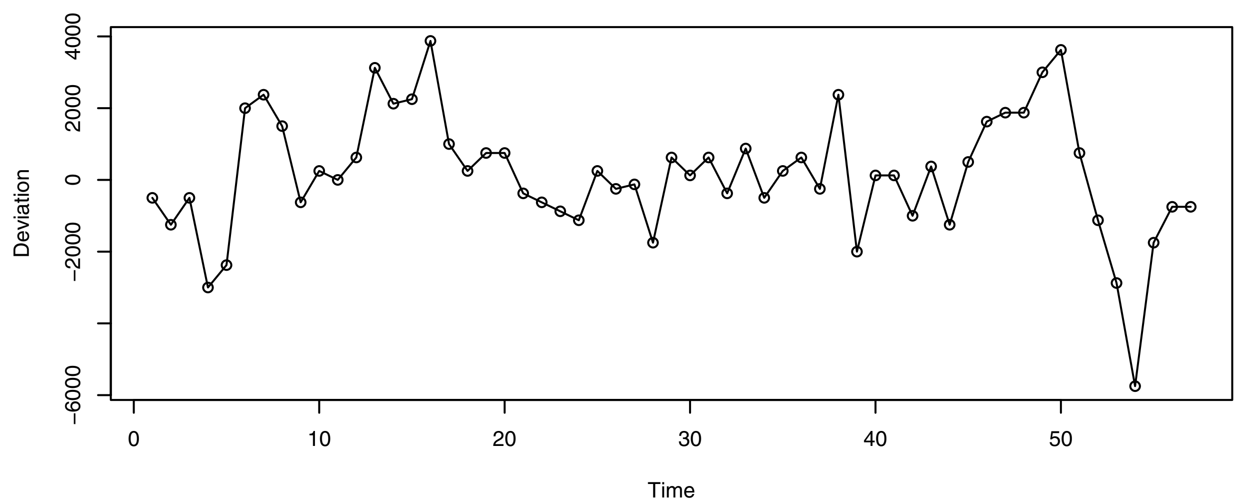



```
> win.graph(width=3.25,height=3,pointsize=8); acf(deere2); pacf(deere2)
```

These plots strongly suggest an AR(1) model for this series. An AR(2) might also be tried as an overfit (See Chapter 8) due to the damped sine wave suggested in the acf (but the pacf does not support this model).

Exercise 6.35 The data file named deere3 contains 57 consecutive measurements recorded from a complex machine tool at Deere & Co. The values given are deviations from a target value in units of ten millionths of an inch. The process employs a control mechanism that resets some of the parameters of the machine tool depending on the magnitude of deviation from target of the last item produced.

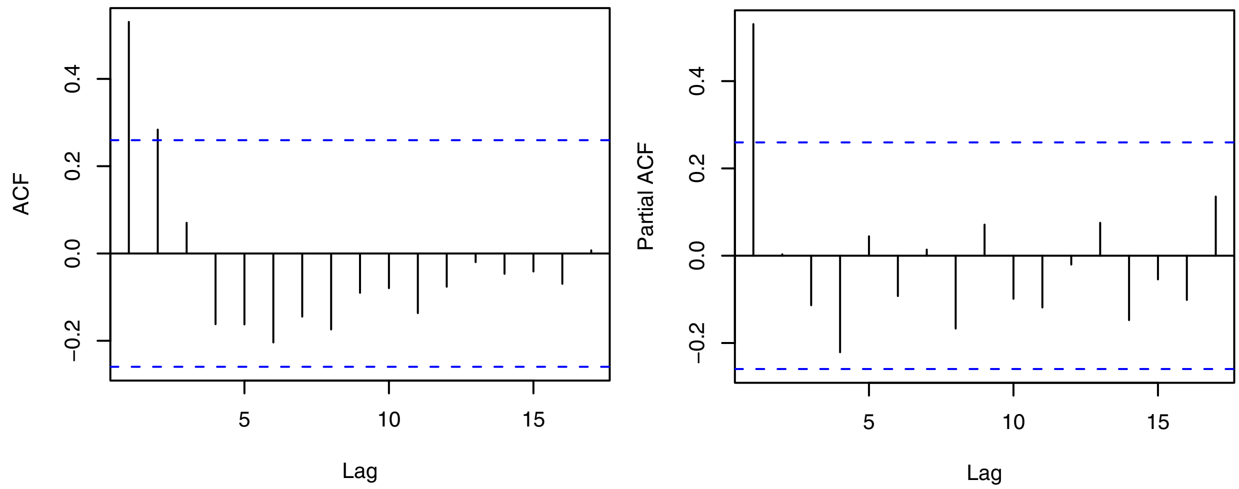
(a) Display the time series plot of this series and comment on its appearance. Would a stationary model be appropriate here?




```
> data(deere3); plot(deere3,type='o',ylab='Deviation')
```

This plot looks reasonably stationary with the possible exception of the last few observations.

(b) Display the sample ACF and PACF for this series and select tentative orders for an ARMA model for the series.

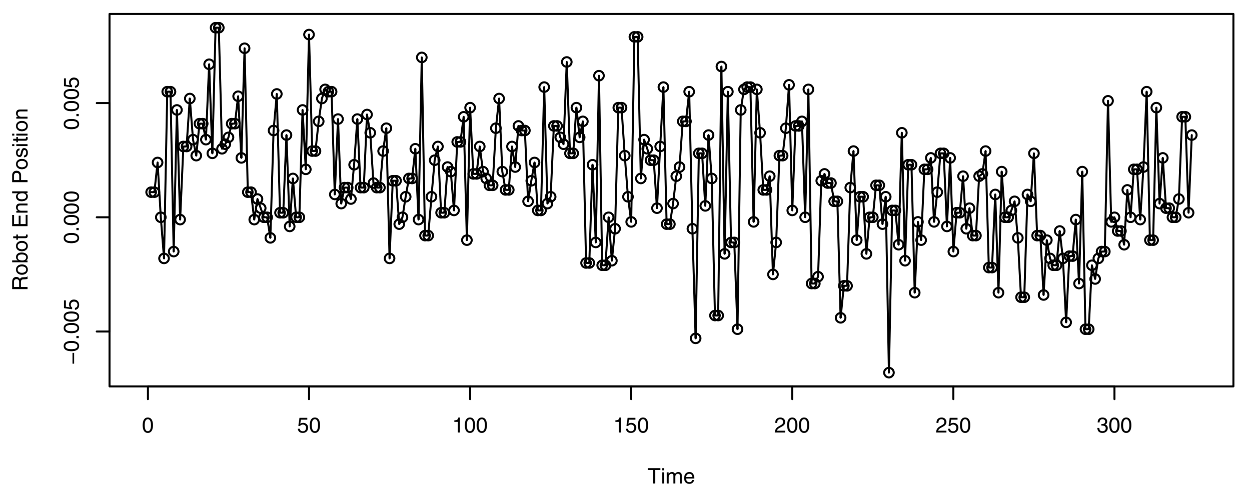



```
> win.graph(width=3.25,height=3,pointsize=8); acf(deere3); pacf(deere3)
```

Based on these displays, we would tentatively specify an AR(1) model for this series.

Exercise 6.36 The data file named robot contains a time series obtained from an industrial robot. The robot was put through a sequence of maneuvers, and the distance from a desired ending point was recorded in inches. This was repeated 324 times to form the time series.

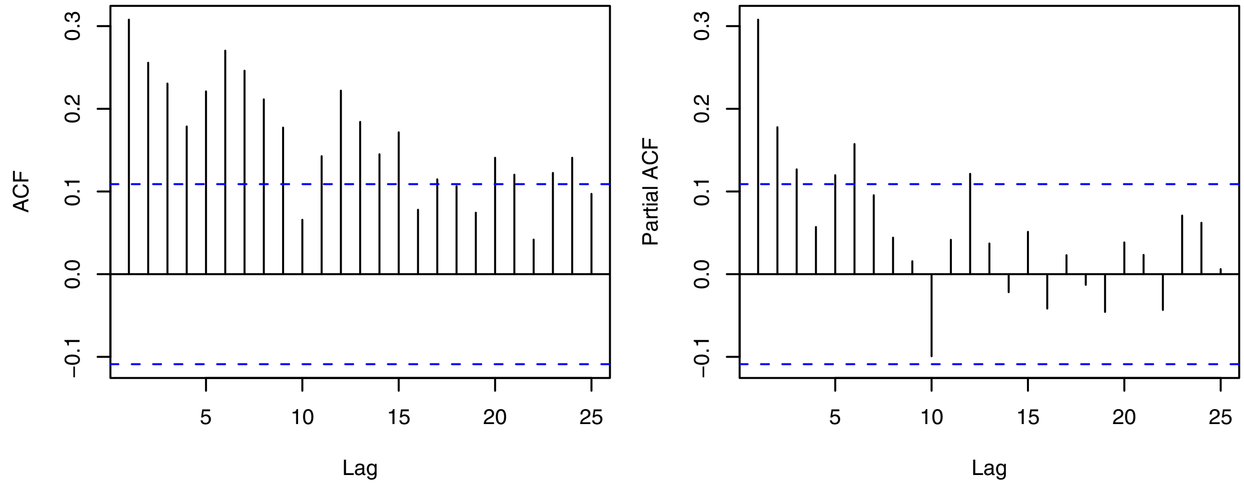
(a) Display the time series plot of the data. Based on this information, do these data appear to come from a stationary or nonstationary process?




```
> data(robot); plot(robot,type='o',ylab='Robot End Position')
```

From this plot we might try a stationary model but there is also enough "drift" that we might also suspect nonstationarity.

(b) Calculate and plot the sample ACF and PACF for these data. Based on this additional information, do these data appear to come from a stationary or nonstationary process?




```
> win.graph(width=3.25,height=3,pointsize=8);acf(robot);pacf(robot)
```

These plots are are not especially definitive, but the pacf suggests possibly an AR(3) model for the series.

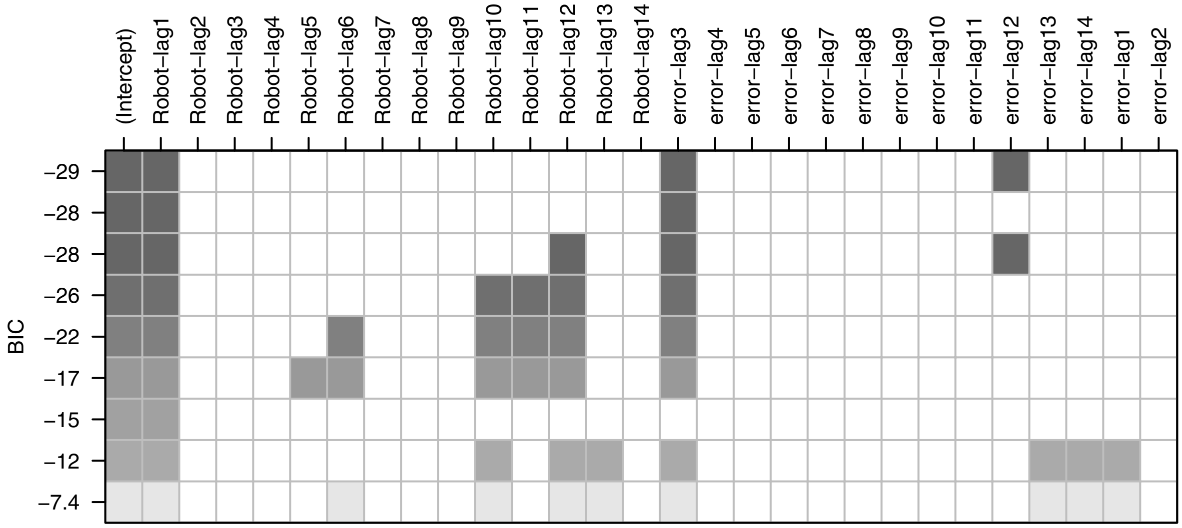
(c) Calculate and interpret the sample EACF.

AR\MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	x	x	x	x	x	x	x	x	o	x	x	x	x
1	x	o	o	o	o	o	o	o	o	o	o	o	o	o
2	x	x	o	o	o	o	o	o	o	o	o	o	o	o
3	x	x	o	o	o	o	o	o	o	o	o	o	o	o
4	x	x	x	x	o	o	o	o	o	o	o	o	x	o
5	x	x	x	o	o	o	o	o	o	o	o	o	x	o
6	x	o	o	o	o	x	o	o	o	o	o	o	o	o
7	x	o	o	x	o	x	x	o	o	o	o	o	o	o

```
> eacf(robot)
```

The EACF suggests an ARMA(1,1) model.

(d) Use the best subsets ARMA approach to specify a model for these data. Compare these results with what you discovered in parts (a), (b), and (c).




```
> plot(armasubsets(y=robot,nar=14,nma=14,y.name='Robot',ar.method='ols'))
```

The best model here includes a lag 1 AR term but lags 3 and 12 in the MA part of the model.

Exercise 6.37 Calculate and interpret the sample EACF for the logarithms of the Los Angeles rainfall series. The data are in the file named larain. Do the results confirm that the logs are white noise?

AR\MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	o	o	o	o	o	o	o	o	o	o	o	o	o	o
1	o	o	o	o	o	o	o	o	o	o	o	o	o	o
2	x	x	o	o	o	o	o	o	o	o	o	o	o	o
3	x	o	o	o	o	o	o	o	o	o	o	o	o	o
4	x	o	o	o	o	o	o	o	o	o	o	o	o	o
5	x	x	x	x	x	o	o	o	o	o	o	o	o	o
6	x	x	o	o	x	o	o	o	o	o	o	o	o	o
7	x	o	x	o	o	o	o	o	o	o	o	o	o	o

```
> eacf(log(larain))
```

This EACF suggests that the logarithms of the L.A. annual rainfall follow a white noise process.

Exercise 6.38 Calculate and interpret the sample EACF for the color property time series. The data are in the color file. Does the sample EACF suggest the same model that was specified by looking at the sample PACF?

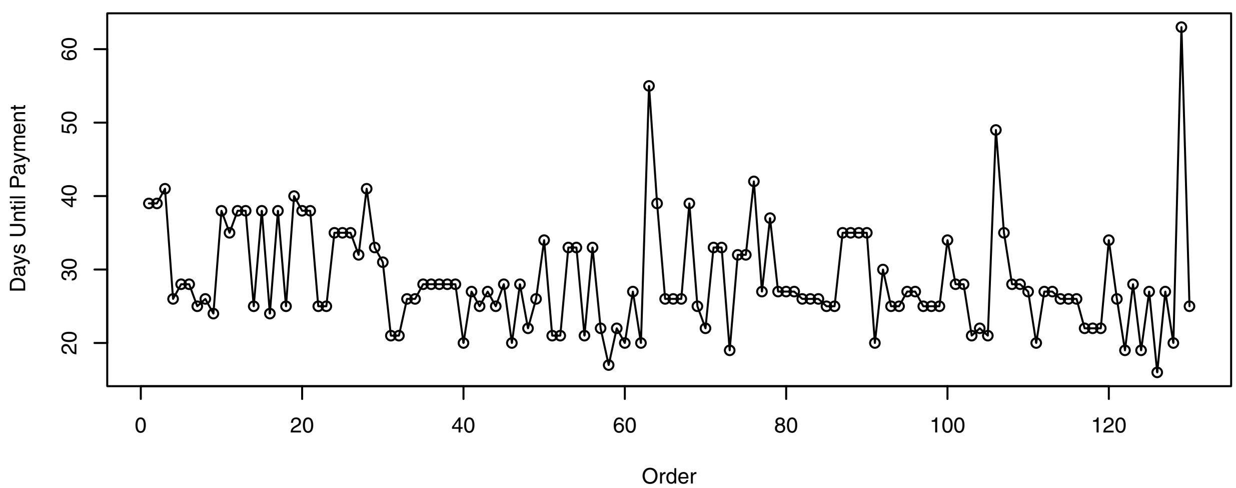
AR\MA	0	1	2	3	4	5	6	7	8	9
0	x	o	o	o	o	o	o	o	o	o
1	o	o	o	o	o	o	o	o	o	o
2	o	o	o	o	o	o	o	o	o	o
3	x	o	o	o	o	o	o	o	o	o
4	o	o	o	o	o	o	o	o	o	o
5	x	o	o	o	o	o	o	o	o	o
6	x	o	o	o	o	o	o	o	o	o
7	x	o	o	o	o	o	o	o	o	o

```
> eacf(color,ar.max=7,ma.max=9)
```

This EACF supports an AR(1) model for this series.

Exercise 6.39 The data file named days contains accounting data from the Winegard Co. of Burlington, Iowa. The data are the number of days until Winegard receives payment for 130 consecutive orders from a particular distributor of Winegard products. (The name of the distributor must remain anonymous for confidentiality reasons.)

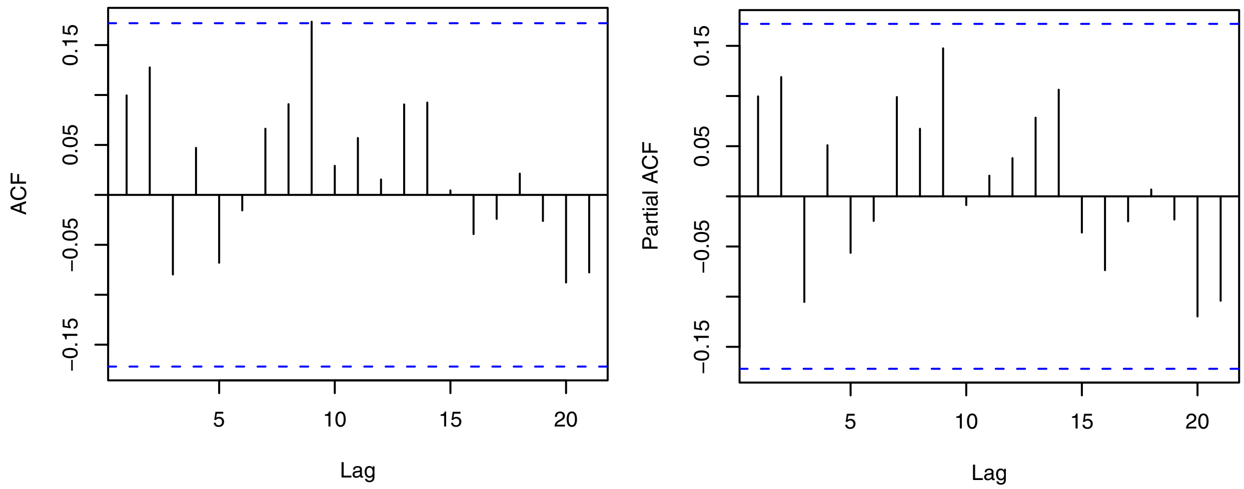
(a) Plot the time series, and comment on the display. Are there any unusual values?




```
> plot(days,type='o',ylab='Days Until Payment',xlab='Order')
```

The values of 55 at order number 63, 49 at order number 106, and 63 at order 129 look rather unusual relative the other values.

(b) Calculate the sample ACF and PACF for this series.

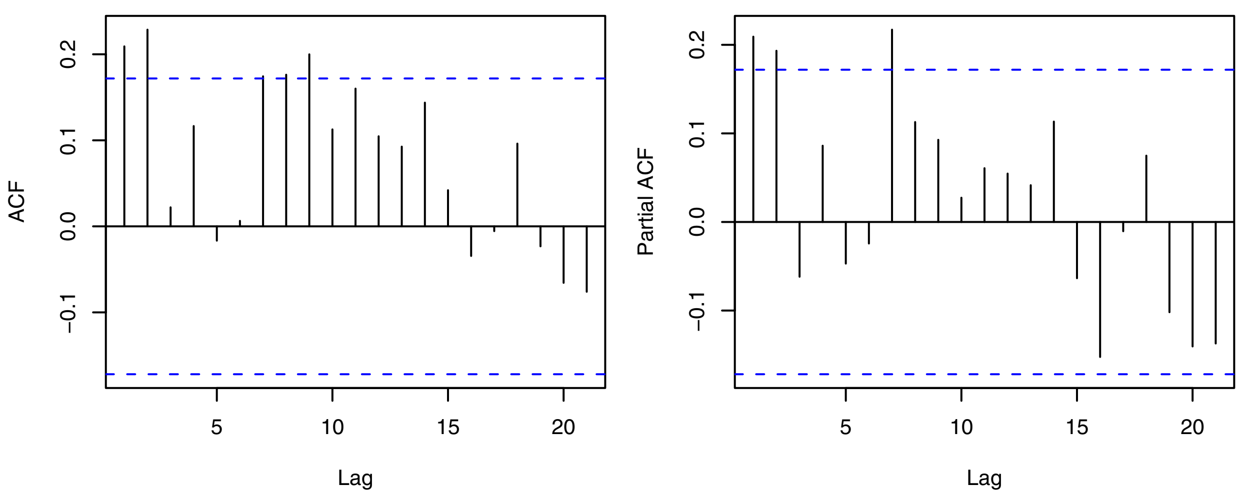



```
> win.graph(width=3.25,height=3,pointsize=8); acf(days); pacf(days)
```

None of the autocorrelations or partial autocorrelations are statistically significantly different from zero. The series looks like white noise.

(c) Now replace each of the unusual values with a value of 35 days—much more typical values—and repeat the calculation of the sample ACF and PACF. What ARMA model would you specify for this series after removing the outliers? (Later we will investigate other ways to handle outliers in time series modeling.)

```
> daysmod=days; daysmod[63]=35; daysmod[106]=35; daysmod[129]=35
> acf(daysmod); pacf(daysmod)
```



After replacing the outliers, we see several significant auto and partial correlations. No clear-cut pattern has emerged, but would certainly want to try both MA(2) and AR(2) and see how these models might fit.

CHAPTER 7

Exercise 7.1 From a series of length 100, we have computed $r_1 = 0.8$, $r_2 = 0.5$, $r_3 = 0.4$, $\bar{Y} = 2$, and a sample variance of 5. If we assume that an AR(2) model with a constant term is appropriate, how can we get (simple) estimates of ϕ_1, ϕ_2, θ_0, and σ_e^2 ?

Using $\hat{\phi}_1 = \dfrac{r_1(1-r_2)}{1-r_1^2}$ and $\hat{\phi}_2 = \dfrac{r_2-r_1^2}{1-r_1^2}$, we have $\hat{\phi}_1 = \dfrac{0.8(1-0.5)}{1-0.8^2} = 1.11$ and $\hat{\phi}_2 = \dfrac{0.5-0.8^2}{1-0.8^2} = -0.389$.

Then from $\theta_0 = \mu(1-\phi_1-\phi_2)$ we have $\hat{\theta}_0 = \hat{\mu}(1-\hat{\phi}_1-\hat{\phi}_2) = 2(1-1.11-(-0.389)) = 0.558$. Finally, from $\hat{\sigma}_e^2 = (1-\hat{\phi}_1 r_1-\hat{\phi}_2 r_2)s^2 = [1-(1.11)(0.8)-(-0.389)(0.5)]5 = 1.5325$.

Exercise 7.2 Assuming that the following data arise from a stationary process, calculate method-of-moments estimates of μ, γ_0, and ρ_1: 6, 5, 4, 6, 4.

$\hat{\mu} = (6+5+4+6+4)/5 = 5$. $\hat{\gamma}_0 = [(6-5)^2+(5-5)^2+(4-5)^2+(6-5)^2+(4-5)^2]/(5-1) = 1$ and $\hat{\rho}_1 = r_1 = [(6-5)(5-5)+(5-5)(4-5)+(4-5)(6-5)+(6-5)(4-5)]/4 = -1/2$

Exercise 7.3 If $\{Y_t\}$ satisfies an AR(1) model with ϕ of about 0.7, how long of a series do we need to estimate $\phi = \rho_1$ with 95% confidence that our estimation error is no more than ± 0.1?

The (large sample) standard error of $\hat{\phi} = r_1$ is $\sqrt{(1-\hat{\phi}^2)/n} = \sqrt{(1-0.7^2)/n} = \sqrt{0.51/n} \approx 0.7141/\sqrt{n}$. Solving $2(0.7141/\sqrt{n}) = 0.1$ yields $n = 204$.

Exercise 7.4 Consider an MA(1) process for which it is *known* that the process mean is zero. Based on a series of length $n = 3$, we observe $Y_1 = 0$, $Y_2 = -1$, and $Y_3 = ½$.

(a) Show that the conditional least-squares estimate of θ is ½.

Using Equation (7.2.14), page 157, we have $e_1 = Y_1 = 0$, $e_2 = Y_2 + \theta e_1 = -1$, and $e_3 = Y_3 + \theta e_2 = ½ + (-1)\theta$

So $S_c(\theta) = \sum(e_t)^2 = 0^2 + (-1)^2 + (1/2-\theta)^2$ and, by inspection, this is minimized when $\theta = ½$.

(b) Find an estimate of the noise variance. (Hint: Iterative methods are not needed in this simple case.)

$\hat{\sigma}_e^2 = \dfrac{S_c(\theta)}{n-1} = \dfrac{\sum(e_t)^2}{n-1} = \dfrac{1}{3-1} = \dfrac{1}{2}$. Note: Since the mean is *known* to be zero, one might reasonably argue that the divisor should be n rather than $n-1$.

Exercise 7.5 Given the data $Y_1 = 10$, $Y_2 = 9$, and $Y_3 = 9.5$, we wish to fit an IMA(1,1) model without a constant term.

(a) Find the conditional least squares estimate of θ. (Hint: Do Exercise 7.4 first.)

After computing first differences, we have data just like in Exercise 7.4. Since fitting an IMA(1,1) model with no constant term to the original data is equivalent to fitting an MA(1) with zero mean to these differences, the answers will be the same as in the previous exercise.

(b) Estimate σ_e^2.

Exercise 7.6 Consider two different parameterizations of the AR(1) process with nonzero mean:

Model I. $Y_t - \mu = \phi(Y_{t-1} - \mu) + e_t$.

Model II. $Y_t = \phi Y_{t-1} + \theta_0 + e_t$.

We want to estimate ϕ and μ or ϕ and θ_0 using conditional least squares conditional on Y_1. Show that with Model I we are led to solve nonlinear equations to obtain the estimates, while with Model II we need only solve linear equations.

Rewriting Model I as $Y_t = \mu(1-\phi) + Y_{t-1} + e_t$ we see that it is not linear in the parameters μ and ϕ. Thus, the equations obtained from setting the partial derivatives of $\sum(e_t)^2$ to zero will not be linear equations. Model II is linear in the parameters and ordinary least squares regression may be used to obtain the conditional least squares estimates.

Exercise 7.7 Verify Equation (7.1.4), page 150.

Rewriting $r_1 = -\frac{\theta}{1+\theta^2}$ we have the quadratic equation in θ $r_1\theta^2 + \theta + r_1 = 0$. The two solutions are $-\frac{1}{2r_1} \pm \sqrt{\frac{1}{4r_1^2} - 1}$. We claim that $\hat{\theta} = \frac{-1 + \sqrt{1 - 4r_1^2}}{2r_1}$ is always the invertible one. We consider four cases:

Case 1: $0 < r_1 < 0.5$. Then $\hat{\theta} < 1 \Leftrightarrow \frac{-1 + \sqrt{1 - 4r_1^2}}{2r_1} < 1 \Leftrightarrow 1 - 4r_1^2 < 4r_1^2 + 4r_1 + 1$ which is clearly true.

Case 2: $0 < r_1 < 0.5$. Then $-1 < \hat{\theta} \Leftrightarrow -1 < \frac{-1 + \sqrt{1 - 4r_1^2}}{2r_1} \Leftrightarrow -2r_1 < -1 + \sqrt{1 - 4r_1^2} \Leftrightarrow 1 - 2r_1 < \sqrt{1 - 4r_1^2} \Leftrightarrow$

$1 - 4r_1 + 4r_1^2 < 1 - 4r_1^2 \Leftrightarrow -1 + r_1 < r_1 \Leftrightarrow r_1 < 0.5$ which is true.

Case 3: $-0.5 < r_1 < 0$. Then $\hat{\theta} < 1 \Leftrightarrow \frac{-1 + \sqrt{1 - 4r_1^2}}{2r_1} < 1 \Leftrightarrow \sqrt{1 - 4r_1^2} > 2r_1 + 1 \Leftrightarrow 1 - 4r_1^2 > 4r_1^2 + 4r_1 + 1 \Leftrightarrow$

$-r_1 < r_1 + 1 \Leftrightarrow -0.5 < r_1$ which is true.

Case 4: $-0.5 < r_1 < 0$. Then $-1 < \hat{\theta} \Leftrightarrow -1 < \frac{-1 + \sqrt{1 - 4r_1^2}}{2r_1} \Leftrightarrow -2r_1 > -1 + \sqrt{1 - 4r_1^2} \Leftrightarrow 1 - 2r_1 > \sqrt{1 - 4r_1^2} \Leftrightarrow$

$1 - 4r_1 + 4r_1^2 > 1 - 4r_1^2 \Leftrightarrow -1 + r_1 < -r_1 \Leftrightarrow r_1 < 0.5$ which is true.

Exercise 7.8 Consider an ARMA(1,1) model with $\phi = 0.5$ and $\theta = 0.45$.

(a) For $n = 48$, evaluate the variances and correlation of the maximum likelihood estimators of ϕ and θ using Equations (7.4.13) on page 161. Comment on the results.

$Var(\hat{\phi}) \approx \left[\frac{1-\phi^2}{n}\right]\left[\frac{1-\phi\theta}{\phi-\theta}\right]^2 = \left[\frac{1-0.5^2}{48}\right]\left[\frac{1-(0.5)(0.45)}{0.5-0.45}\right]^2 = 3.75$ and $\sqrt{3.75} = 1.94$

$Var(\hat{\theta}) \approx \left[\frac{1-\theta^2}{n}\right]\left[\frac{1-\phi\theta}{\phi-\theta}\right]^2 = \left[\frac{1-0.45^2}{48}\right]\left[\frac{1-(0.5)(0.45)}{0.5-0.45}\right]^2 = 3.99$ and $\sqrt{3.99} = 1.997$. The standard errors are quite large relative to the quantities being estimated. This is because of the near cancellation of the AR and MA parameters. This is a rather unstable model approaching white noise.

Furthermore, $Corr(\hat{\phi}, \hat{\theta}) \approx \frac{\sqrt{(1-\phi^2)(1-\theta^2)}}{1-\phi\theta} = \frac{\sqrt{(1-0.5^2)(1-0.45^2)}}{1-(0.5)(0.45)} = 0.998$. The estimates are very highly correlated.

(b) Repeat part (a) but now with $n = 120$. Comment on the new results.

$Var(\hat{\phi}) \approx \left[\frac{1-0.5^2}{120}\right]\left[\frac{1-(0.5)(0.45)}{0.5-0.45}\right]^2 = 1.50$ and $\sqrt{1.50} = 1.22$.

$Var(\hat{\theta}) \approx \left[\frac{1-0.45^2}{120}\right]\left[\frac{1-(0.5)(0.45)}{0.5-0.45}\right]^2 = 1.60$ and $\sqrt{1.60} = 1.26$. The standard errors are only slightly smaller.

$Corr(\hat{\phi}, \hat{\theta}) \approx 0.998$ does not change with n!

Exercise 7.9 Simulate an MA(1) series with $\theta = 0.8$ and $n = 48$.

```
> set.seed(15234); series=arima.sim(n=48,list(ma=-0.8))
```

(a) Find the method-of-moments estimate of θ.

```
> # Below is a function that computes the method of moments estimator of
> # the MA(1) coefficient of an MA(1) model.
> estimate.ma1.mom=function(x){r=acf(x,plot=F)$acf[1]; if (abs(r)<0.5)
> return((-1+sqrt(1-4*r^2))/(2*r)) else return(NA)}
> estimate.ma1.mom(series)
```

```
theta hat = 0.685599
```

(b) Find the conditional least squares estimate of θ and compare it with part (a).

```
> arima(series,order=c(0,0,1),method='CSS')
```

```
Call:
arima(x = series, order = c(0, 0, 1), method = "CSS")

Coefficients:
          ma1  intercept
      -0.7529     0.0109
s.e.   0.1089     0.0422

sigma^2 estimated as 1.247:  part log likelihood = -73.41
```

With our sign convention, the estimate of theta is +0.7529. This estimate is closer to the truth than the method-of-moments estimate in part (a).

(c) Find the maximum likelihood estimate of θ and compare it with parts (a) and (b).

```
> arima(series,order=c(0,0,1),method='ML')
```

```
Call:
arima(x = series, order = c(0, 0, 1), method = "ML")

Coefficients:
          ma1  intercept
      -0.7700     0.0033
s.e.   0.1183     0.0405

sigma^2 estimated as 1.226:  log likelihood = -73.45,  aic = 150.9
```

With our sign convention, the estimate of theta is +0.7700 and this is the best of the three estimates for this particular simulation.

(d) Repeat parts (a), (b), and (c) with a new simulated series using the same parameters and same sample size. Compare your results with your results from the first simulation.

Exercise 7.10 Simulate an MA(1) series with θ = –0.6 and n = 36.

```
> set.seed(135246); series=arima.sim(n=36,list(ma=0.6))
```

(a) Find the method-of-moments estimate of θ.

```
> estimate.ma1.mom(series)
```

```
theta hat = -0.7510313
```

(b) Find the conditional least squares estimate of θ and compare it with part (a).

```
> arima(series,order=c(0,0,1),method='CSS')
```

```
Call:
arima(x = series, order = c(0, 0, 1), method = "CSS")

Coefficients:
         ma1  intercept
      0.9227     0.3786
s.e.  0.0618     0.2821

sigma^2 estimated as 0.8075:  part log likelihood = -47.23
```

The conditional least squares estimate is –0.9227 which in this case is worse than the method-of-moments estimate of –0.7510313.

(c) Find the maximum likelihood estimate of θ and compare it with parts (a) and (b).

```
> arima(series,order=c(0,0,1),method='ML')
```

```
Call:
arima(x = series, order = c(0, 0, 1), method = "ML")

Coefficients:
         ma1  intercept
      0.9248     0.4390
s.e.  0.0922     0.2831

sigma^2 estimated as 0.7993:  log likelihood = -48.01,  aic = 100.03
```

The maximum likelihood estimate is quite similar to the CSS estimate and, again, the MOM estimate is better in this case.

(d) Repeat parts (a), (b), and (c) with a new simulated series using the same parameters and same sample size. Compare your results with your results from the first simulation.

Exercise 7.11 Simulate an MA(1) series with $\theta = -0.6$ and $n = 48$.

```
> set.seed(1352); series=arima.sim(n=48,list(ma=0.6))
```

(a) Find the maximum likelihood estimate of θ.

```
> arima(series,order=c(0,0,1),method='ML')$coef[1]
```

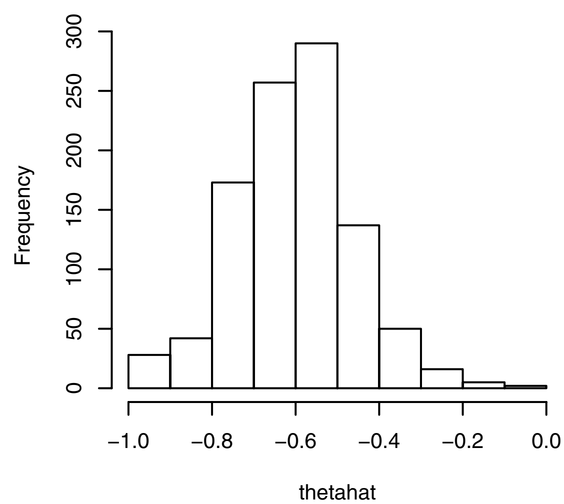
```
ma1
0.5081146
```

(Recall, that our sign convention is oposite that of the R software.)

(b) If your software permits, repeat part (a) many times with a new simulated series using the same parameters and same sample size.

```
> set.seed(1352); thetahat=rep(NA,1000)
> for (k in 1:1000) {series=arima.sim(n=48,list(ma=0.6));
>     thetahat[k]=-arima(series,order=c(0,0,1),method='ML')$coef[1]}
```

(c) Form the sampling distribution of the maximum likelihood estimates of θ.



The distribution is roughly normal but skewed a little toward zero.

(d) Are the estimates (approximately) unbiased?

```
> mean(thetahat)
```

The mean of the sampling distribution is –0.6042001 so the estimate is nearly unbiased.

(e) Calculate the variance of your sampling distribution and compare it with the large-sample result in Equation (7.4.11), page 161.

```
> sd(thetahat)^2
```

The sample variance from the 1000 replications is 0.02080749. Compare this to the large-sample result of

$Var(\hat{\theta}) \approx \frac{1-\theta^2}{n} = \frac{1-(-0.6)^2}{48} = 0.013$. Things are somewhat better in standard deviation terms. The comparison is 0.144 for the sampling distribution versus 0.115 from the large-sample theory.

Exercise 7.12 Repeat Exercise 7.11 using a sample size of n = 120.

Simulate an MA(1) series with $\theta = -0.6$ and n = 120.

```
> set.seed(1352); series=arima.sim(n=120,list(ma=0.6))
```

(a) Find the maximum likelihood estimate of θ.

```
> arima(series,order=c(0,0,1)) # Maximum likelihood is the default estimation method
```

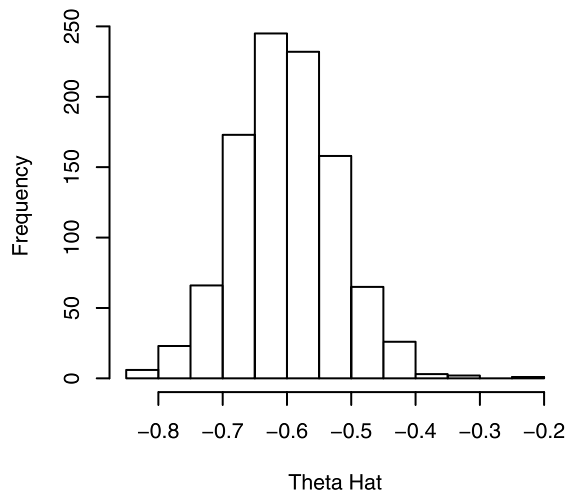
```
Call:
arima(x = series, order = c(0, 0, 1), method = "ML")

Coefficients:
         ma1  intercept
      0.5275    -0.0600
s.e.  0.0856     0.1367

sigma^2 estimated as 0.966:  log likelihood = -168.36,  aic = 340.72
```

Remembering the sign convention and noting the size of the standard error, we have an excellent estimate of θ in this simulation.

(b) If your software permits, repeat part (a) many times with a new simulated series using the same parameters and same sample size.

```
> set.seed(1352); thetahat=rep(NA,1000)
> for (k in 1:1000) {series=arima.sim(n=120,list(ma=0.6));
>              thetahat[k]=-arima(series,order=c(0,0,1))$coef[1]}
```

(c) Form the sampling distribution of the maximum likelihood estimates of θ.




```
> win.graph(width=3,height=3,pointsize=8)
> hist(thetahat,xlab='Theta Hat')
```

The histogram is roughly symmetric around $\theta = -0.6$.

(d) Are the estimates (approximately) unbiased?

```
> mean(thetahat)
```

```
[1] -0.6012115
```

The mean of the sampling distribution, -0.6012115, is very close to the true value of $\theta = -0.6$.

(e) Calculate the variance of your sampling distribution and compare it with the large-sample result in Equation (7.4.11), page 161.

```
> sd(thetahat)^2
```

```
[1] 0.006199142
```

$Var(\hat{\theta}) \approx (1-\theta^2)/n = (1-(-0.6)^2)/120 = 0.0053$ so the simulated result compares very favorably with the large-sample result.

Exercise 7.13 Simulate an AR(1) series with $\phi = 0.8$ and $n = 48$.

```
> set.seed(4321); series=arima.sim(n=48,list(ar=0.8))
```

(a) Find the method-of-moments estimate of ϕ.

```
> acf(series)$acf[1]
```

```
0.8285387
```

(a) Find the conditional least squares estimate of ϕ and compare it with part (a).

```
> arima(series,order=c(1,0,0),method='CSS')$coef[1]
```

```
ar1
0.8367125
```

For this AR(1) series the two estimates are quite close.

(b) Find the maximum likelihood estimate of ϕ and compare it with parts (a) and (b).

```
> arima(series,order=c(1,0,0),method='ML')$coef[1]
```

```
ar1
0.849501
```

All three methods produce very similar estimates for this AR(1) series.

(c) Repeat parts (a), (b), and (c) with a new simulated series using the same parameters and same sample size. Compare your results with your results from the first simulation.

Exercise 7.14 Simulate an AR(1) series with $\phi = -0.5$ and $n = 60$.

```
> set.seed(35431); series=arima.sim(n=60,list(ar=-0.5))
```

(a) Find the method-of-moments estimate of ϕ.

```
> acf(series)$acf[1]
```

```
-0.5469125
```

(b) Find the conditional least squares estimate of ϕ and compare it with part (a).

```
> arima(series,order=c(1,0,0),method='CSS')$coef[1]
```

```
ar1
-0.5596042
```

These two estimates are very close to each other and not far from the true value of –0.5.

(c) Find the maximum likelihood estimate of ϕ and compare it with parts (a) and (b).

```
> arima(series,order=c(1,0,0),method='ML')$coef[1]
```

```
ar1
-0.5580715
```

The MLE and CSS estimates are nearly the same.

(d) Repeat parts (a), (b), and (c) with a new simulated series using the same parameters and same sample size. Compare your results with your results from the first simulation.

Exercise 7.15 Simulate an AR(1) series with $\phi = 0.7$ and $n = 100$.

```
> set.seed(12352); series=arima.sim(n=100,list(ar=0.7))
```

(a) Find the maximum likelihood estimate of ϕ.

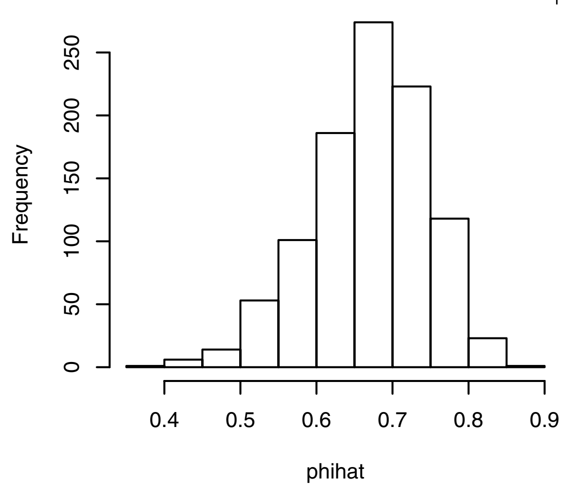
```
> arima(series,order=c(1,0,0),method='ML')$coef[1]
```

```
ar1
0.7025283
```

(b) If your software permits, repeat part (a) many times with a new simulated series using the same parameters and same sample size.

```
> for (k in 1:1000) {series=arima.sim(n=100,list(ar=0.7));
>               phihat[k]=arima(series,order=c(1,0,0),method='ML')$coef[1]}
```

(c) Form the sampling distribution of the maximum likelihood estimates of ϕ.



The distribution is roughly normal with a little skewness toward the lower values.

(d) Are the estimates (approximately) unbiased? Yes.

```
> mean(phihat)
```

```
0.6691908
```

(e) Calculate the variance of your sampling distribution and compare it with the large-sample result in Equation (7.4.9), page 161.

```
> sd(phihat)
```

The standard deviation of the simulated sampling distribution is 0.0774 whereas

$\sqrt{Var(\hat{\phi})} \approx \sqrt{\frac{1-\hat{\phi}^2}{n}} = \sqrt{\frac{1-(0.7)^2}{100}} = 0.0714$ and the large-sample result gives an excellent approximation.

Exercise 7.16 Simulate an AR(2) series with $\phi_1 = 0.6$, $\phi_2 = 0.3$, and $n = 60$.

```
> set.seed(12345); series=arima.sim(n=60,list(ar=c(0.6,0.3)))
```

(a) Find the method-of-moments estimates of ϕ_1 and ϕ_2.

```
> ar(series,aic=F,order.max=2,method='yw') # yw stands for Yule-Walker
```

```
Call:
ar(x = series, aic = F, order.max = 2, method = "yw")

Coefficients:
     1       2
0.5423  0.3161

Order selected 2  sigma^2 estimated as  1.366
```

(b) Find the conditional least squares estimates of ϕ_1 and ϕ_2 and compare them with part (a).

```
> ar(series,aic=F,order.max=2,method='ols') # ols stands for Ordinary Least Squares
```

```
Call:
ar(x = series, aic = F, order.max = 2, method = "ols")

Coefficients:
     1       2
0.5363  0.3288

Intercept: -0.05011 (0.1497)

Order selected 2  sigma^2 estimated as  1.299
```

The estimates are very similar to those obtained in parts (a) and (b).

(c) Find the maximum likelihood estimates of ϕ_1 and ϕ_2 and compare them with parts (a) and (b).

```
> ar(series,aic=F,order.max=2,method='mle')
> # mle stands for Maximum Likelihood Estimator
```

```
Call:
ar(x = series, aic = F, order.max = 2, method = "mle")

Coefficients:
     1       2
0.5330  0.3188

Order selected 2  sigma^2 estimated as  1.265
```

The estimates are very similar with all three methods.

(d) Repeat parts (a), (b), and (c) with a new simulated series using the same parameters and same sample size. Compare these results to your results from the first simulation.

Exercise 7.17 Simulate an ARMA(1,1) series with $\phi = 0.7$, $\theta = 0.4$, and $n = 72$.

```
> set.seed(54321); series=arima.sim(n=72,list(ar=0.7,ma=-0.4))
```

(a) Find the method-of-moments estimates of ϕ and θ.

```
> acf(series)$acf
```

```
[1,]  0.549397661
[2,]  0.388131613
```

So $\hat{\phi} = \frac{r_2}{r_1} = \frac{0.38813161}{0.54939766} = 0.70646754$. Recall that $r_1 = \frac{(1-\theta\hat{\phi})(\hat{\phi}-\theta)}{1-2\theta\hat{\phi}+\theta^2}$. The solutions for θ of the quadratic equation implied by this relationship are complex valued so that no method-of-moments estimate of θ exists for this series.

(b) Find the conditional least squares estimates of ϕ and θ and compare them with part (a).

```
> arima(series,order=c(1,0,1),method='CSS')
```

```
Call:
arima(x = series, order = c(1, 0, 1), method = "CSS")

Coefficients:
         ar1      ma1  intercept
      0.7655  -0.3605    -0.2444
s.e.  0.0961   0.1480     0.3075

sigma^2 estimated as 0.868:  part log likelihood = -97.07
```

The estimate of ϕ here is larger than the one obtained by the method-of-moments. However, taking standard errors into account, the two are not significantly different.

(c) Find the maximum likelihood estimates of ϕ and θ and compare them with parts (a) and (b).

```
> arima(series,order=c(1,0,1),method='ML')
```

```
Call:
arima(x = series, order = c(1, 0, 1), method = "ML")

Coefficients:
         ar1      ma1  intercept
      0.7771  -0.3055    -0.0201
s.e.  0.1190   0.1647     0.3410

sigma^2 estimated as 0.9147:  log likelihood = -99.2,  aic = 204.39
```

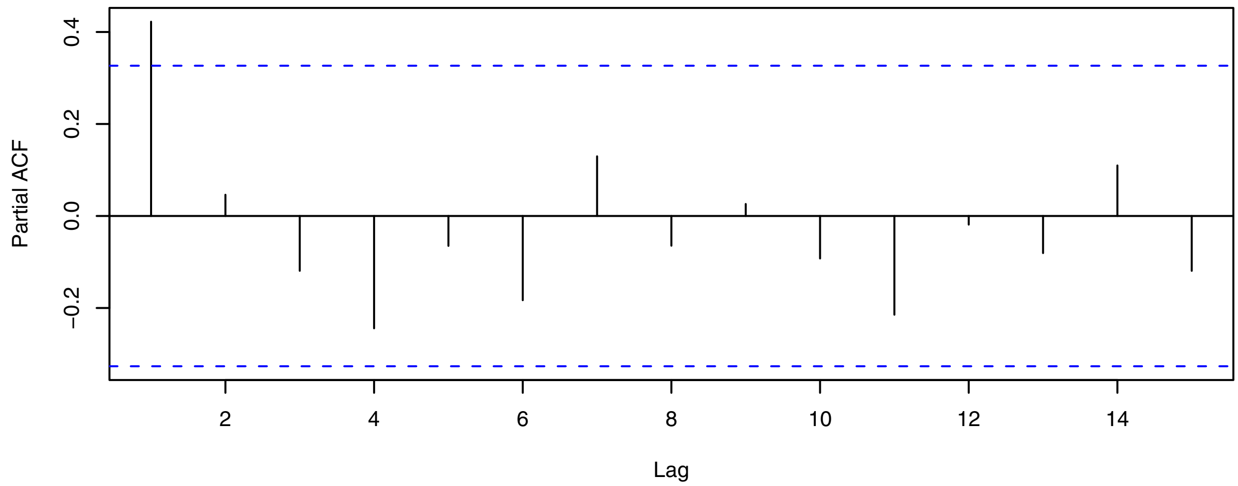
The CSS and ML estimates are very close to each other and easily within two standard errors of their true values.

(d) Repeat parts (a), (b), and (c) with a new simulated series using the same parameters and same sample size. Compare your new results with your results from the first simulation.

Exercise 7.18 Simulate an AR(1) series with $\phi = 0.6$, $n = 36$ but with error terms from a t-distribution with 3 degrees of freedom.

```
> set.seed(54321); series=arima.sim(n=36,list(ar=0.6),innov=rt(n,3))
```

(a) Display the sample PACF of the series. Is an AR(1) model suggested?




```
> pacf(series)
```

This PACF suggests an AR(1) model (in spite of the heavy tail errors).

(b) Estimate ϕ from the series and comment on the results.

```
> acf(series)$acf[1]; arima(series,order=c(1,0,0),method='ML')
```

```
Method of moments estimate of phi is 0.422392.

> arima(series,order=c(1,0,0),method='ML')

Call:
arima(x = series, order = c(1, 0, 0), method = "ML")

Coefficients:
         ar1  intercept
      0.4218     0.5067
s.e.  0.1496     0.3866

sigma^2 estimated as 1.871:  log likelihood = -62.46,  aic = 128.92
```

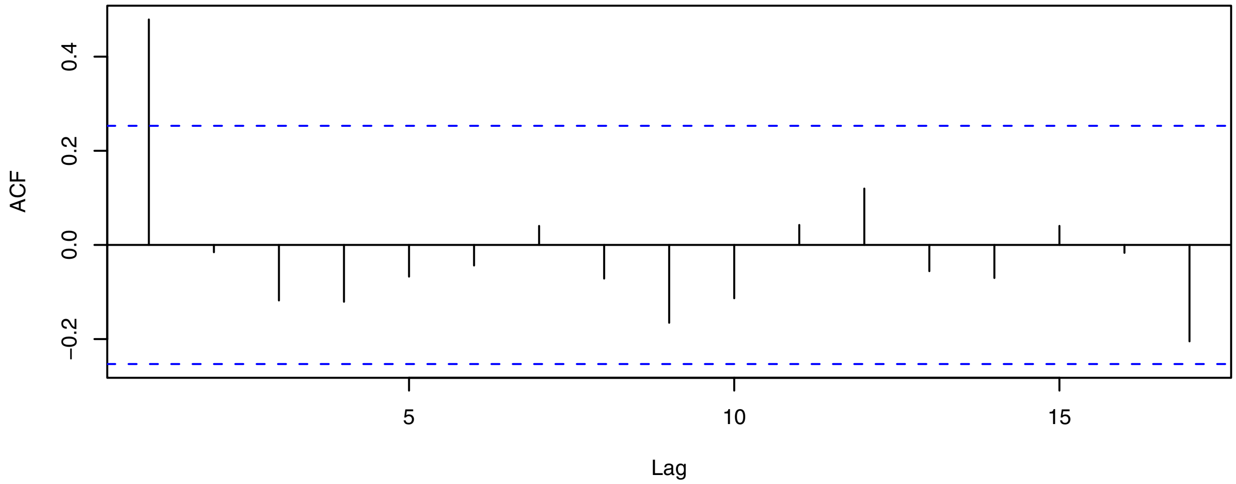
The mle estimation eroneously assumes *normal* errors but still estimates quite similarly to the method of moments.

(c) Repeat parts (a) and (b) with a new simulated series under the same conditions.

Exercise 7.19 Simulate an MA(1) series with $\theta = -0.8$, $n = 60$ but with error terms from a t-distribution with 4 degrees of freedom.

```
> set.seed(54321); series=arima.sim(n=60,list(ma=0.8),innov=rt(n,4))
```

(a) Display the sample ACF of the series. Is an MA(1) model suggested?



In spite of the heavy tailed errors, this sample acf strongly suggests an MA(1) model.

(b) Estimate θ from the series and comment on the results.

```
> estimate.ma1.mom(series); arima(series,order=c(0,0,1),method='ML')
```

```
Method of of moments estimate of theta = -0.7443132
Call:
arima(x = series, order = c(0, 0, 1), method = "ML")

Coefficients:
         ma1  intercept
      0.8170    -0.0471
s.e.  0.1166     0.2853

sigma^2 estimated as 1.500:  log likelihood = -97.84,  aic = 199.68
```

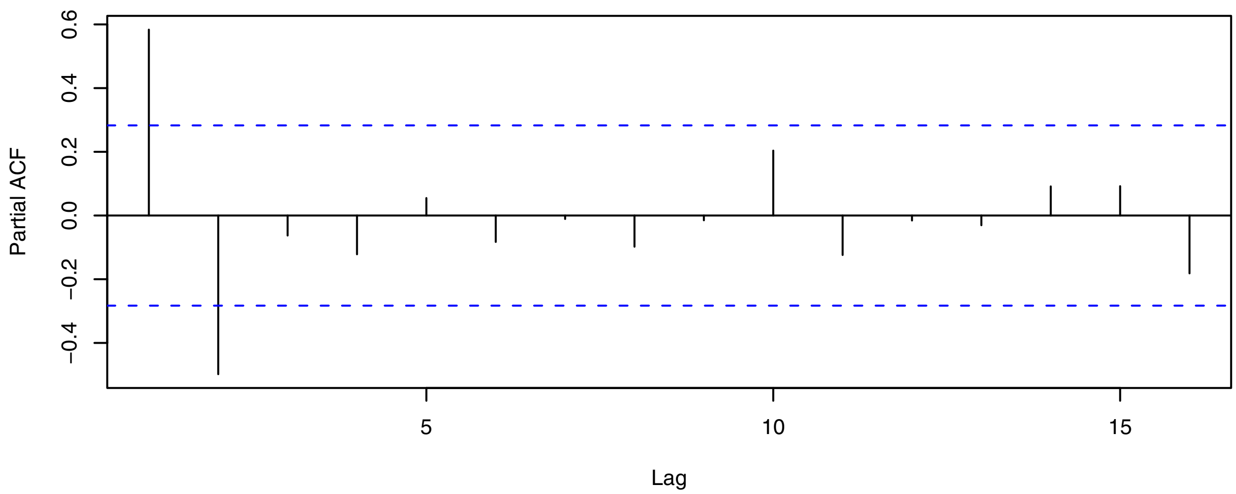
Given the (approximate) standard errors of the estimates, the method of moments and mle estimates are similar. (Recall that we use the oposite sign for the MA parameter.)

(c) Repeat parts (a) and (b) with a new simulated series under the same conditions.

Exercise 7.20 Simulate an AR(2) series with $\phi_1 = 1.0$, $\phi_2 = -0.6$, $n = 48$ but with error terms from a t-distribution with 5 degrees of freedom.

```
> set.seed(54321); series=arima.sim(n=48,list(ar=c(1,-0.6)),innov=rt(n,5))
```

(a) Display the sample PACF of the series. Is an AR(2) model suggested?




```
> pacf(series)
```

The partial acf gives a clear indication of an AR(2) model.

(b) Estimate ϕ_1 and ϕ_2 from the series and comment on the results.

```
> ar(series,order.max=2,aic=F,method='yw'); arima(series,order=c(2,0,0),method='ML')
```

```
Call:
ar(x = series, aic = F, order.max = 2, method = "yw")

Coefficients:
      1        2
 0.8742  -0.4982

Order selected 2  sigma^2 estimated as  1.609

Call:
arima(x = series, order = c(2, 0, 0), method = "ML")

Coefficients:
         ar1      ar2  intercept
      0.9454  -0.5648     0.0466
s.e.  0.1223   0.1209     0.2727

sigma^2 estimated as 1.348:  log likelihood = -75.89,  aic = 157.79
```

In the light of the appropriate standard errors, both Yule-Walker and (pseudo) maximum likelihood, give quite similar estimates of the AR(2) parameters.

(c) Repeat parts (a) and (b) with a new simulated series under the same conditions.

Exercise 7.21 Simulate an ARMA(1,1) series with $\phi = 0.7$, $\theta = -0.6$, $n = 48$ but with error terms from a t-distribution with 6 degrees of freedom.

```
> set.seed(54321); series=arima.sim(n=48,list(ar=0.7,ma=0.6),innov=rt(n,6))
```

(a) Display the sample EACF of the series. Is an ARMA(1,1) model suggested?

```
> eacf(series)
```

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	x	x	o	o	o	o	o	o	o	o	o	o	o
1	x	o	o	o	o	o	o	o	o	o	o	o	o	o
2	x	o	o	o	o	o	o	o	o	o	o	o	o	o
3	x	x	o	o	o	o	o	o	o	o	o	o	o	o
4	o	x	o	o	o	o	o	o	o	o	o	o	o	o
5	o	o	x	x	o	o	o	o	o	o	o	o	o	o
6	x	o	o	o	o	x	o	o	o	o	o	o	o	o
7	x	o	o	o	o	o	o	o	o	o	o	o	o	o

The sample eacf provides a rather clear indication of an ARMA(,1) model.

(b) Estimate ϕ and θ from the series and comment on the results.

```
> arima(series,order=c(1,0,1))
```

```
Call:
arima(x = series, order = c(1, 0, 1))

Coefficients:
         ar1     ma1  intercept
      0.6606  0.7156     0.9246
s.e.  0.1240  0.1465     0.8133

sigma^2 estimated as 1.332:  log likelihood = -76.02,  aic = 158.03
```

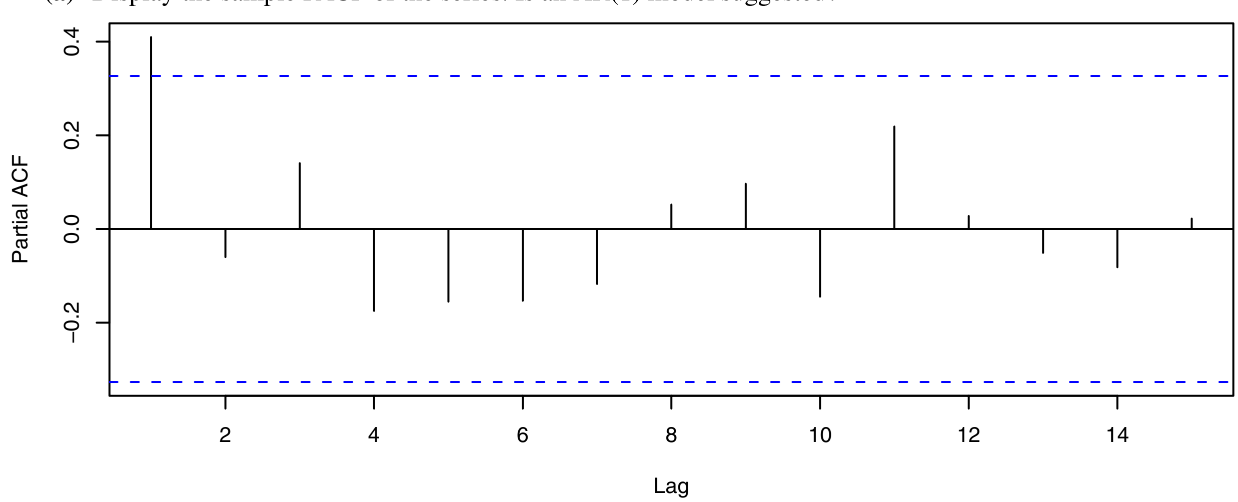
In spite of the nonnormal errors, maximum "likelihood" produces excellent estimates of the ϕ and θ parameters.

(c) Repeat parts (a) and (b) with a new simulated series under the same conditions.

Exercise 7.22 Simulate an AR(1) series with $\phi = 0.6$, $n = 36$ but with error terms from a chi-square distribution with 6 degrees of freedom.

```
> set.seed(54321); series=arima.sim(n=36,list(ar=0.6),innov=rchisq(n,5))
```

(a) Display the sample PACF of the series. Is an AR(1) model suggested?




```
> pacf(series)
```

The sample pacf gives a clear indication of an AR(1) model even though the error terms are chi-square distributed.

(b) Estimate ϕ from the series and comment on the results.

```
> ar(series,order.max=1,aic=F,method='yw')
> arima(series,order=c(1,0,0)) # Maximum likelihood is the default estimation method
```

```
Call:
ar(x = series, aic = F, order.max = 1, method = "yw")

Coefficients:
     1
0.4097

Order selected 1  sigma^2 estimated as  21.1

Call:
arima(x = series, order = c(1, 0, 0))

Coefficients:
         ar1  intercept
      0.4695    13.6086
s.e.  0.1582     1.3442

sigma^2 estimated as 19.19:  log likelihood = -104.38,  aic = 212.77
```

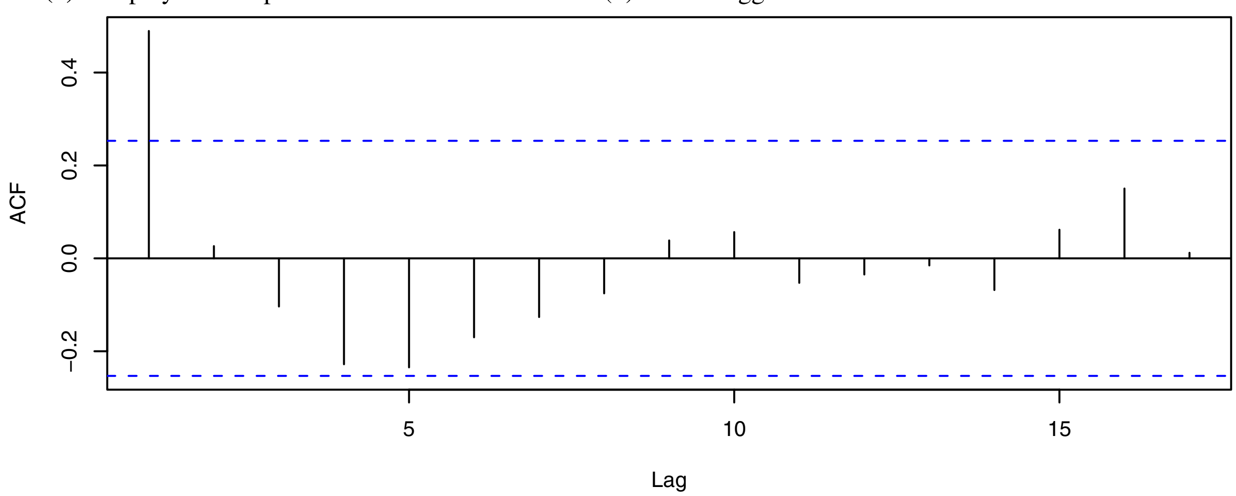
The Yule-Walker (method of moments) estimate and (pseudo) maximum likelihood estimates are quite similar with the ML estimate a little better.

(c) Repeat parts (a) and (b) with a new simulated series under the same conditions.

Exercise 7.23 Simulate an MA(1) series with $\theta = -0.8$, $n = 60$ but with error terms from a chi-square distribution with 7 degrees of freedom.

```
> set.seed(54321); series=arima.sim(n=60,list(ma=0.8),innov=rchisq(n,7))
```

(a) Display the sample ACF of the series. Is an MA(1) model suggested?




```
> acf(series)
```

The sample acf points to the MA(1) model quite clearly.

(b) Estimate θ from the series and comment on the results.

```
> estimate.ma1.mom(series); arima(series,order=c(0,0,1))
```

```
Method of moments estimate of theta = -0.8131914

Call:
arima(x = series, order = c(0, 0, 1))
```

```
Coefficients:
         ma1  intercept
      0.7954    12.6476
s.e.  0.0743     0.9679

sigma^2 estimated as 17.69:  log likelihood = -171.83,  aic = 347.67
```

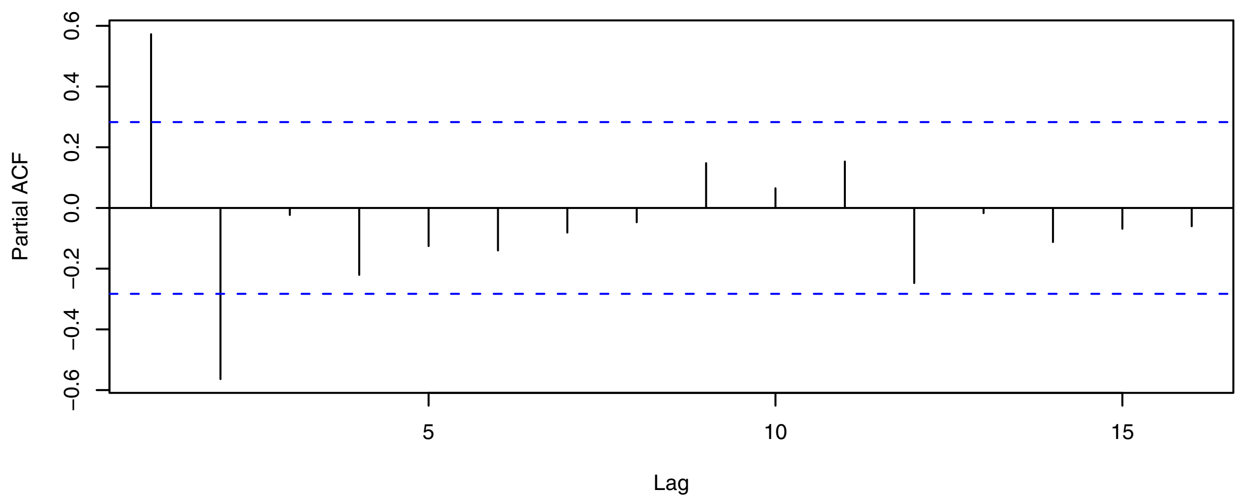
Both method of moments and (pseudo) maximum likelihood give similar good estimates of theta in this simulation.

(c) Repeat parts (a) and (b) with a new simulated series under the same conditions.

Exercise 7.24 Simulate an AR(2) series with $\phi_1 = 1.0$, $\phi_2 = -0.6$, $n = 48$ but with error terms from a chi-square distribution with 8 degrees of freedom.

```
> set.seed(54321); series=arima.sim(n=48,list(ar=c(1,-0.6)),innov=rchisq(n,8))
```

(a) Display the sample PACF of the series. Is an AR(2) model suggested?




```
> pacf(series)
```

The sample pacf gives a clear indication of the AR(2) model.

(b) Estimate ϕ_1 and ϕ_2 from the series and comment on the results.

```
> arima(series,order=c(2,0,0))
```

```
Call:
arima(x = series, order = c(2, 0, 0), method = "ML")

Coefficients:
         ar1      ar2  intercept
      0.9529  -0.6074    13.3780
s.e.  0.1155   0.1129     1.0207

sigma^2 estimated as 21.05:  log likelihood = -141.91,  aic = 289.82
```

The (pseudo) maximum li9kelihood estimates are quiten close to the true values.

(c) Repeat parts (a) and (b) with a new simulated series under the same conditions.

Exercise 7.25 Simulate an ARMA(1,1) series with $\phi = 0.7$, $\theta = -0.6$, $n = 48$ but with error terms from a chi-square distribution with 9 degrees of freedom.

```
> set.seed(54321); series=arima.sim(n=48,list(ar=0.7,ma=0.6),innov=rchisq(n,9))
```

(a) Display the sample EACF of the series. Is an ARMA(1,1) model suggested?

```
> eacf(series)
```

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	x	o	o	o	o	o	o	o	o	o	o	o	o
1	x	o	o	o	o	o	o	o	o	o	o	o	o	o
2	x	x	o	o	o	o	o	o	o	o	o	o	o	o
3	x	x	x	o	o	o	o	o	o	o	o	o	o	o
4	x	o	o	o	o	o	o	o	o	o	o	o	o	o
5	o	o	o	o	o	o	o	o	o	o	o	o	o	o
6	o	x	o	o	o	o	o	o	o	o	o	o	o	o
7	x	o	o	o	o	o	o	o	o	o	o	o	o	o

The sample eacf gives a clear indication of the mixed ARMA(1,1) model.

(b) Estimate ϕ and θ from the series and comment on the results.

```
> arima(series,order=c(1,0,1))
```

```
Call:
arima(x = series, order = c(1, 0, 1))

Coefficients:
         ar1     ma1  intercept
      0.6413  0.8027    45.8962
s.e.  0.1420  0.1163     3.7751

sigma^2 estimated as 29.10:  log likelihood = -150.2,  aic = 306.4
```

Relative to their standard errors, the estimates are not significantly different from their true values.

(c) Repeat parts (a) and (b) with a new series under the same conditions.

Exercise 7.26 Consider the AR(1) model specified for the color property time series displayed in Exhibit (1.3), page 3. The data are in the file named color.

(a) Find the method-of-moments estimate of ϕ.

```
> data(color); acf(color)$acf[1]
```

Method of momemts estimate of phi = 0.5282091

(b) Find the maximum likelihood estimate of ϕ and compare it with part (a).

```
> arima(color,order=c(1,0,0)) # Maximum likelihood is the default estimation method
```

```
Call:
arima(x = color, order = c(1, 0, 0))

Coefficients:
         ar1  intercept
      0.5706    74.3293
s.e.  0.1435     1.9151

sigma^2 estimated as 24.83:  log likelihood = -106.07,  aic = 216.15
```

Relative to the standard errors, the two methods give similar estimates of phi. See Exhibit (7.7), page 165, for an alternative solution to this exercise.

Exercise 7.27 Exhibit (6.31), page 139, suggested specifying either an AR(1) or possibly an AR(4) model for the difference of the logarithms of the oil price series. The data are in the file named oil.price.

(a) Estimate both of these models using maximum likelihood and compare it with the results using the AIC criteria.

```
> data(oil.price); arima(log(oil.price),order=c(1,1,0))
```

```
Call:
arima(x = log(oil.price), order = c(1, 1, 0))

Coefficients:
         ar1
      0.2364
s.e.  0.0660

sigma^2 estimated as 0.006787:  log likelihood = 258.55,  aic = -515.11
```

```
> arima(log(oil.price),order=c(4,1,0))
```

```
Call:
arima(x = log(oil.price), order = c(4, 1, 0))

Coefficients:
         ar1      ar2     ar3      ar4
      0.2673  -0.1550  0.0238  -0.0970
s.e.  0.0669   0.0691  0.0691   0.0681

sigma^2 estimated as 0.006603:  log likelihood = 261.82,  aic = -515.64
```

These two models are very similar. The ar3 and ar4 coefficients in the AR(4) model are not significantly different from zero.

(b) Exhibit (6.32), page 140 suggested specifying an MA(1) model for the difference of the logs. Estimate this model by maximum likelihood and compare to your results in part (a).

```
> arima(log(oil.price),order=c(0,1,1))
```

```
Call:
arima(x = log(oil.price), order = c(0, 1, 1))

Coefficients:
         ma1
      0.2956
s.e.  0.0693

sigma^2 estimated as 0.006689:  log likelihood = 260.29,  aic = -518.58
```

There is actually very little difference among these three models.

Exercise 7.28 The data file named deere3 contains 57 consecutive values from a complex machine tool at Deere & Co. The values given are deviations from a target value in units of ten millionths of an inch. The process employs a control mechanism that resets some of the parameters of the machine tool depending on the magnitude of deviation from target of the last item produced.

(a) Estimate the parameters of an AR(1) model for this series.

```
> data(deere3); arima(deere3,order=c(1,0,0))
```

```
Call:
arima(x = deere3, order = c(1, 0, 0))

Coefficients:
         ar1  intercept
      0.5256   124.3524
s.e.  0.1108   394.2320

sigma^2 estimated as 2069354:  log likelihood = -495.51,  aic = 995.02
```

The ar1 (= phi1hat) coefficient is significantly different from zero.

(b) Estimate the parameters of an AR(2) model for this series and compare the results with those in part (a).

```
> arima(deere3,order=c(2,0,0))
```

```
Call:
arima(x = deere3, order = c(2, 0, 0))

Coefficients:
         ar1     ar2  intercept
      0.5211  0.0083   123.2418
s.e.  0.1310  0.1315   397.5991

sigma^2 estimated as 2069209:  log likelihood = -495.51,  aic = 997.01
```

The ar2 (= phi2hat) coefficient is not statistically significant so the AR(1) model still looks good.

Exercise 7.29 The data file named robot contains a time series obtained from an industrial robot. The robot was put through a sequence of maneuvers, and the distance from a desired ending point was recorded in inches. This was repeated 324 times to form the time series.

(a) Estimate the parameters of an AR(1) model for these data.

```
> data(robot); arima(robot,order=c(1,0,0))
```

```
Call:
arima(x = robot, order = c(1, 0, 0))

Coefficients:
         ar1  intercept
      0.3076     0.0015
s.e.  0.0528     0.0002

sigma^2 estimated as 6.482e-06:  log likelihood = 1475.54,  aic = -2947.08
```

Notice that both the ar1 and intercept coefficients are significantly different from zero statistically.

(b) Estimate the parameters of an IMA(1,1) model for these data.

```
> arima(robot,order=c(0,1,1))
```

```
Call:
arima(x = robot, order = c(0, 1, 1))

Coefficients:
          ma1
      -0.8713
s.e.   0.0389

sigma^2 estimated as 6.069e-06:  log likelihood = 1480.95,  aic = -2959.9
```

The ma1 coefficient is significantly different from zero statistically.

(c) Compare the results from parts (a) and (b) in terms of AIC.

The nonstationary IMA(1,1) model has a slightly smaller AIC value but the log likelihoods and AIC values are very close to each other.

Exercise 7.30 The data file named days contains accounting data from the Winegard Co. of Burlington, Iowa. The data are the number of days until Winegard receives payment for 130 consecutive orders from a particular distributor of Winegard products. (The name of the distributor must remain anonymous for confidentiality reasons.) The time series contains outliers that are quite obvious in the time series plot.

(a) Replace each of the unusual values with a value of 35 days, a much more typical value, and then estimate the parameters of an MA(2) model.

```
> data(days); daysmod=days; daysmod[63]=35; daysmod[106]=35; daysmod[129]=35
> arima(daysmod,order=c(0,0,2))
```

```
Call:
arima(x = daysmod, order = c(0, 0, 2))

Coefficients:
         ma1     ma2  intercept
```

```
        0.1893  0.1958    28.1957
s.e.    0.0894  0.0740     0.6980

sigma^2 estimated as 33.22:  log likelihood = -412.23,  aic = 830.45
```

All three estimated coefficients are significantly different from zero statistically.

(b) Now assume an MA(5) model and estimate the parameters. Compare these results with those obtained in part (a).

```
> arima(daysmod,order=c(0,0,5))
```

```
Call:
arima(x = daysmod, order = c(0, 0, 5))

Coefficients:
         ma1     ma2     ma3     ma4      ma5  intercept
      0.1844  0.2680  0.0305  0.1717  -0.0859    28.2351
s.e.  0.0898  0.0929  0.1033  0.0850   0.0932     0.7755

sigma^2 estimated as 32.02:  log likelihood = -409.93,  aic = 831.86
```

In this model only the ma1, ma2, ma4 (just barely) coefficients and intercept are significantly different from zero statistically. The AIC of the MA(5) model is actually a little larger than that for the MA(2) model.

Exercise 7.31 Simulate a time series of length $n = 48$ from an AR(1) model with $\phi = 0.7$. Use that series as if it were real data. Now compare the theoretical asymptotic distribution of the estimator of ϕ with the distribution of the bootstrap estimator of ϕ.

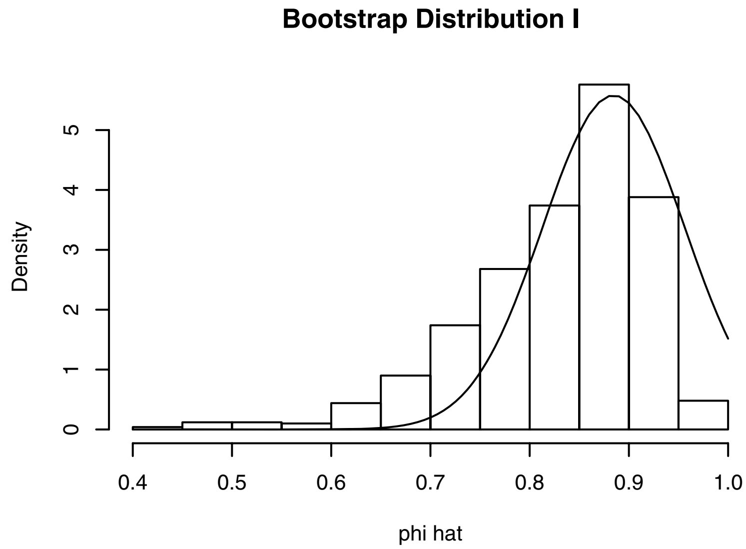
```
> set.seed(54321); series=arima.sim(n=48,list(ar=0.7))
> result = arima(series,order=c(1,0,0))
```

```
Call:
arima(x = series, order = c(1, 0, 0))

Coefficients:
         ar1  intercept
      0.8846    -0.0489
s.e.  0.0715     1.0160
```

The s.e. of 0.0715 is based on large-sample theory.

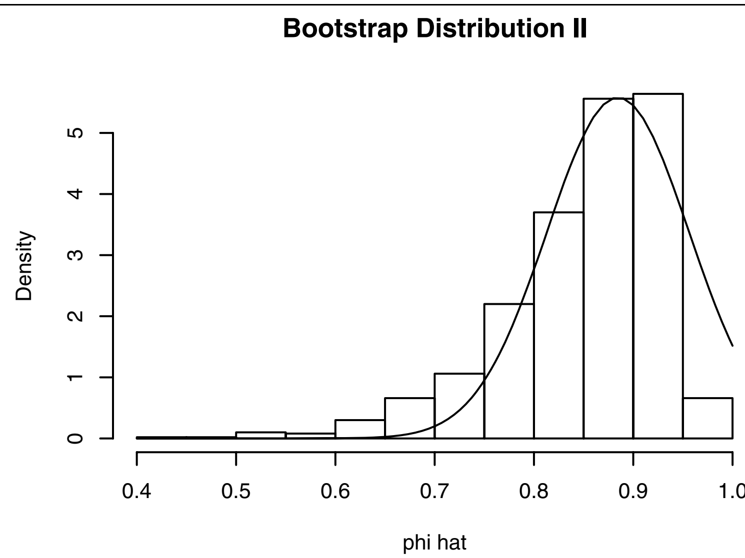
```
> x=seq(from=0.4,to=1,by=0.01); y=dnorm(x,mean=0.8846,sd=0.0715)
> # Setup asymptotic distribution.
```

```
> set.seed(12345) # Bootstrap Method I
> coefmI=arima.boot(result,cond.boot=T,is.normal=T,B=1000,init=series)
> win.graph(width=4,height=3,pointsize=8)
> hist(coefmI[,1],xlab='phi hat',main='Bootstrap Distribution I',freq=F)
> lines(x,y,type='l')
```



The bootstrap distribution is skewed somewhat strongly toward lower values and, of course, the asymptotic normal distribution is symmetric. The asymptotic distribution has significant probablity above $\phi = 1$.

```
> set.seed(12345) # Method II
> coefmII=arima.boot(result,cond.boot=T,is.normal=F,B=1000,init=series)
> hist(coefmII[,1],xlab='phi hat',main='Bootstrap Distribution II')
> lines(x,y,type='l')
```



This bootstrap distribution is also skewed strongly toward lower values and, of course, the asymptotic normal distribution is symmetric. The asymptotic distribution has significant probablity above $\phi = 1$.

Exercise 7.32 The industrial color property time series was fitted quite well by an AR(1) model. However, the series is rather short, with $n = 35$. Compare the theoretical asymptotic distribution of the estimator of ϕ with the distribution of the bootstrap estimator of ϕ. The data are in the file named color.

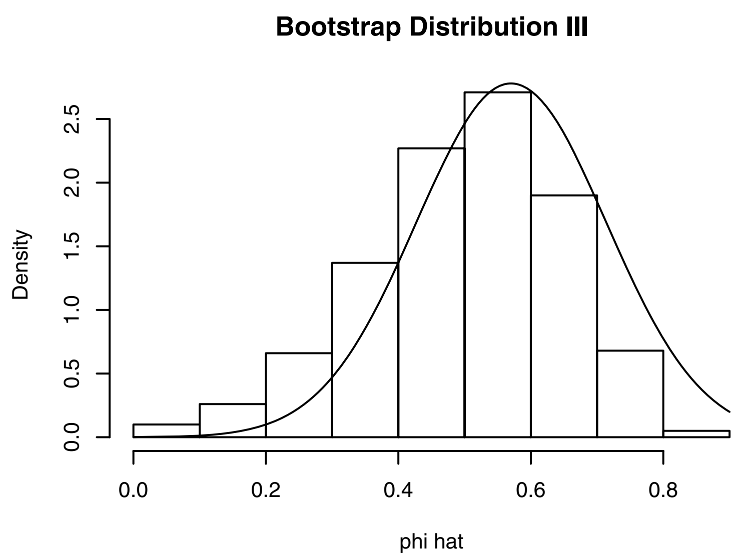
```
> data(color); arima(color,order=c(1,0,0))
```

```
Call:
arima(x = color, order = c(1, 0, 0))

Coefficients:
```

```
            ar1  intercept
         0.5705    74.3293
   s.e.  0.1435     1.9151

   sigma^2 estimated as 24.83:  log likelihood = -106.07,  aic = 216.15
```

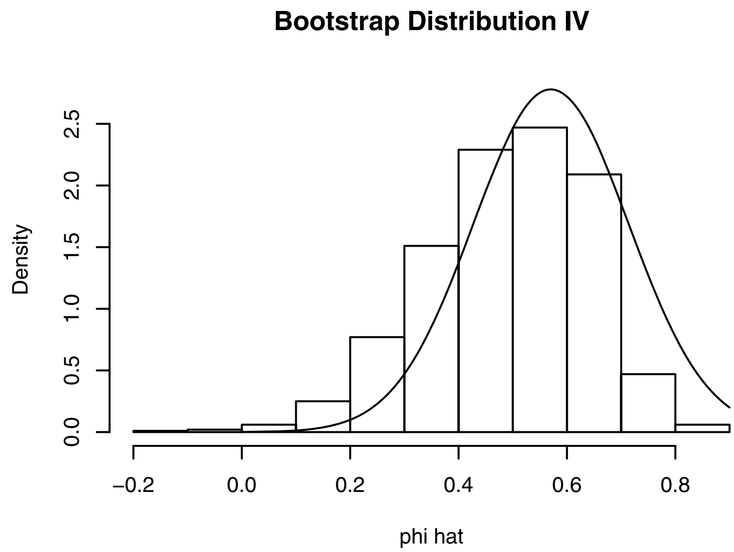
```
> x=seq(from=0.0,to=0.9,by=0.01); y=dnorm(x,mean=0.5705,sd=0.1435)
> # Setup asymptotic distribution.
```

```
> set.seed(12345) # Bootstrap Method III
> coefmIII=arima.boot(result,cond.boot=F,is.normal=T,ntrans=100,B=1000,init=color)
> hist(coefmIII[,1],xlab='phi hat',main='Bootstrap Distribution III',freq=F)
> lines(x,y,type='l')
```



The bootstrap distribution is skewed strongly toward lower values and, of course, the asymptotic normal distribution is symmetric.

```
> set.seed(12345) # Method IV
> coefmIV=arima.boot(result,cond.boot=F,is.normal=F,ntrans=100,B=1000,init=color)
> hist(coefmIV[,1],xlab='phi hat', main='Bootstrap Distribution IV',
   freq=F,ylim=c(0,2.8))
> lines(x,y,type='l')
```



This bootstrap distribution is also skewed strongly toward lower values while the asymptotic normal distribution is symmetric.

CHAPTER 8

Exercise 8.1 For an AR(1) model with $\phi \approx 0.5$ and $n = 100$, the lag 1 sample autocorrelation of the residuals is 0.5. Should we consider this unusual? Why or why not?

From Equation (8.1.5), page 180, we have that $\sqrt{Var(\hat{r}_1)} \approx \sqrt{\phi^2/n} = \sqrt{0.5^2/100} = 0.05$ so that would expect the lag 1 sample autocorrelation of the residuals to be within ±0.1. The residual autocorrelation of 0.5 is most unusual.

Exercise 8.2 Repeat Exercise 8.1 for an MA(1) model with $\theta \approx 0.5$ and $n = 100$.

From Equation (8.1.5), page 180, with θ replacing φ as indicated on page 183, we have that $\sqrt{Var(\hat{r}_1)} \approx \sqrt{\theta^2/n}$ $= \sqrt{0.5^2/100} = 0.05$ so that again a lag 1 residual autocorrelation of 0.5 is most unusual.

Exercise 8.3 Based on a series of length $n = 200$, we fit an AR(2) model and obtain residual autocorrelations of $\hat{r}_1 = 0.13$, $\hat{r}_2 = 0.13$, and $\hat{r}_3 = 0.12$. If $\hat{\phi}_1 = 1.1$ and $\hat{\phi}_2 = -0.8$, do these residual autocorrelations support the AR(2) specification? Individually? Jointly?

From Equation (8.1.8), page 182, $\sqrt{Var(\hat{r}_1)} \approx \sqrt{\phi_2^2/n} = \sqrt{(-0.8)^2/200} = 0.057$ so that $\hat{r}_1 = 0.13$ is "too large."

From Equation (8.1.9), page 182, $\sqrt{Var(\hat{r}_2)} \approx \sqrt{(\phi_2^2 + \phi_1^2(1 + \phi_2)^2)/n} = 0.059$ so that $\hat{r}_2 = 0.13$ is also "too large."

From Equation (8.1.10), page 182, for $k > 2$, $\sqrt{Var(\hat{r}_k)} \approx \sqrt{(1/200)} = 0.071$ so that $\hat{r}_3 = 0.12$ is OK.

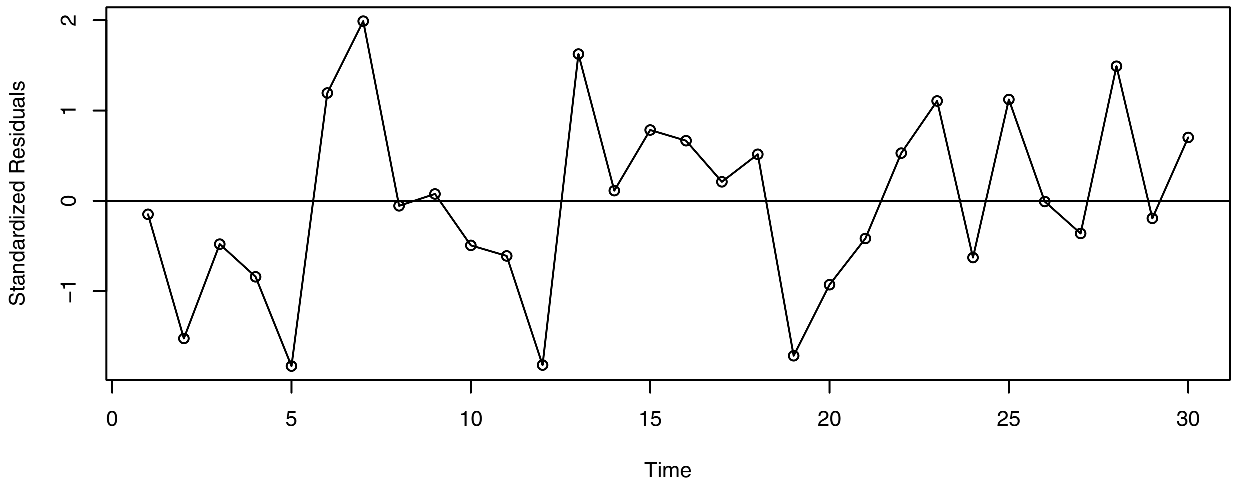
The Ljung-Box**,** statistic, $Q_* = n(n+2)\left(\frac{\hat{r}_1^2}{n-1} + \frac{\hat{r}_2^2}{n-2} + \frac{\hat{r}_3^2}{n-3}\right) = 200(202)\left(\frac{(0.13)^2}{199} + \frac{(0.13)^2}{198} + \frac{(0.12)^2}{197}\right) = 9.83$. If the AR(2) specification is correct, then Q_* has (approximately) a chi-square distribution with $3 - 2 = 1$ degree of freedom. However, $Pr[\chi_1^2 > 9.83] = 0.0017$ so that these residual autocorrelations are (jointly) too large to support the AR(2) model.

Exercise 8.4 Simulate an AR(1) model with $n = 30$ and $\phi = 0.5$.

```
> set.seed(12347); series=arima.sim(n=30,list(ar=0.5))
```

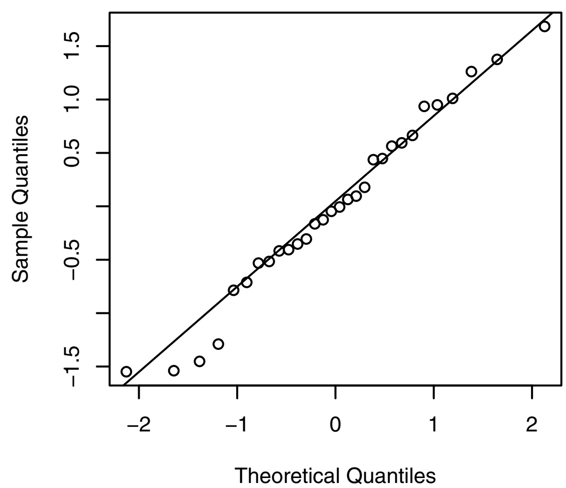
(a) Fit the correctly specified AR(1) model and look at a time series plot of the residuals. Does the plot support the AR(1) specification?

```
> model=arima(series,order=c(1,0,0)); win.graph(width=6.5,height=3,pointsize=8)
> plot(rstandard(model),ylab ='Standardized Residuals', type='o'); abline(h=0)
```



These standardized residuals look fairly “random” with no particular patterns.

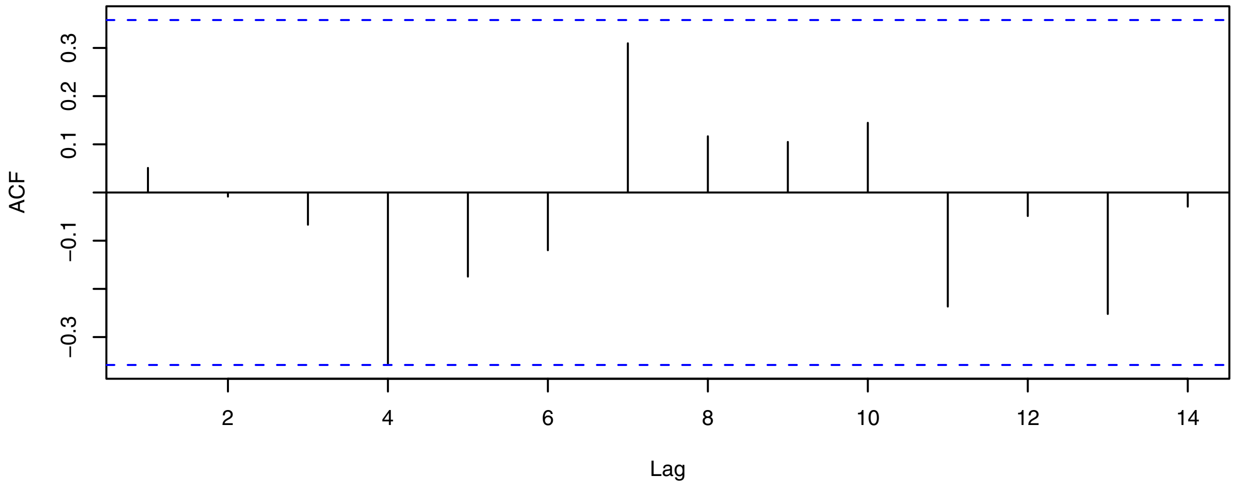
(b) Display a normal quantile-quantile plot of the standardized residuals. Does the plot support the AR(1) specification?




```
> win.graph(width=3,height=3,pointsize=8)
> qqnorm(residuals(model)); qqline(residuals(model))
```

With a few minor exceptions in the lower tail, the Q-Q plot of the standardized residuals looks reasonably “normal.”

(c) Display the sample ACF of the residuals. Does the plot support the AR(1) specification?




```
> acf(residuals(model))
```

The sample acf at lag 4 is the only individual autocorrelation that comes close to being "significant."

(d) Calculate the Ljung-Box statistic summing to $K = 8$. Does this statistic support the AR(1) specification?

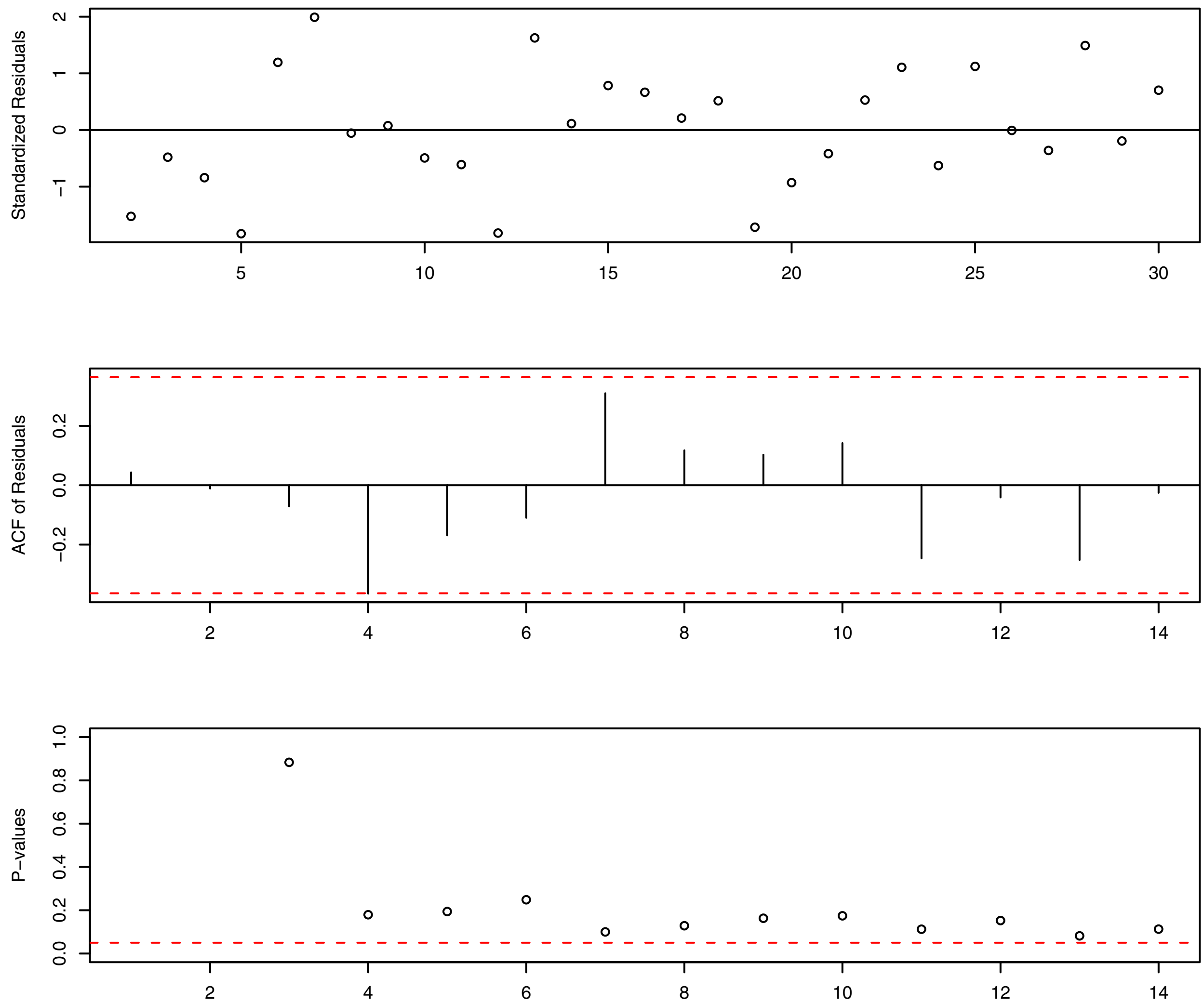
```
> LB.test(model,lag=8)
```

```
        Box-Ljung test

data:  residuals from  model
X-squared = 11.2399, df = 7, p-value = 0.1285
```

This test does not reject randomness of the error terms based on the first eight autocorrelations of the residuals.

Note: The `tsdiag` function will produce a display that answer all three parts (a), (c), and (d) of this exercise. See below.

```
> win.graph(width=6.5,height=6,pointsize=8); tsdiag(model)
```

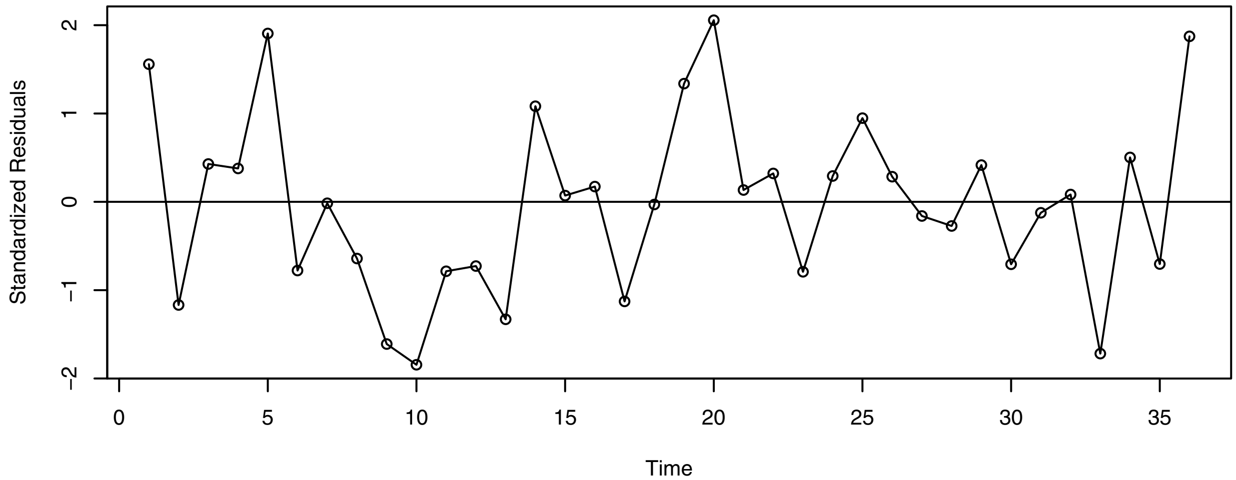


The bottom display shows the p-values of the Ljung-Box test for a variety of values of the "K" parameter—the highest lag used in the sum. The top display will flag potential outliers, if any, using the Bonferroni criteria.

Exercise 8.5 Simulate an MA(1) model with $n = 36$ and $\theta = -0.5$.

```
> set.seed(64231); series=arima.sim(n=36,list(ma=0.5))
```

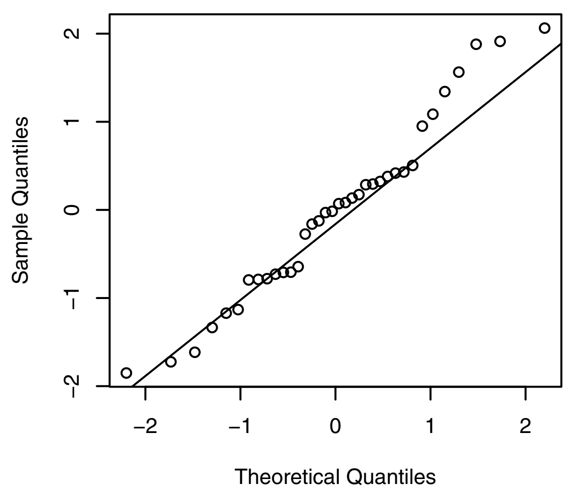
(a) Fit the correctly specified MA(1) model and look at a time series plot of the residuals. Does the plot support the MA(1) specification?




```
> model=arima(series,order=c(0,0,1))
> win.graph(width=6.5,height=3,pointsize=8)
> plot(rstandard(model),ylab ='Standardized Residuals', type='o'); abline(h=0)
```

The sequence plot of the standardized residuals looks fairly "random."

(b) Display a normal quantile-quantile plot of the standardized residuals. Does the plot support the MA(1) specification?




```
> win.graph(width=3,height=3,pointsize=8)
> qqnorm(residuals(model)); qqline(residuals(model))
```

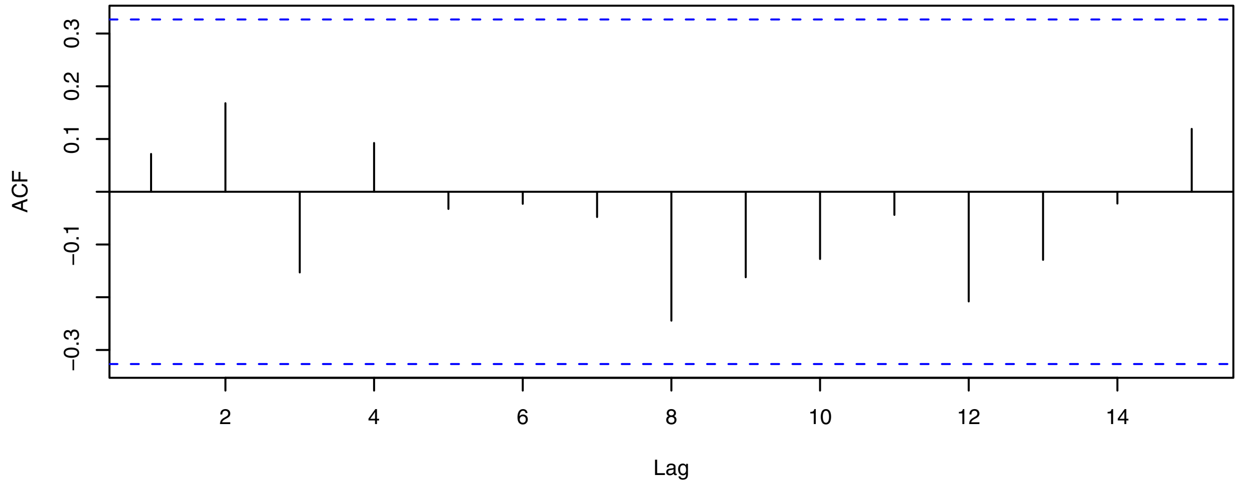
The Q-Q plot is fairly straight but there may be some problem with the upper tail but sample size is quite small. Could do a Shapiro-Wilk test to check further.

```
> shapiro.test(residuals(model))
```

```
Shapiro-Wilk normality test

data:  residuals(model)
W = 0.968, p-value = 0.374
```

We would not reject normality based on these test results.

(c) Display the sample ACF of the residuals. Does the plot support the MA(1) specification?




```
> win.graph(width=6.5,height=3,pointsize=8); acf(residuals(model))
```

No problem with significant residual autocorrelations in this simulation.

(d) Calculate the Ljung-Box statistic summing to $K = 6$. Does this statistic support the MA(1) specification?

```
> LB.test(model,lag=6)
```

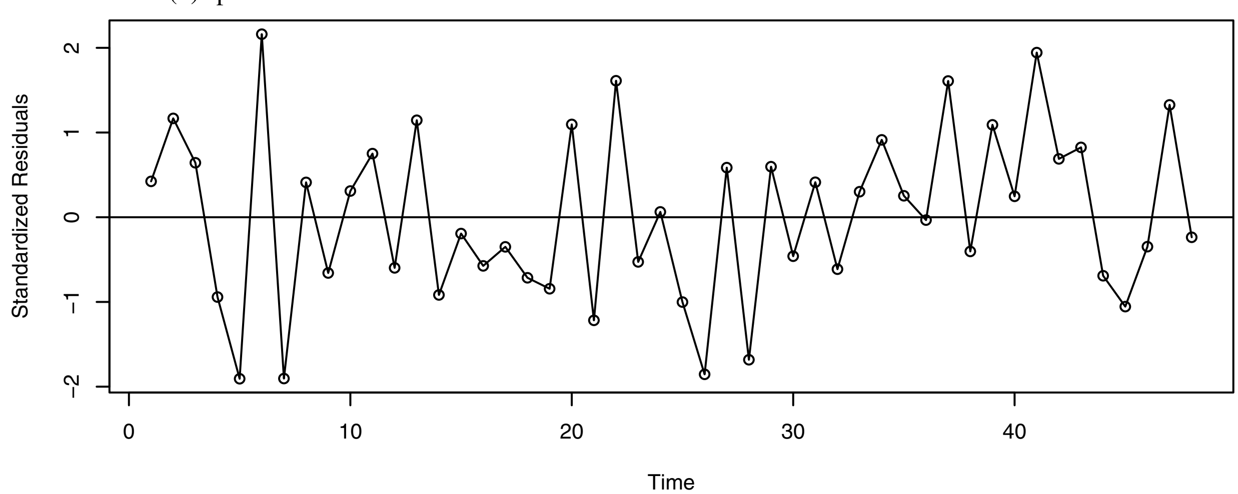
```
        Box-Ljung test

data:  residuals from  model
X-squared = 6.6437, df = 5, p-value = 0.2485
```

No problem with large residual autocorrelations jointly out to lag 6.

Exercise 8.6 Simulate an AR(2) model with $n = 48$, $\phi_1 = 1.5$, and $\phi_2 = -0.75$.

```
> set.seed(65423); series=arima.sim(n=48,list(ar=c(1.5,-0.75)))
```

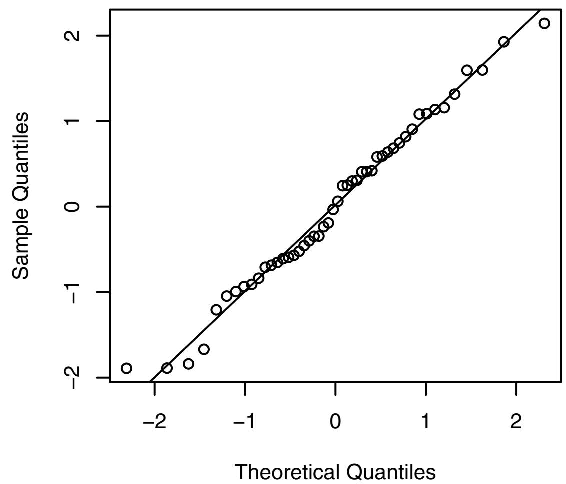
(a) Fit the correctly specified AR(2) model and look at a time series plot of the residuals. Does the plot support the AR(2) specification?




```
> model=arima(series,order=c(2,0,0)); win.graph(width=6.5,height=3,pointsize=8)
> plot(rstandard(model),ylab ='Standardized Residuals', type='o'); abline(h=0)
```

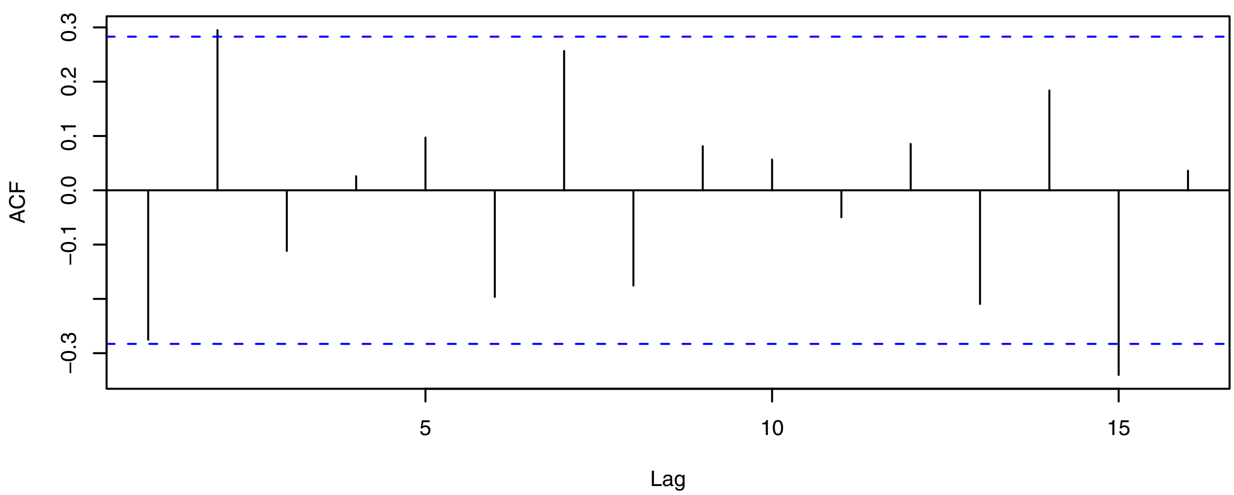
The residuals look “random.”

(b) Display a normal quantile-quantile plot of the standardized residuals. Does the plot support the AR(2) specification?



No problem with normality of the error terms.

(c) Display the sample ACF of the residuals. Does the plot support the AR(2) specification?




```
> win.graph(width=6.5,height=3,pointsize=8); acf(rstandard(model))
```

There are two residual autocorrelations that are "signifnicant:" at lags 2 and 15.

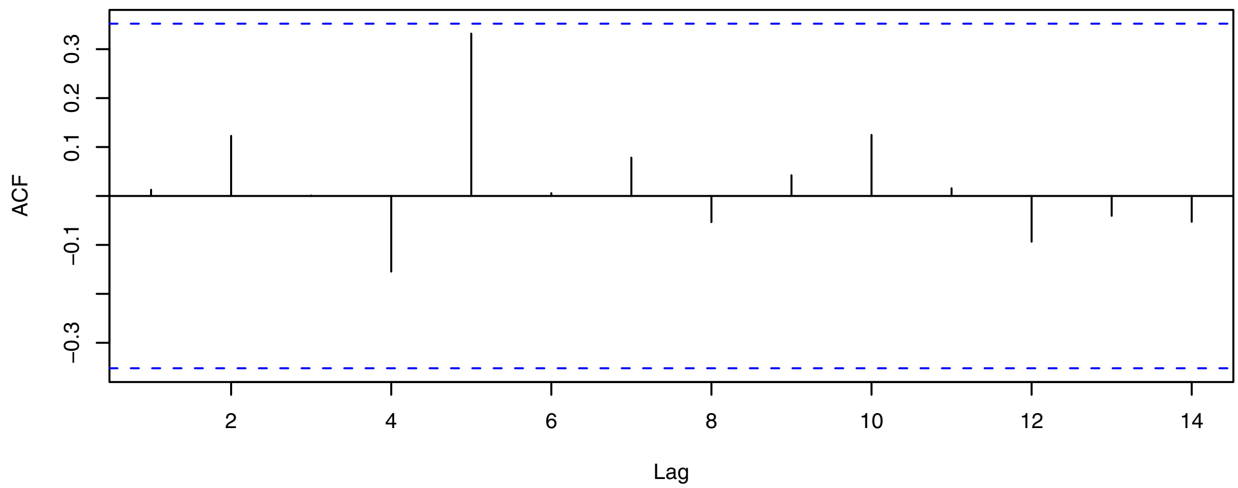
(d) Calculate the Ljung-Box statistic summing to $K = 12$. Does this statistic support the AR(2) specification?

```
> LB.test(model,lag=12)

      Box-Ljung test

      data:  residuals from  model
      X-squared = 18.7997, df = 10, p-value = 0.04288
```

Based on this test, we would reject the assumption of independent error terms at the 5% significance level for this simulation.

Exercise 8.7 Fit an AR(3) model by maximum likelihood to the square root of the hare abundance series (filename hare).

(a) Plot the sample ACF of the residuals. Comment on the size of the correlations.




```
> data(hare); model=arima(sqrt(hare),order=c(3,0,0))
> win.graph(width=6.5,height=3,pointsize=8); acf(rstandard(model))
```

These residual autocorrelations look excellent.

(b) Calculate the Ljung-Box statistic summing to $K = 9$. Does this statistic support the AR(3) specification?

```
> LB.test(model,lag=9)
```

```
Box-Ljung test

data:  residuals from  model
X-squared = 6.2475, df = 6, p-value = 0.3960
```

As we would suspect from the results in part (a), the Ljung-Box test does not reject independence of the error terms.

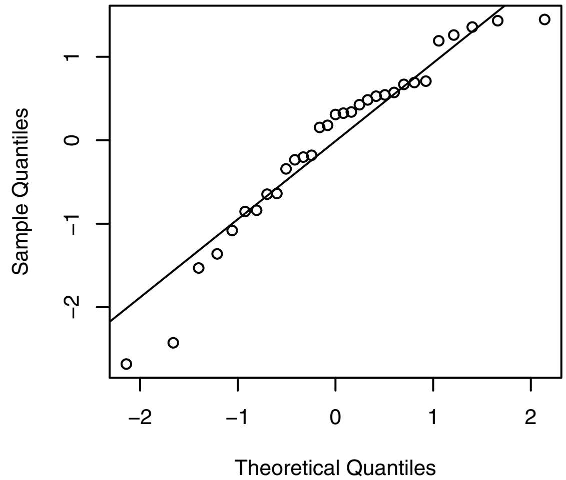
(c) Perform a runs test on the residuals and comment on the results.

```
> runs(rstandard(model))
```

```
$pvalue
[1] 0.602
$observed.runs
[1] 18
$expected.runs
[1] 16.09677
$n1
[1] 13
$n2
[1] 18
$k
[1] 0
```

The *p*-value does not permit us to reject independence of the error terms. The number of runs is not unusual!

(d) Display the quantile-quantile normal plot of the residuals. Comment on the plot.




```
> win.graph(width=3,height=3,pointsize=8)
> qqnorm(residuals(model)); qqline(residuals(model))
```

There is some minor curvature to this plot with possible outliers at both extremes.

(e) Perform the Shapiro-Wilk test of normality on the residuals.

```
> shapiro.test(residuals(model))
```

```
        Shapiro-Wilk normality test

data:  residuals(model)
W = 0.9351, p-value = 0.06043
```

We would not reject normality of the error terms at the usual significance levels.

Exercise 8.8 Consider the oil filter sales data shown in Exhibit (1.8), page 7. The data are in the file named oilfilters.

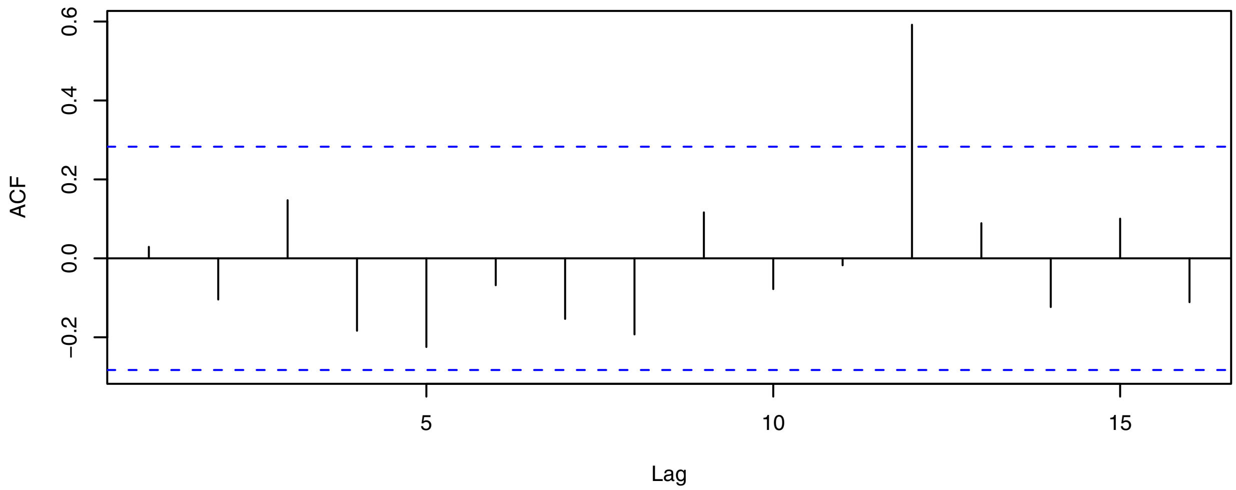
(a) Fit an AR(1) model to this series. Is the estimate of the ϕ parameter significantly different from zero statistically?

```
> data(oilfilters); model=arima(oilfilters,order=c(1,0,0)); model
```

```
Call:
arima(x = oilfilters, order = c(1, 0, 0))

Coefficients:
         ar1  intercept
      0.3115  3370.6744
s.e.  0.1368   253.1498

sigma^2 estimated as 1482802:  log likelihood = -409.19,  aic = 822.37
```

The estimate of ϕ is more than two standard errors away from zero and would be deemed significant at the usual significance levels.

(b) Display the sample ACF of the residuals from the AR(1) fitted model. Comment on the display.




```
> acf(as.numeric(rstandard(model)))
```

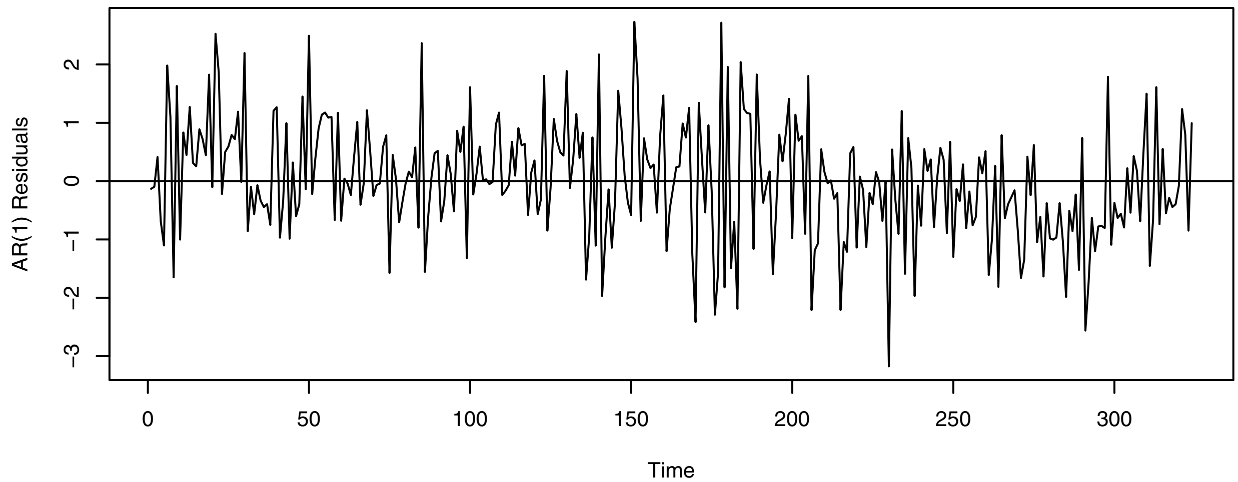
The sample autocorrelation of the residuals displays a highly significant correlation at lag 12. This series contains substantial seasonality that this model does not capture.

Exercise 8.9 The data file named robot contains a time series obtained from an industrial robot. The robot was put through a sequence of maneuvers, and the distance from a desired ending point was recorded in inches. This was repeated 324 times to form the time series. Compare the fits of an AR(1) model and an IMA(1,1) model for these data in terms of the diagnostic tests discussed in this chapter.

```
> data(robot); mod1=arima(robot,order=c(1,0,0)); res1=rstandard(mod1); mod1
> mod2=arima(robot,order=c(1,0,1)); res2=rstandard(mod2); mod2
```

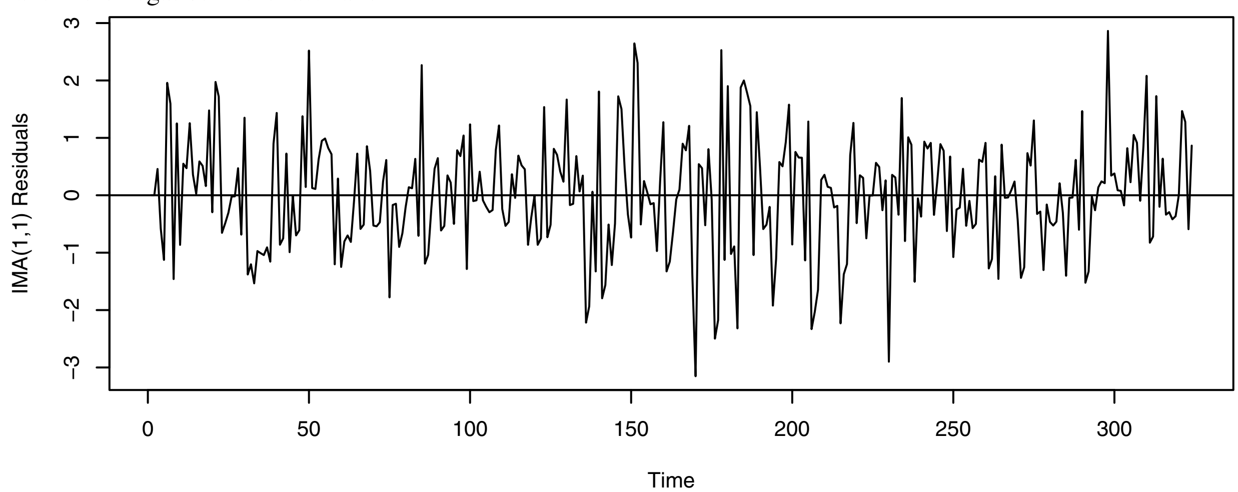
```
Call:
arima(x = robot, order = c(1, 0, 0))

Coefficients:
         ar1  intercept
      0.3074     0.0015
s.e.  0.0528     0.0002

sigma^2 estimated as 6.482e-06:  log likelihood = 1475.54,  aic = -2947.08

Call:
arima(x = robot, order = c(0, 1, 1))

Coefficients:
          ma1
      -0.8713
s.e.   0.0389

sigma^2 estimated as 6.069e-06:  log likelihood = 1480.95,  aic = -2959.9
```

Both models have statistically significant parameter estimates. The log likelihood and AIC values are just a little better in the IMA(1,1) model.



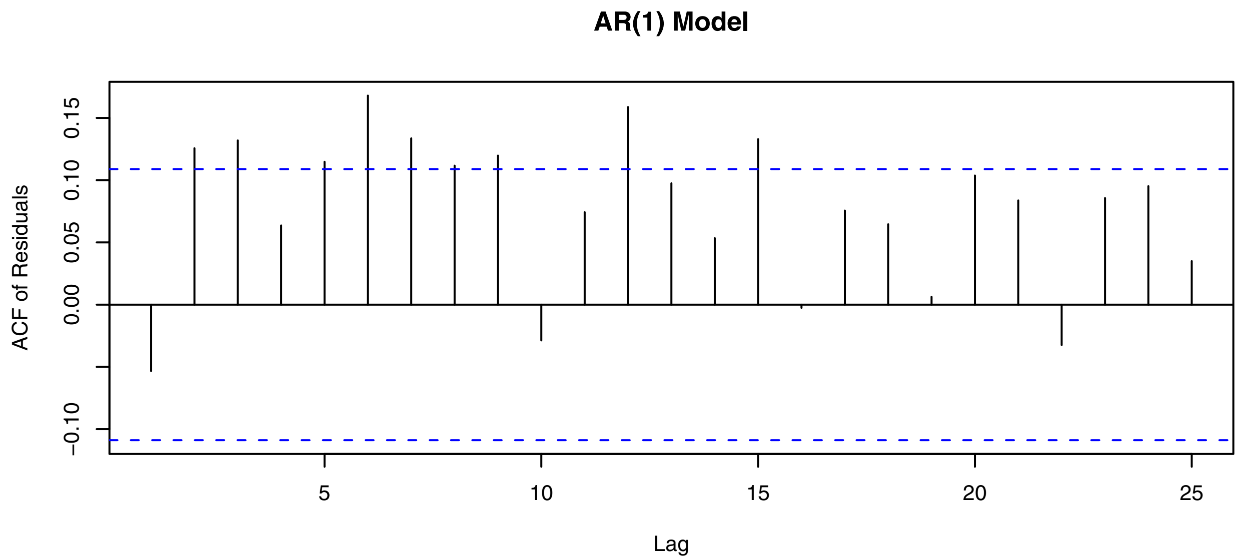

```
> win.graph(width=6.5,height=3,pointsize=8)
> plot(res1,ylab='AR(1) Residuals'); abline(h=0)
```

There may be a little drift in these residuals over time. There are more positive residuals in the first half of the series and more negative in the last half.




```
> plot(res2,ylab='IMA(1,1) Residuals'); abline(h=0)
```

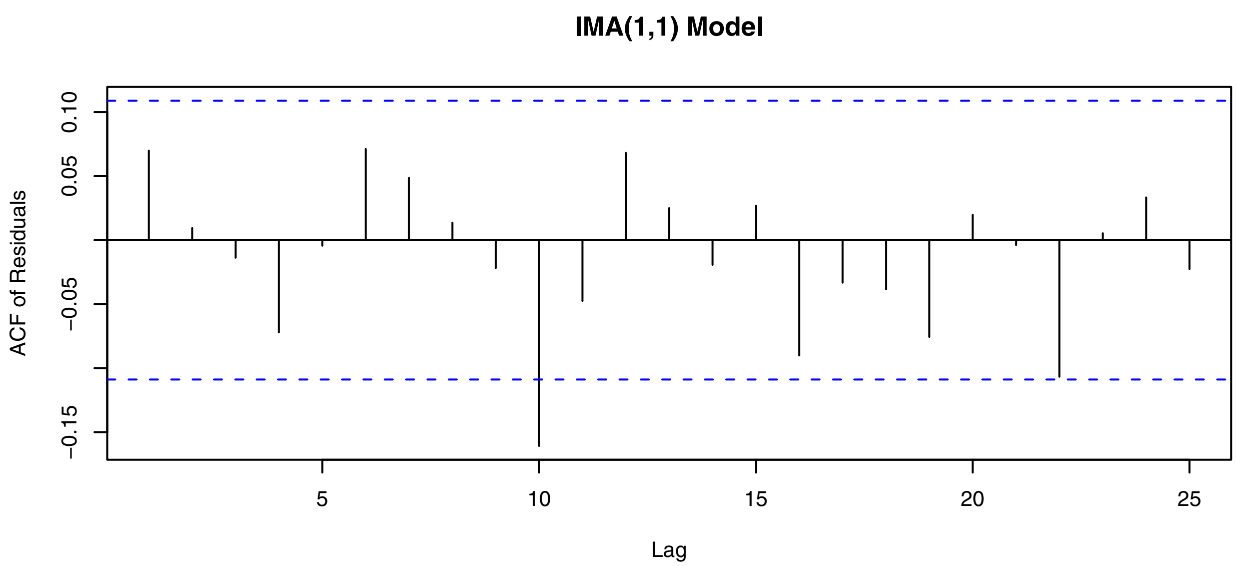
The drift observed in the residuals of the AR(1) model does not seem to appear with the IMA(1,1) model residuals. We proceed to look at correlation in the residuals.




```
> acf(residuals(mod1), main='AR(1) Model',ylab='ACF of Residuals'); LB.test(mod1)
```

```
        Box-Ljung test

data:  residuals from  mod1
X-squared = 52.5123, df = 11, p-value = 2.201e-07
```

The residuals from the AR(1) model clearly have too much autocorrelation.

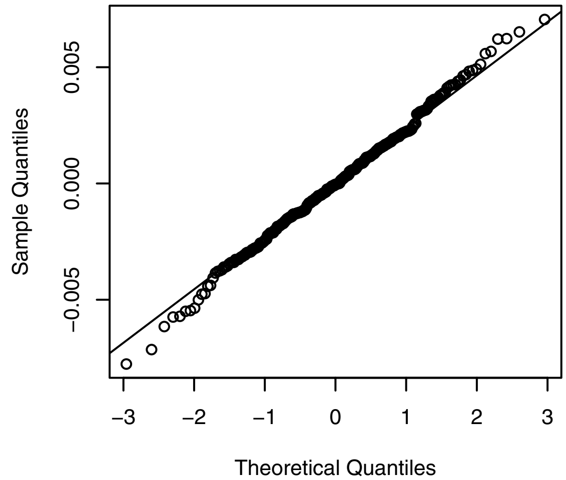



```
> acf(residuals(mod2), main='IMA(1,1) Model',ylab='ACF of Residuals'); LB.test(mod2)
```

```
        Box-Ljung test

data:  residuals from  mod2
X-squared = 17.0808, df = 11, p-value = 0.1055
```

The residuals from the IMA(1,1) model are much less correlated with only one significant autocorrelation at lag 10. The Ljung-Box test indicates that, jointly, the residual autocorrelations are not too large.

Next we check out normality of the error terms by first displayinga Q-Q plot of the residuals.




```
> win.graph(width=3,height=3,pointsize=8)
> qqnorm(residuals(mod2)); qqline(residuals(mod2))
```

The Q-Q plot looks good. Let's confirm this with the Shapiro-Wilk test.

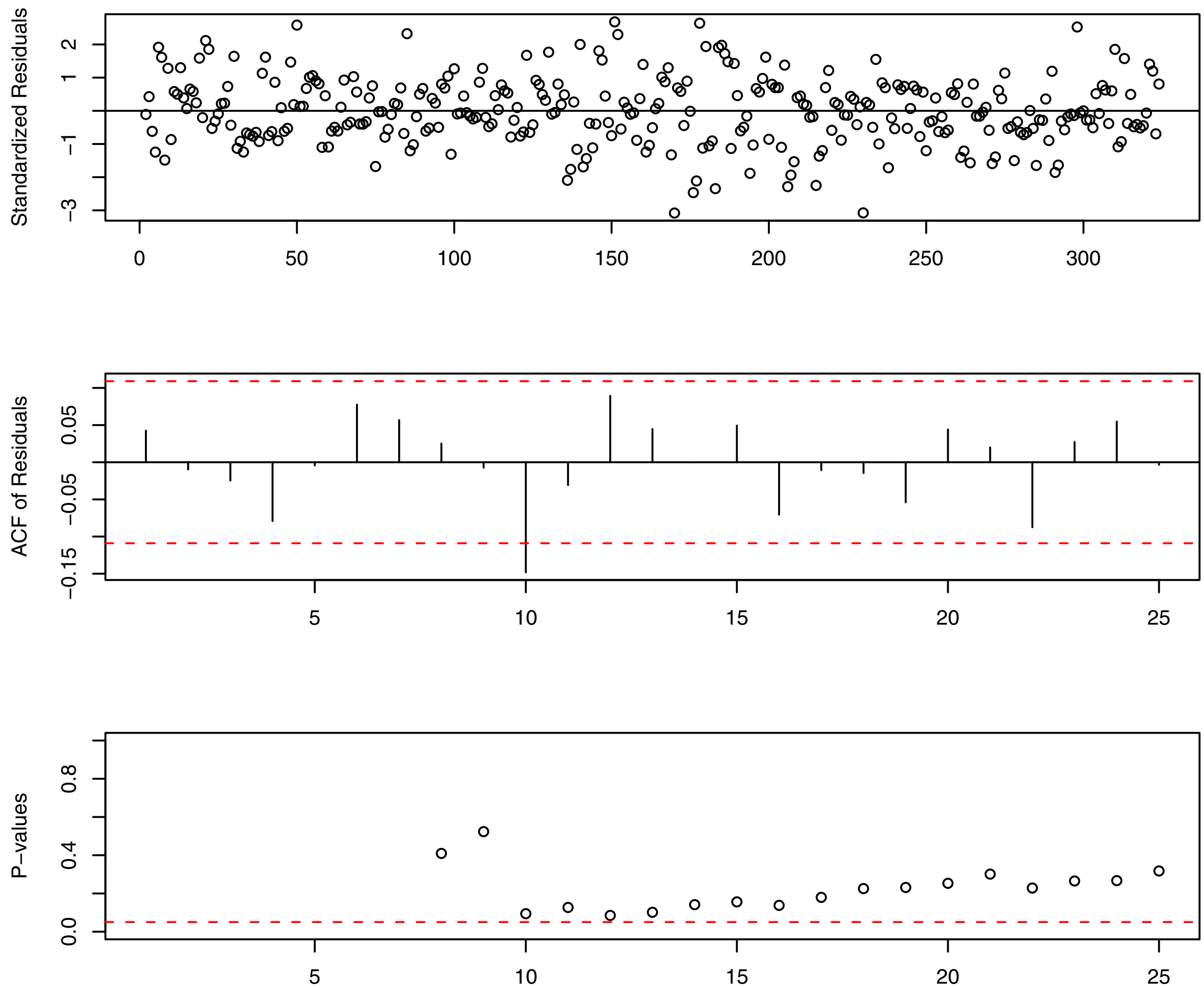
```
> shapiro.test(residuals(mod2))
```

```
        Shapiro-Wilk normality test

data:  residuals(mod2)
W = 0.9969, p-value = 0.7909
```

Normality looks like a viable assumption for the error terms in the IMA(1,1) model for the robot time series.

Finally, let's look at the results from the `tsdiag` command.

```
> win.graph(height=6,width=6.5,pointsize=8); tsdiag(mod2)
```



Summarizing: The robot time series seems to be well-represented by the IMA(1,1) model.

Exercise 8.10 The data file named deere3 contains 57 consecutive values from a complex machine tool at Deere & Co. The values given are deviations from a target value in units of ten millionths of an inch. The process employs a control mechanism that resets some of the parameters of the machine tool depending on the magnitude of deviation from target of the last item produced. Diagnose the fit of an AR(1) model for these data in terms of the tests discussed in this chapter.

```
> data(deere3); model=arima(deere3,order=c(1,0,0)); model
```

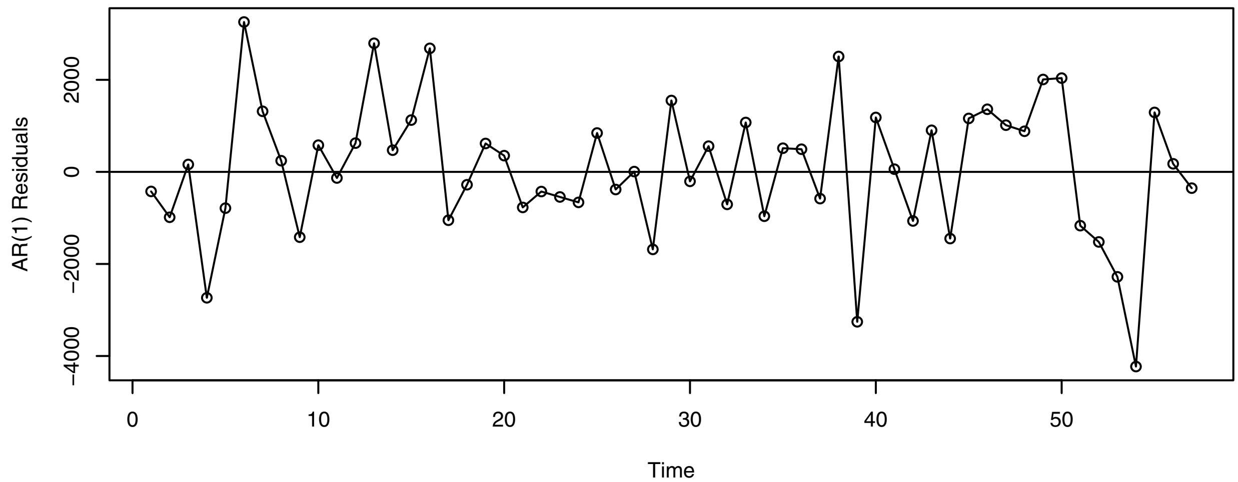
```
Call:
arima(x = deere3, order = c(1, 0, 0))

Coefficients:
         ar1  intercept
      0.5255   124.3832
s.e.  0.1108   394.2069

sigma^2 estimated as 2069355:  log likelihood = -495.51,  aic = 995.02
```

The ar1 parameter estimate (of ϕ) is statistically significant but the intercept could be removed from the model. This is not too surprising since the data are deviations from a target value. We will fit the model excluding a mean or intercept term.

```
> model=arima(deere3,order=c(1,0,0),include.mean=F); model
```

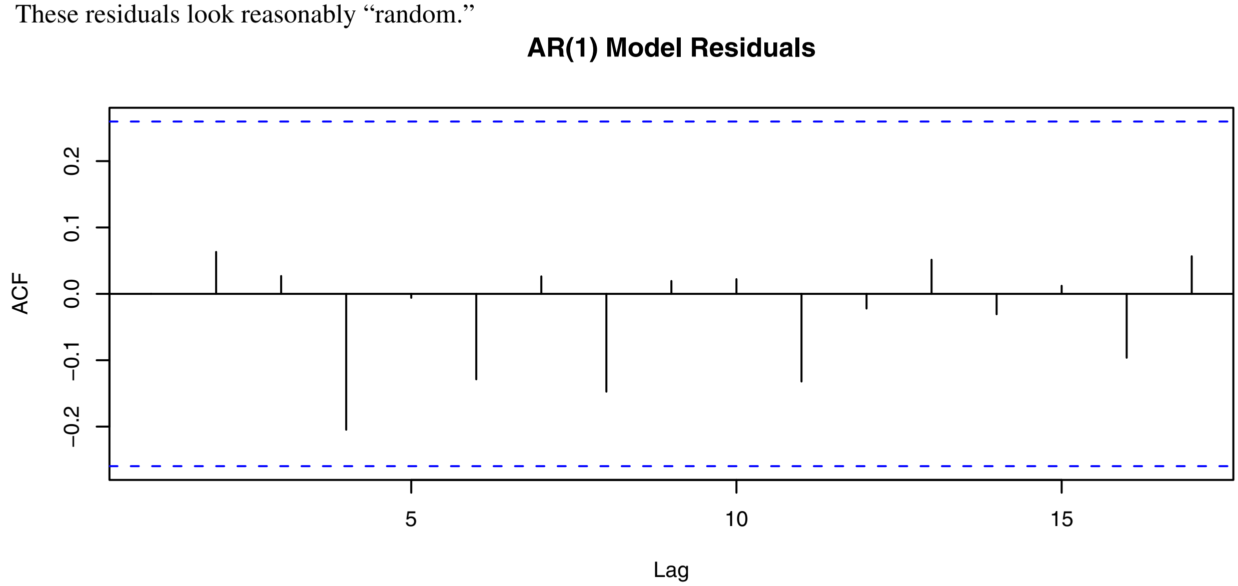
```
Call:
arima(x = deere3, order = c(1, 0, 0), include.mean = F)

Coefficients:
         ar1
      0.5291
s.e.  0.1103

sigma^2 estimated as 2072748:  log likelihood = -495.56,  aic = 993.12
```

There is very little change in the estimate of ϕ. The AIC is actually a little worse but we will continue our analysis with this simpler model.




```
> res=residuals(model); plot(res,ylab='AR(1) Residuals'); abline(h=0)
```

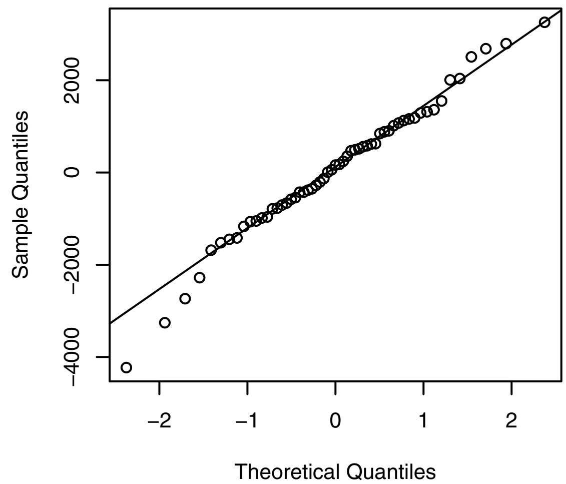
These residuals look reasonably “random.”




```
> acf(res,main='AR(1) Model Residuals')
```

There is little evidence of autocorrelation in the error terms for this model.

Next, on to normality.

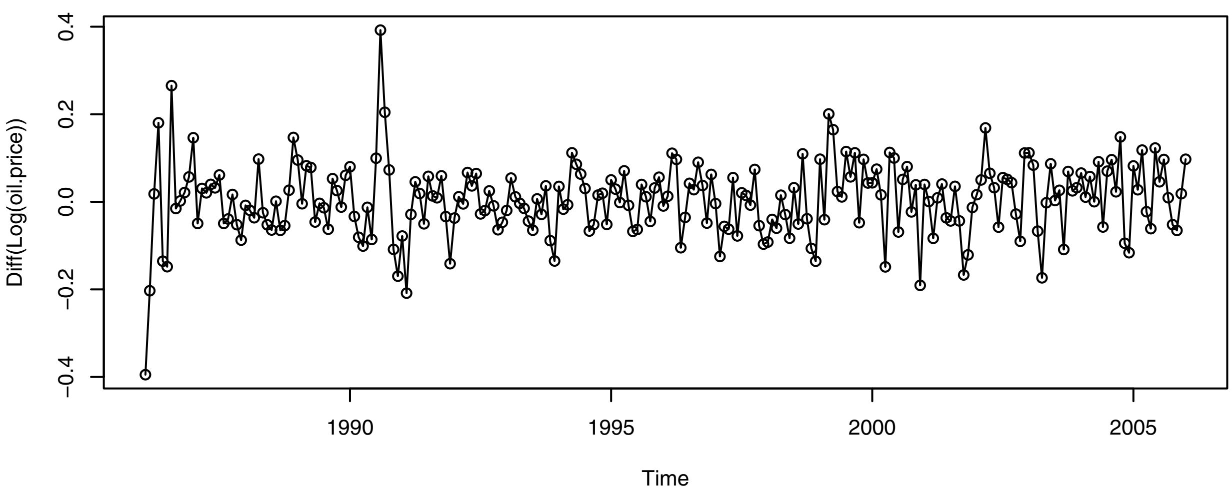



```
> qqnorm(res); qqline(res); shapiro.test(res)
```

```
        Shapiro-Wilk normality test

data:  res
W = 0.9829, p-value = 0.5966
```

The Q-Q plot shows some deviation froma straight line at the low end but the Shapiro-Wilk test is quite clear. We cannot reject normality for the error terms in this model. In summary, this series may be adequately modeled as an AR(1) series with no intercept (or mean) parameter and with uncorrelated, normal error terms.

Exercise 8.11 Exhibit (6.31), page 139, suggested specifying either an AR(1) or possibly an AR(4) model for the difference of the logarithms of the oil price series. (The filename is oil.price). For refernce we first plot the differences of the logs.




```
> plot(diff(log(oil.price)),type='o',ylab='Diff(Log(oil.price))')
```

The difference of the logarithms looks fairly stable except for possible outliers at the beginning (February 1986) and at August 1990.

(a) Estimate both of these models using maximum likelihood and compare the results using the diagnostic tests considered in this chapter.

```
> data(oil.price); mod1=arima(log(oil.price),order=c(1,1,0)); mod1
> mod2=arima(log(oil.price),order=c(4,1,0)); mod2
```

```
Call:
arima(x = log(oil.price), order = c(1, 1, 0))

Coefficients:
         ar1
      0.2364
s.e.  0.0660

sigma^2 estimated as 0.006787:  log likelihood = 258.55,  aic = -515.11

Call:
arima(x = log(oil.price), order = c(4, 1, 0))

Coefficients:
         ar1      ar2     ar3      ar4
      0.2673  -0.1550  0.0238  -0.0970
s.e.  0.0669   0.0691  0.0691   0.0681

sigma^2 estimated as 0.006603:  log likelihood = 261.82,  aic = -515.64
```

The ar3 and ar4 coefficients are not significant in the ARIMA(4,1,0) model and the AIC value is a tiny bit better in the simpler ARIMA(1,1,0) case. Furthermore, given the standard error of the ar1 coefficients, there is no real difference beteen the estimates of the ar1 coefficients in the two models. Let's try an ARIMA(2,1,0) model for comparison.

```
> mod3=arima(log(oil.price),order=c(2,1,0)); mod3
```

```
Call:
arima(x = log(oil.price), order = c(2, 1, 0))

Coefficients:
         ar1      ar2
      0.2630  -0.1436
s.e.  0.0666   0.0673

sigma^2 estimated as 0.00666:  log likelihood = 260.81,  aic = -517.61
```

This model has the best (smallest) AIC value of the three considered so far.

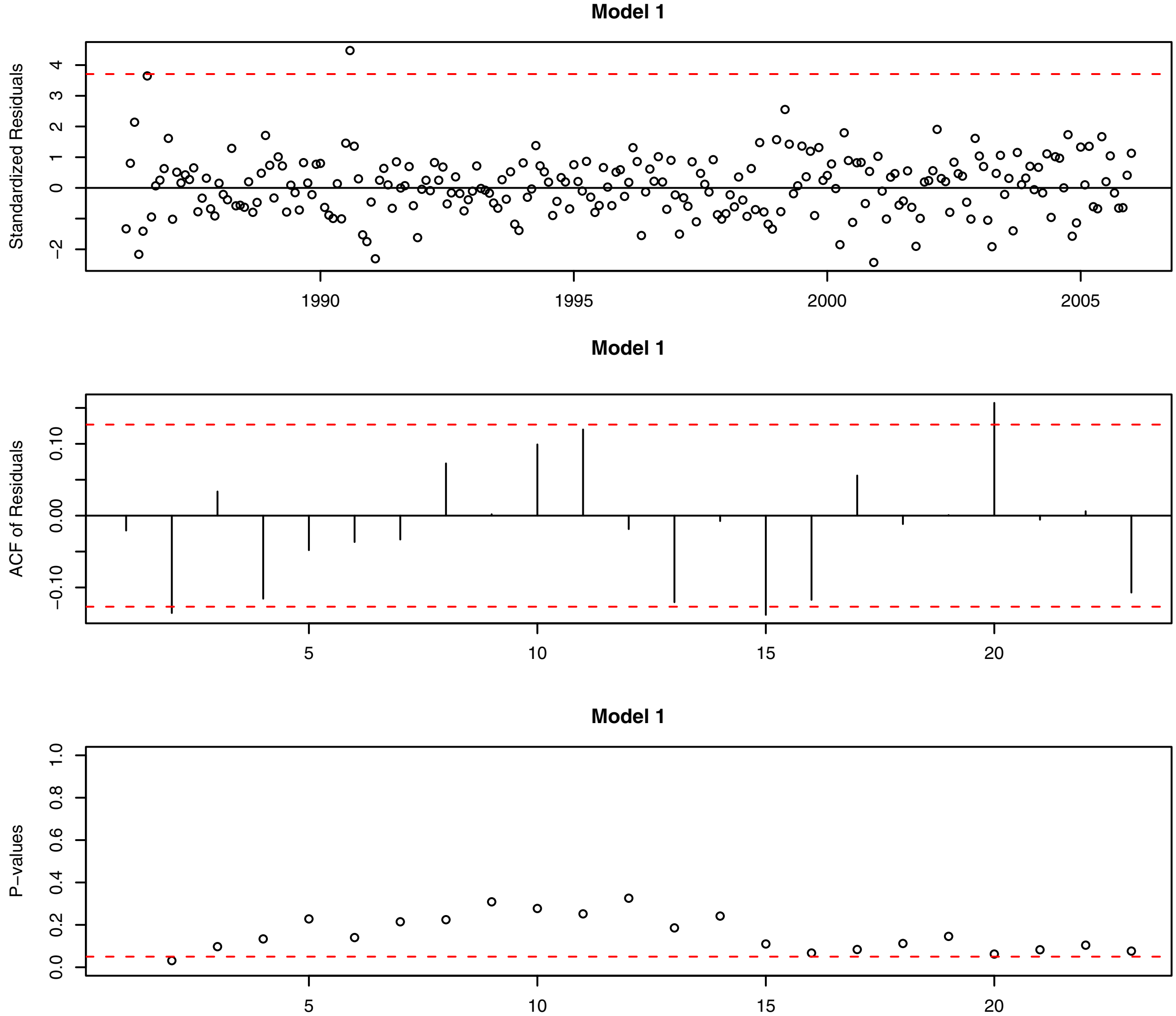
(b) Exhibit (6.32), page 140, suggested specifying an MA(1) model for the difference of the logs. Estimate this model by maximum likelihood and perform the diagnostic tests considered in this chapter.

```
> mod4=arima(log(oil.price),order=c(0,1,1)); mod4
```

```
Call:
arima(x = log(oil.price), order = c(0, 1, 1))

Coefficients:
         ma1
      0.2956
s.e.  0.0693

sigma^2 estimated as 0.006689:  log likelihood = 260.29,  aic = -518.58
```

This model has a significant ma1 coefficient and log-ikelihood and AIC values quite similar to the ARIMA(1,1,0) and ARIMA(0,1,1) models. This IMA(1,1) model does have the best AIC value. We will look at the diagnostics for these three models in part (c) before we decide which we prefer.

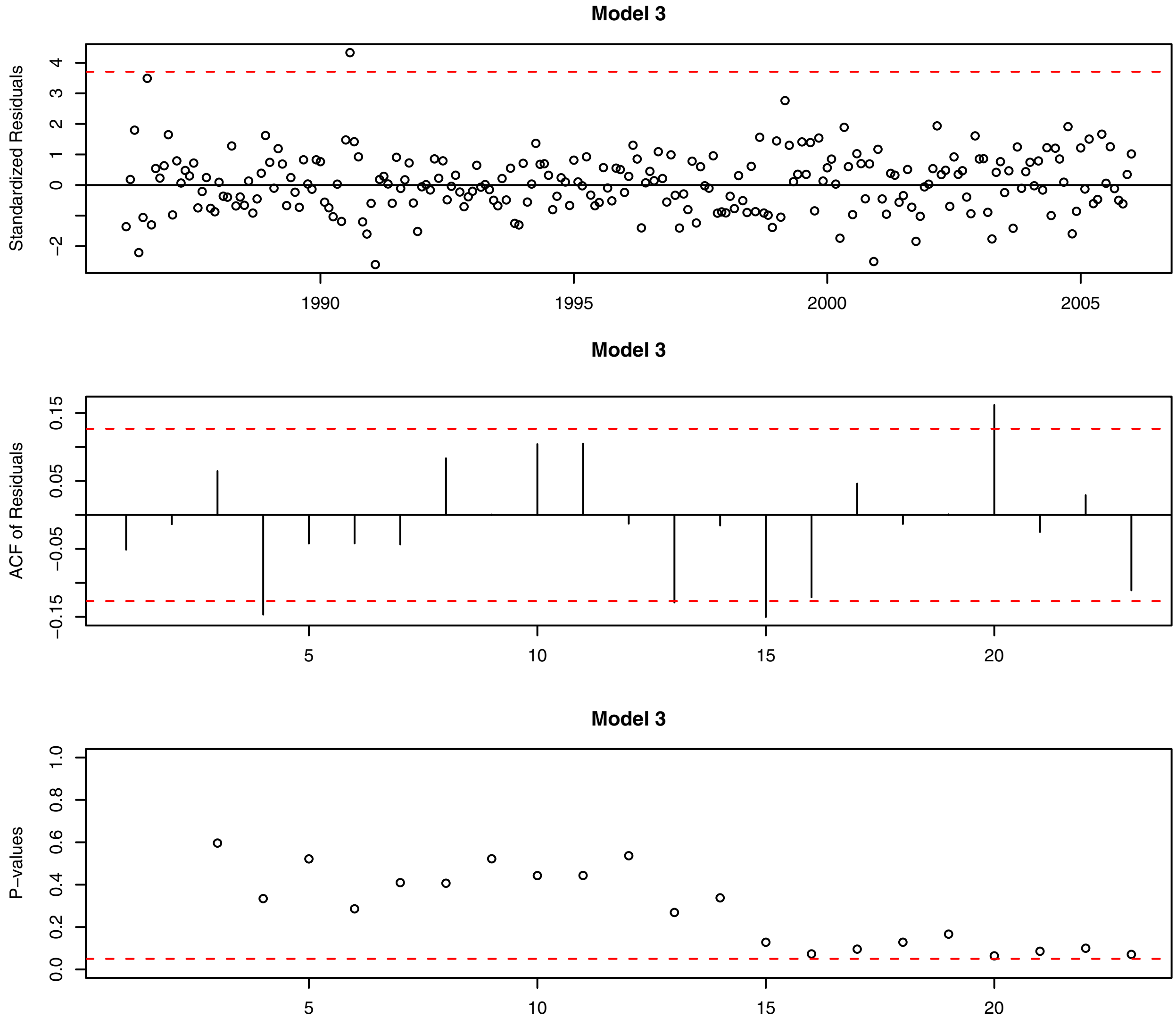
(c) Which of the three models AR(1), AR(4), or MA(1) would you prefer given the results of parts (a) and (b)?




```
> win.graph(width=6.5,height=6,pointsize=10); tsdiag(mod1,main='Model 1')
```

The possible outlier in August 1990 stands out in the plot of residuals and is “flagged” by the Bonferroni rule. There are also three residual acf values outside the critical limits.

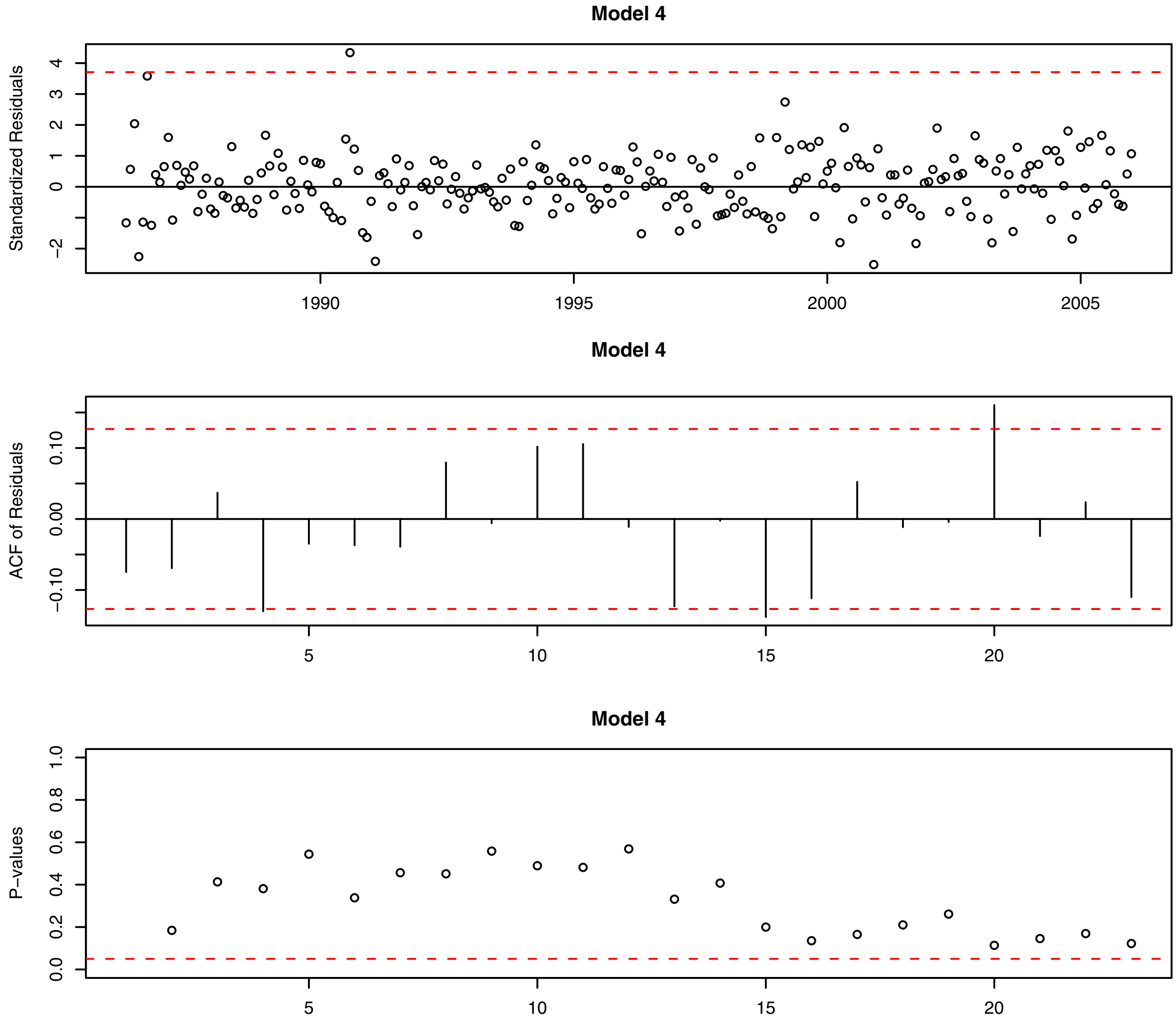
Let's look at similar results for Model 3.




```
> tsdiag(mod3,main='Model 3')
```

Model 3 diagnostics are similar to those for Model 1 with the exception that the Ljung-Box statistics are better as shown in the bottom display.

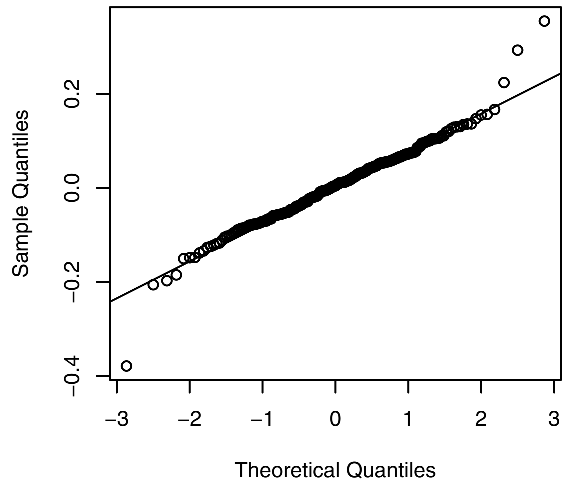
On to Model 4.




```
> tsdiag(mod4,main='Model 4')
```

Model 4 diagnostics are similar to those for Models 1 and 3. The Ljung-Box statistics are the best of the lot as shown in the bottom display.

Let's look at normality of the error terms for Model 4.




```
> win.graph(width=3,height=3,pointsize=8)
> qqnorm(residuals(mod4)); qqline(residuals(mod4))
> shapiro.test(residuals(mod4))
```

```
        Shapiro-Wilk normality test

data:  residuals(mod4)
W = 0.9688, p-value = 3.937e-05
```

Both the Q-Q plot and the results of the Shapiro-Wilk test indicate that we should reject of normality for the error terms in this model but this could be caused by the suspected outliers in the series. The possible outliers and the IM(1,1) for the logarithms is considered further in Exercise 11.21.

CHAPTER 9

Exercise 9.1 For an AR(1) model with $Y_t = 12.2$, $\phi = -0.5$, and $\mu = 10.8$,

(a) Find $\hat{Y}_t(1)$.

From Equation (9.3.6), page 193, $\hat{Y}_t(1) = \mu + \phi(Y_t - \mu) = 10.8 + (-0.5)(12.2 - 10.8) = 10.1$.

(b) Calculate $\hat{Y}_t(2)$ in two different ways.

From Equation (9.3.7), page 193, $\hat{Y}_t(2) = \mu + \phi[\hat{Y}_t(1) - \mu] = 10.8 + (-0.5)[10.1 - 10.8] = 11.15$.

Alternatively, from Equation (9.3.8), page 194, $\hat{Y}_t(2) = \mu + \phi^2(Y_t - \mu) = 10.8 + (-0.5)^2(12.2 - 10.8) = 11.15$.

(c) Calculate $\hat{Y}_t(10)$.

From Equation (9.3.8), page 194, $\hat{Y}_t(10) = \mu + \phi^{10}(Y_t - \mu) = 10.8 + (-0.5)^{10}(12.2 - 10.8) = 10.801367 \approx \mu$.

Exercise 9.2 Suppose that annual sales (in millions of dollars) of the Acme Corporation follow the AR(2) model $Y_t = 5 + 1.1Y_{t-1} - 0.5Y_{t-2} + e_t$ with $\sigma_e^2 = 2$.

(a) If sales for 2005, 2006, and 2007 were \$9 million, \$11 million, and \$10 million, respectively, forecast sales for 2008 and 2009.

From Equation (9.3.28), page 199, $\hat{Y}_{2007}(1) = 5 + 1.1Y_{2007} - 0.5Y_{2006} = 5 + 1.1(10) - 0.5(11) = 10.5$.

Furthermore, $\hat{Y}_{2007}(2) = 5 + 1.1\hat{Y}_{2008} - 0.5Y_{2007} = 5 + 1.1(10.5) - 0.5(10) = 11.55$.

(b) Show that $\psi_1 = 1.1$ for this model.

From Equations (4.3.21) on page 75, $\psi_1 - \phi_1\psi_0 = 0$ with $\psi_0 = 1$ so that $\psi_1 = \phi_1 = 1.1$.

(c) Calculate 95% prediction limits for your forecast in part (a) for 2008.

Using $Var(e_t(\ell)) = \sigma_e^2 \sum_{j=0}^{\ell-1} \psi_j^2$ from page 203, the prediction limits are $\hat{Y}_{2007}(1) \pm 2[\sqrt{\sigma_e^2}]$ or $10.5 \pm 2\sqrt{2}$ which is 10.5 ± 2.83. We are 95% confident that the 2008 value will be between 7.67 and 13.33.

(d) If sales in 2008 turn out to be \$12 million, update your forecast for 2009.

The updating Equation (9.6.1), page 207, is $\hat{Y}_{t+1}(\ell) = \hat{Y}_t(\ell+1) + \psi_\ell[Y_{t+1} - \hat{Y}_t(1)]$.

So we need to update as $\hat{Y}_{2008}(1) = \hat{Y}_{2007}(2) + \psi_1[Y_{2008} - \hat{Y}_{2007}(1)] = 11.55 + 1.1[12 - 10.5] = 13.2$.

Exercise 9.3 Using the estimated cosine trend on page 192:

(a) Forecast the average monthly temperature in Dubuque, Iowa, for April 1976.

From page 192 (or page 35), $\hat{\mu}_t = 46.2660 + (-26.7079)\cos(2\pi t) + (-2.1697)\sin(2\pi t)$. In this forecast, time is measured in years, with January 1964 as the starting value and a frequency of 1 per year. So the time value for April 1976 is 12 + (3/12) and the forecast is

$$\hat{\mu}_t = 46.2660 + (-26.7079)\cos\left(2\pi\left(12 + \frac{3}{12}\right)\right) + (-2.1697)\sin\left(2\pi\left(12 + \frac{3}{12}\right)\right) = 44.1\,°$$

(b) Find a 95% prediction interval for that April forecast. (The estimate of $\sqrt{\gamma_0}$ for this model is 3.719°F.)

The 95% prediction limits are $44.1 \pm 2(3.719)$ or 44.1 ± 7.4. Based on this model, we are 95% confident that the April 1976 average temperature in Dubuque, Iowa "will" be between 36.7°F and 51.5°F.

(c) What is the forecast for April, 1977? For April 2009?

With this deterministic model, all Aprils are forecast as the same value.

Exercise 9.4 Using the estimated cosine trend on page 192:

(a) Forecast the average monthly temperature in Dubuque, Iowa, for May 1976.

For May we use $t = 12 + (4/12)$ (see Exercise 9.3) and the forecast is

$$\hat{\mu}_t = 46.2660 + (-26.7079)\cos\left(2\pi\left(12 + \frac{4}{12}\right)\right) + (-2.1697)\sin\left(2\pi\left(12 + \frac{4}{12}\right)\right) = 57.7\text{ °F.}$$

(b) Find a 95% prediction interval for that May 1976 forecast. (The estimate of $\sqrt{\gamma_0}$ for this model is 3.719°F.)

The 95% prediction limits are $57.7 \pm 2(3.719)$ or 57.7 ± 7.4. Based on this model, we are 95% confident that the May 1976 average temperature in Dubuque, Iowa "will" be between 50.3°F and 65.1°F.

Exercise 9.5 Using the seasonal means model *without* an intercept shown in Exhibit (3.3), page 32:

(a) Forecast the average monthly temperature in Dubuque, Iowa, for April, 1976.

The forecast is just the coefficient estimate for all Aprils of 46.525 which we round to 46.5°F.

(b) Find a 95% prediction interval for that April forecast. (The estimate of $\sqrt{\gamma_0}$ for this model is 3.419°F.)

The 95% prediction interval is $46.5 \pm 2(3.419)$ or 46.5 ± 6.8. With this model, we are 95% confident that the April 1976 average temperature in Dubuque, Iowa "will" be between 39.7°F and 53.3°F.

(c) Compare your forecast with the one obtained in Exercise 9.3.

This model forecasts a slightly higher average temperature, 46.5° versus 44.1°, and it has a slightly narrower prediction interval: width = 6.8°F versus width = 7.4°F.

(d) What is the forecast for April 1977? April 2009?

With this model, all April forecasts are identical so the forecast is 46.5°F.

Exercise 9.6 Using the seasonal means model *with* an intercept shown in Exhibit (3.4), page 33:

(a) Forecast the average monthly temperature in Dubuque, Iowa, for April 1976.

With this model formulation the forecast is the Intercept + April coefficient so 16.608 + 29.917 = 46.525 or 46.5°F.

(b) Find a 95% prediction interval for that April forecast. (The estimate of $\sqrt{\gamma_0}$ for this model is 3.419°F.)

The 95% prediction interval is $46.5 \pm 2(3.419)$ or 46.5 ± 6.8. With this model, we are 95% confident that the April 1976 average temperature in Dubuque, Iowa “will” be between 39.7°F and 53.3°F.

(c) Compare your forecast with the one obtained in Exercise 9.5.

These two models will always produce identical forecasts and prediction intervals.

Exercise 9.7 Using the seasonal means model *with* an intercept shown in Exhibit (3.4), page 33

(a) Forecast the average monthly temperature in Dubuque, Iowa, for January 1976.

With this model, all Januarys are forecast with the estimated intercept, namely, 16.608 or 16.6°F.

(b) Find a 95% prediction interval for that January forecast. (The estimate of $\sqrt{\gamma_0}$ for this model is 3.419°F.)

The 95% prediction interval is $16.6 \pm 2(3.419)$ or 16.6 ± 6.8. With this model, we are 95% confident that the April 1976 average temperature in Dubuque, Iowa “will” be between 9.8 degrees and 23.4°F.

Exercise 9.8 Consider the monthly electricity generation time series shown in Exhibit (5.8), page 99. The data are in the file named electricity.

(a) Fit a deterministic trend model containing seasonal means together with a linear time trend to the logarithms of the electricity values.

```
> data(electricity); month.=season(electricity) # First method
> model=lm(log(electricity)~month.+time(electricity)); summary(model)
```

```
Call:
lm(formula = log(electricity) ~ month. + time(electricity))

Residuals:
       Min         1Q     Median         3Q        Max
-0.0962741 -0.0291892  0.0003147  0.0255065  0.1349765

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)       -3.783e+01  4.283e-01 -88.321  < 2e-16 ***
month.February    -1.246e-01  1.004e-02 -12.408  < 2e-16 ***
month.March       -9.080e-02  1.004e-02  -9.040  < 2e-16 ***
month.April       -1.642e-01  1.004e-02 -16.344  < 2e-16 ***
month.May         -1.000e-01  1.004e-02  -9.959  < 2e-16 ***
month.June        -2.016e-02  1.004e-02  -2.007   0.0455 *
month.July         7.675e-02  1.004e-02   7.641 1.75e-13 ***
month.August       7.368e-02  1.004e-02   7.335 1.33e-12 ***
month.September   -6.473e-02  1.004e-02  -6.444 3.49e-10 ***
month.October     -1.148e-01  1.005e-02 -11.431  < 2e-16 ***
month.November    -1.346e-01  1.005e-02 -13.400  < 2e-16 ***
month.December    -4.481e-02  1.005e-02  -4.460 1.08e-05 ***
time(electricity)  2.526e-02  2.153e-04 117.310  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0408 on 383 degrees of freedom
Multiple R-Squared: 0.9753,     Adjusted R-squared: 0.9746
F-statistic:  1262 on 12 and 383 DF,  p-value: < 2.2e-16
```

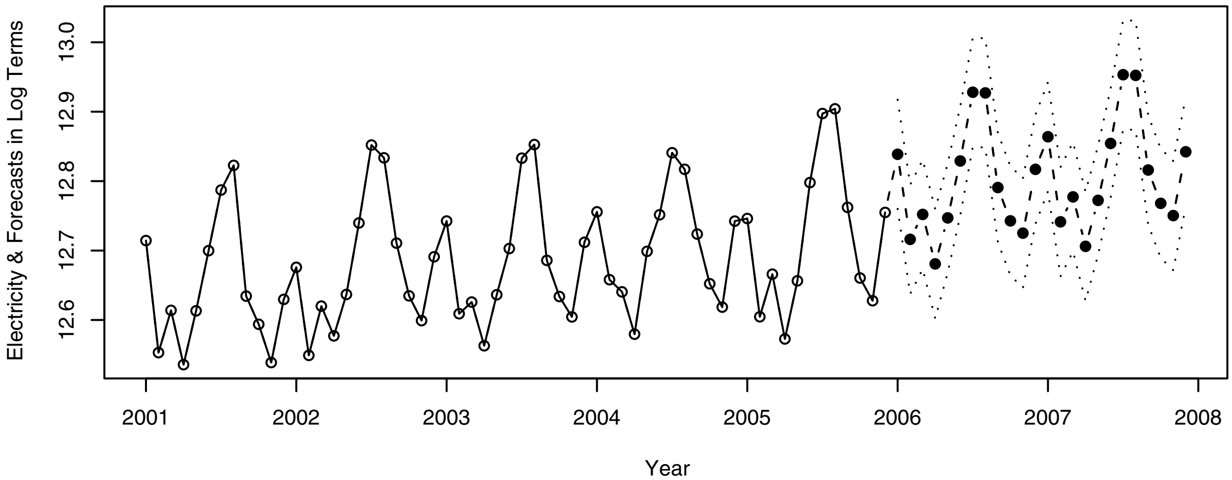
Notice that all coefficients are statistically significant. Alternative R code using the `arima` function for this calculation follows. This second method facillitates plotting of the forecasts.

```
> month.=season(electricity); time.trend=time(electricity) # Alternative formulation
> determ=as.matrix(model.matrix(~month.+time.trend-1))[,-1]
> model2=arima(log(electricity),order=c(0,0,0),xreg=determ); model2
```

```
Call:
arima(x = log(electricity), order = c(0, 0, 0), xreg = determ)

Coefficients:
      intercept  month.February  month.March  month.April  month.May
       -37.8299         -0.1246      -0.0908      -0.1642    -0.1000
```

```
s.e.      0.4212          0.0099       0.0099       0.0099     0.0099
      month.June  month.July  month.August  month.September  month.October
         -0.0202      0.0768        0.0737          -0.0647        -0.1148
s.e.      0.0099      0.0099        0.0099           0.0099         0.0099
      month.November  month.December  time.trend
             -0.1346         -0.0448      0.0253
s.e.          0.0099          0.0099      0.0002

sigma^2 estimated as 0.00161:  log likelihood = 711.56,  aic = -1397.11
```

(b) Plot the last five years of the series together with two years of forecasts and the 95% forecast limits. Interpret the plot.




```
> win.graph(width=6.5,height=3,pointsize=8)
> newmonth.=season(ts(rep(1,24),start=c(2006,1),freq=12))
> newtrend=time(electricity)[length(electricity)]+(1:24)*deltat(electricity)
> plot(model2,n.ahead=24,n1=c(2001,1),xlab='Year',pch=19,ylab='Electricity & Forecasts
   in Log Terms', newxreg=as.matrix(model.matrix(~newmonth.+newtrend-1))[,-1])
> # The second formulation using arima facillitates plotting the forecasts
```

The two years of forecasts mimic the upward trend and seasonal nature of the series quite well. The widths of the prediction intervals are reasonably narrow for these forecasts.

Exercise 9.9 Simulate an AR(1) process with $\phi = 0.8$ and $\mu = 100$. Simulate 48 values but set aside the last 8 values to compare forecasts to actual values.

```
> set.seed(132456); series=arima.sim(n=48,list(ar=0.8))+100
> future=window(series,start=41); series=window(series,end=40) # Set aside future
```

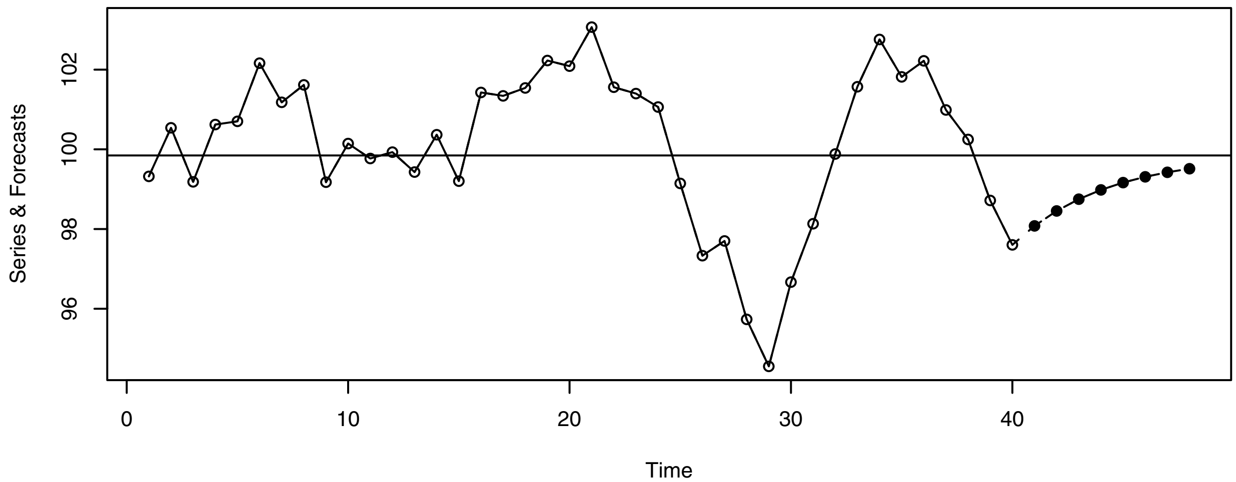
(a) Using the first 40 values of the series, find the values for the maximum likelihood estimates of ϕ and μ.

```
> model=arima(series,order=c(1,0,0)); model
```

```
Call:
arima(x = series, order = c(1, 0, 0))

Coefficients:
         ar1  intercept
      0.7878    99.8465
s.e.  0.0943     0.8110

sigma^2 estimated as 1.372:  log likelihood = -63.57,  aic = 131.14
```

The maximum likelihood estimates are quite accurate in this particular simulation.

(b) Using the estimated model, forecast the next eight values of the series. Plot the series together with the eight forecasts. Place a horizontal line at the estimate of the process mean.



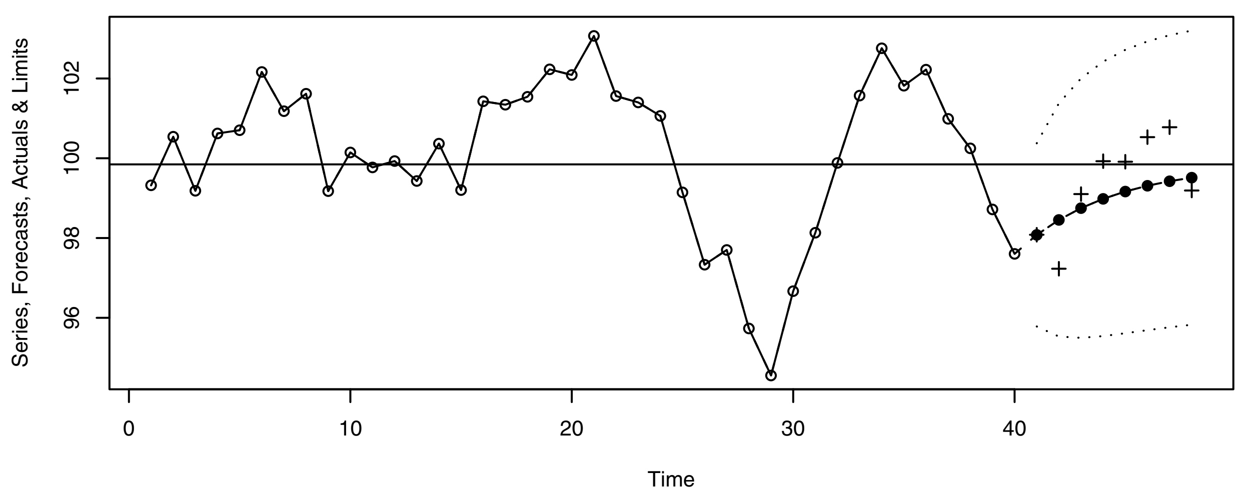

```
> plot(model,n.ahead=8,ylab='Series & Forecasts',col=NULL,pch=19)
> # col=NULL suppresses plotting the prediction intervals
> abline(h=coef(model)[names(coef(model))=='intercept'])
```

The forecasts are plotted as solid circles and have the chacteristic exponential decay toward the series mean.

(c) Compare the eight forecasts with the actual values that you set aside.

See part (d).

(d) Plot the forecasts together with 95% forecast limits. Do the actual values fall within the forecast limits?




```
> plot(model,n.ahead=8,ylab='Series, Forecasts, Actuals & Limits',pch=19)
> points(x=(41:48),y=future,pch=3) # Add the actual future values to the plot
> abline(h=coef(model)[names(coef(model))=='intercept'])
```

The actual future series values are plotted as plus signs (+) and they fall well within the forecast prediction limits shown as dotted lines. The forecasts, plotted as solid circles, are, of course, much smoother than the actual values.

(e) Repeat parts (a) through (d) with a new simulated series using the same values of the parameters and the same sample size.

Exercise 9.10 Simulate an AR(2) process with $\phi_1 = 1.5$, $\phi_2 = -0.75$, and $\mu = 100$. Simulate 52 values but set aside the last 12 values to compare forecasts to actual values.

```
> set.seed(132456); series=arima.sim(n=52,list(ar=c(1.5,-0.75)))+100
> actual=window(series,start=41); series=window(series,end=40)
```

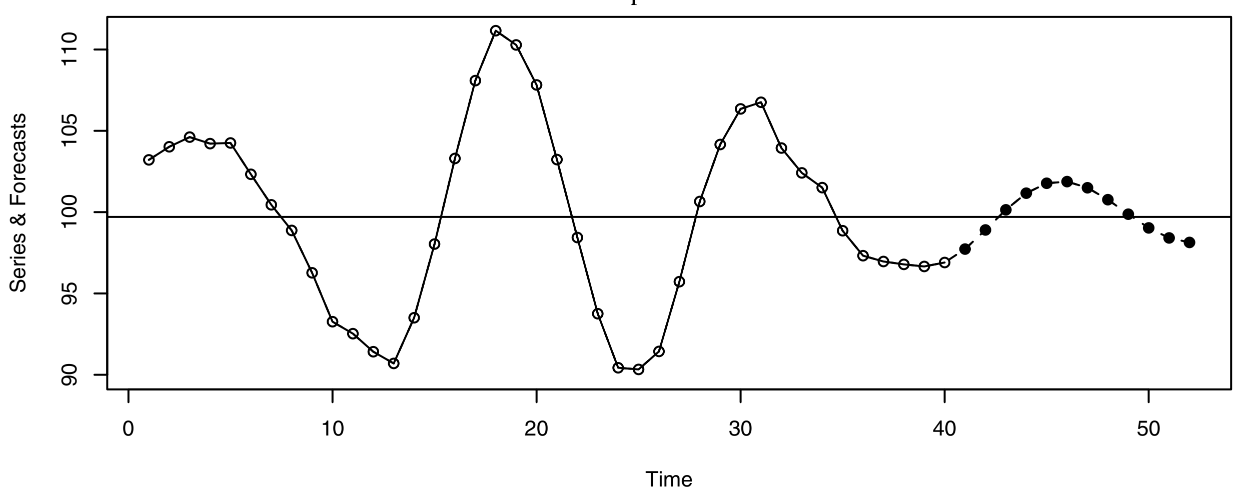
(a) Using the first 40 values of the series, find the values for the maximum likelihood estimates of the ϕ's and μ.

```
> model=arima(series,order=c(2,0,0)); model
```

```
Call:
arima(x = series, order = c(2, 0, 0))

Coefficients:
         ar1      ar2  intercept
      1.6874  -0.9043    99.7057
s.e.  0.0523   0.0509     0.7142

sigma^2 estimated as 0.9335:  log likelihood = -57.85,  aic = 121.71
```

Notice that for this simulation, the ar estimates are slightly more than two standard errors away from the "true" values.

(b) Using the estimated model, forecast the next 12 values of the series. Plot the series together with the 12 forecasts. Place a horizontal line at the estimate of the process mean.




```
> result=plot(model,n.ahead=12,ylab='Series & Forecasts',col=NULL,pch=19)
> abline(h=coef(model)[names(coef(model))=='intercept'])
```

The forecasts mimic the pseudo periodic nature of the series but also decay toward the series mean as they go further into the future.

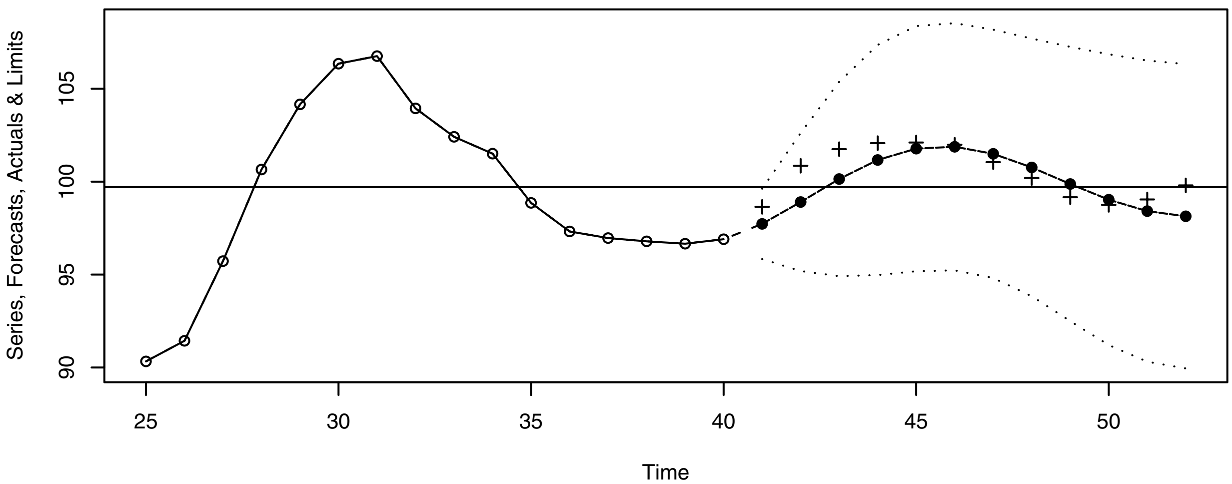
(c) Compare the 12 forecasts with the actual values that you set aside.

```
> forecast=result$pred; cbind(actual,forecast)
      Time Series:
      Start = 41
      End = 52
      Frequency = 1
            actual  forecasts
      41  98.64384  97.72871
      42 100.85458  98.90509
      43 101.74739 100.14264
      44 102.07439 101.16701
      45 102.10662 101.77635
      46 101.98472 101.87816
      47 101.04909 101.49890
      48 100.19580 100.76687
      49  99.16138  99.87465
```

```
50  98.74983  99.03113
51  99.04251  98.41467
52  99.79968  98.13729
```

Thye comparison is much easier to see in the plot in part (d).

(d) Plot the forecasts together with 95% forecast limits. Do the actual values fall within the forecast limits?




```
> plot(model,n1=25,n.ahead=12,ylab='Series, Forecasts, Actuals & Limits', pch=19)
> points(x=(41:52),y=future,pch=3) # future actual values as + signs
> abline(h=coef(model)[names(coef(model))=='intercept'])
```

The forecasts, plotted as solid circles, follow the same general “stochastic trend” as the actual future values (the pluses). The forecasts are well within the forecast prediction limits.

(e) Repeat parts (a) through (d) with a new simulated series using the same values of the parameters and same sample size.

Exercise 9.11 Simulate an MA(1) process with $\theta = 0.6$ and $\mu = 100$. Simulate 36 values but set aside the last 4 values to compare forecasts to actual values.

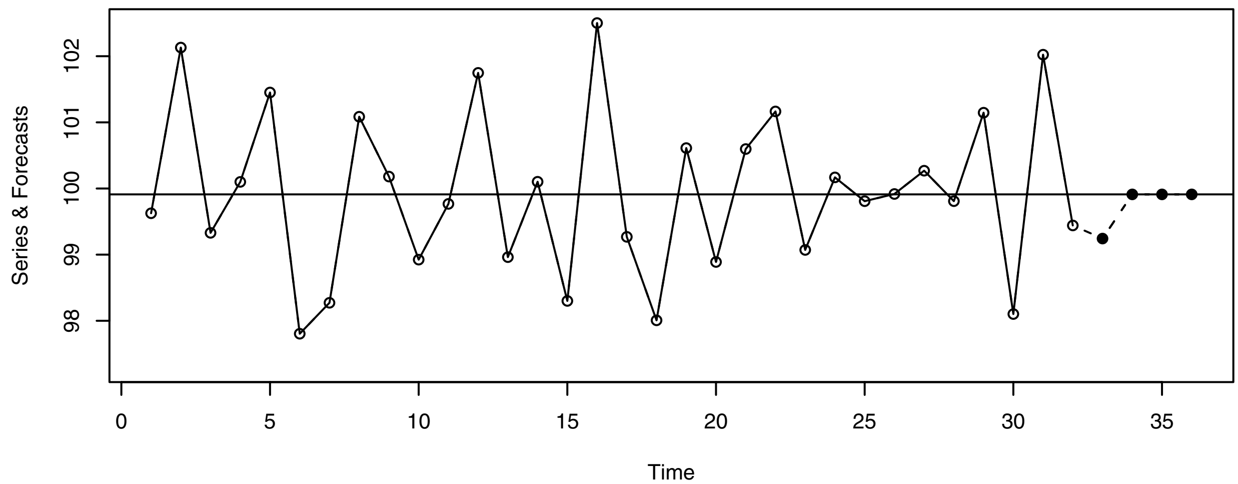
```
> set.seed(132456); series=arima.sim(n=52,list(ma=-0.6))+100
> actual=window(series,start=33); series=window(series,end=32)
```

(a) Using the first 32 values of the series, find the values for the maximum likelihood estimates of the θ and μ.

```
> model=arima(series,order=c(0,0,1)); model
```

```
Call:
arima(x = series, order = c(0, 0, 1))

Coefficients:
          ma1  intercept
      -0.8200    99.9109
s.e.   0.2066     0.0431

sigma^2 estimated as 1.001:  log likelihood = -45.99,  aic = 95.97
```

(b) Using the estimated model, forecast the next four values of the series. Plot the series together with the four forecasts. Place a horizontal line at the estimate of the process mean.




```
> result=plot(model,n.ahead=4,ylab='Series & Forecasts',col=NULL,pch=19)
> abline(h=coef(model)[names(coef(model))=='intercept'])
```

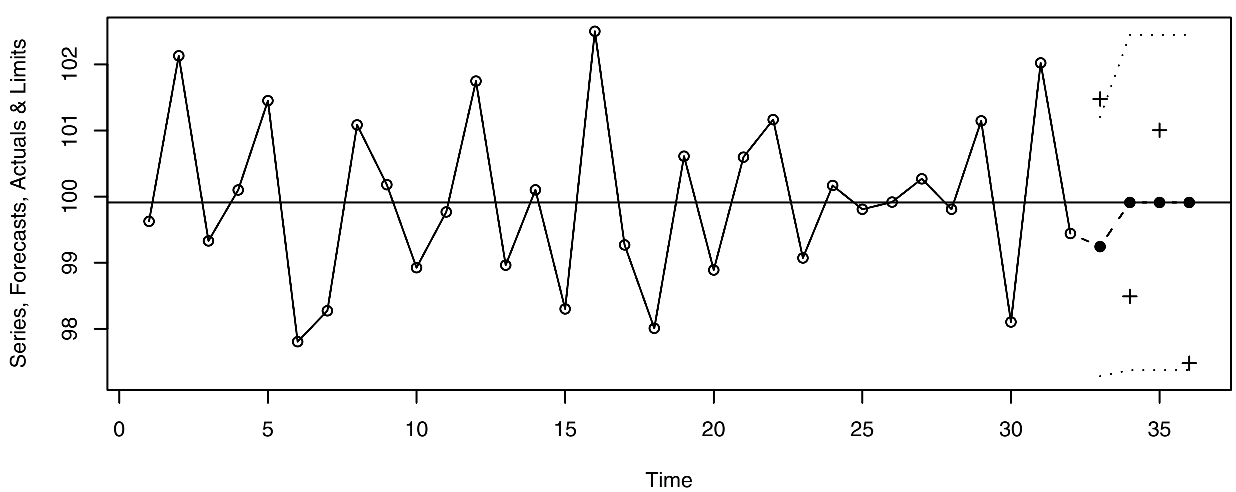
(c) Compare the four forecasts with the actual values that you set aside.

```
> forecast=result$pred; cbind(actual,forecast)
```

```
Time Series:
Start = 33
End = 36
Frequency = 1
      actual forecast
33 101.47674 99.24312
34  98.48984 99.91093
35 101.00347 99.91093
36  97.47793 99.91093
```

With the MA(1) model, only the lead 1 forecast is different from the process mean. There is little in this model to help improve the forecasts.

(d) Plot the forecasts together with 95% forecast limits. Do the actual values fall within the forecast limits?




```
> plot(model,n.ahead=4,ylab='Series, Forecasts, Actuals & Limits',pch=19)
> points(x=(33:36),y=actual,pch=3)
> abline(h=coef(model)[names(coef(model))=='intercept'])
```

The lead 1 forecast is outside the prediction limits for this simulation. The lead 4 forecast is nearly outside the limits and is quite extreme relative to the whole series. Of course, a new simulation would show different results.

(e) Repeat parts (a) through (d) with a new simulated series using the same values of the parameters and same sample size.

Exercise 9.12 Simulate an MA(2) process with $\theta_1 = 1$, $\theta_2 = -0.6$, and $\mu = 100$. Simulate 36 values but set aside the last 4 values with compare forecasts to actual values.

```
> set.seed(1432756); series=arima.sim(n=36,list(ma=c(-1,0.6)))+100
> actual=window(series,start=33); series=window(series,end=32)
```

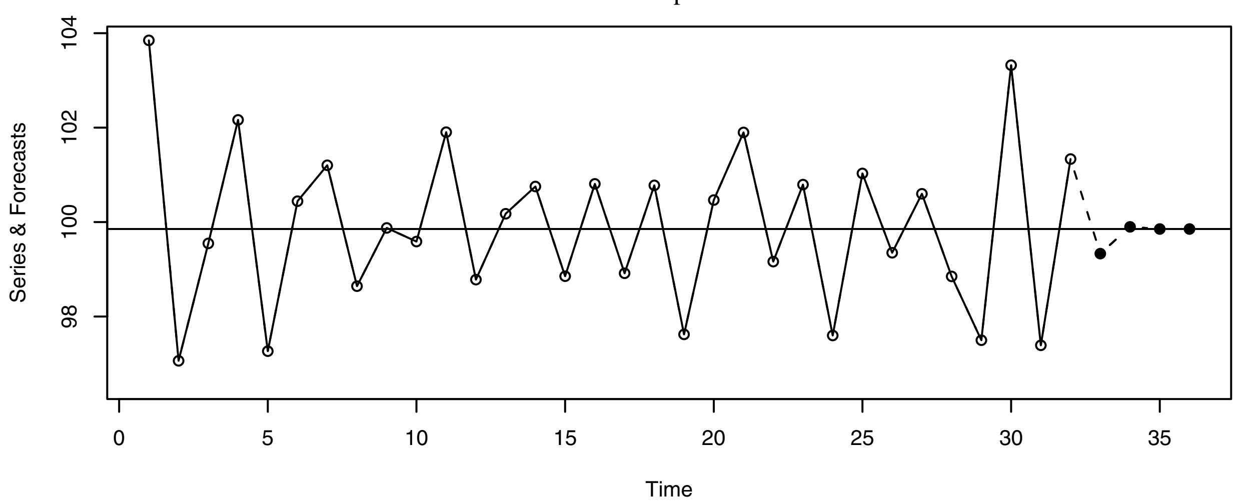
(a) Using the first 32 values of the series, find the values for the maximum likelihood estimates of the θ's and μ.

```
> model=arima(series,order=c(0,0,2)); model
```

```
Call:
arima(x = series, order = c(0, 0, 2))

Coefficients:
          ma1     ma2  intercept
      -1.0972  0.3257    99.8535
s.e.   0.1539  0.1684     0.0498

sigma^2 estimated as 1.234:  log likelihood = -49.46,  aic = 104.92
```

Taking the standard errors into account, the maximum likehood estimates are reasonably close to the true values in this simulation.

(b) Using the estimated model, forecast the next four values of the series. Plot the series together with the four forecasts. Place a horizontal line at the estimate of the process mean.




```
> result=plot(model,n.ahead=4,ylab='Series & Forecasts',col=NULL,pch=19)
> abline(h=coef(model)[names(coef(model))=='intercept'])
```

(c) What is special about the forecasts at lead times 3 and 4?

For the MA(2) model they are simply the estimated process mean.

(d) Compare the four forecasts with the actual values that you set aside.

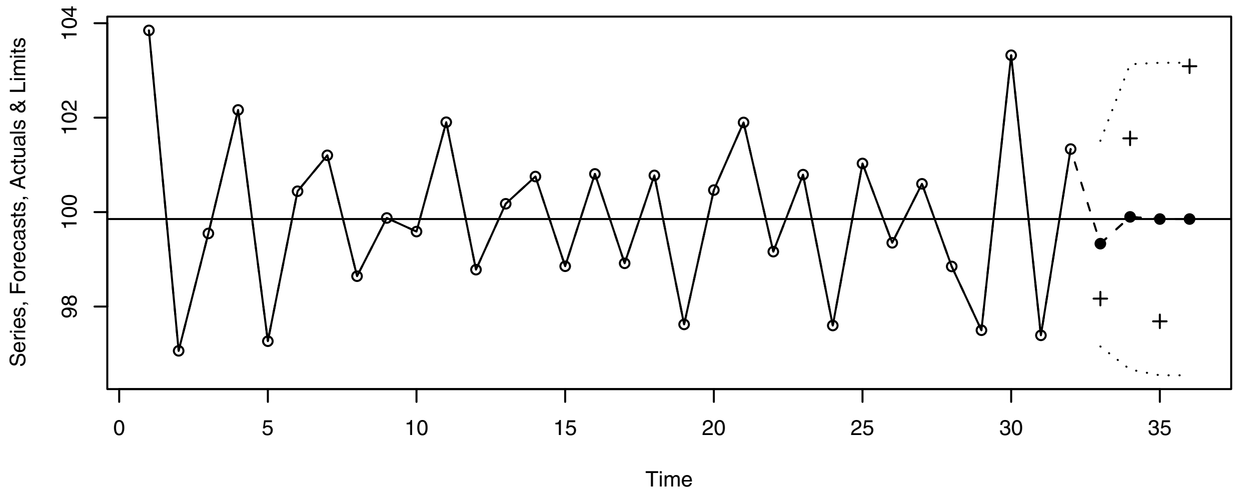
```
> forecast=result$pred; cbind(actual,forecast)
```

```
Time Series:
Start = 33
End = 36
Frequency = 1
      actual forecast
33  98.16822 99.33236
```

```
34 101.56150 99.89905
35  97.68750 99.85353
36 103.08874 99.85353
```

The comparison is best done with a graph. See part (e).

(e) Plot the forecasts together with 95% forecast limits. Do the actual values fall within the forecast limits?




```
> plot(model,n.ahead=4,ylab='Series, Forecasts, Actuals & Limits',type='o',pch=19)
> points(x=(33:36),y=actual,pch=3)
> abline(h=coef(model)[names(coef(model))=='intercept'])
```

For this simulation and this model, the forecasts are rather far from the actual values. However, the actuals are all within the forecast limits.

(f) Repeat parts (a) through (e) with a new simulated series using the same values of the parameters and same sample size.

Exercise 9.13 Simulate an ARMA(1,1) process with $\phi = 0.7$, $\theta = -0.5$, and $\mu = 100$. Simulate 50 values but set aside the last 10 values to compare forecasts with actual values.

```
> set.seed(172456); series=arima.sim(n=50,list(ar=0.7,ma=0.5))+100
> actual=window(series,start=41); series=window(series,end=40)
```

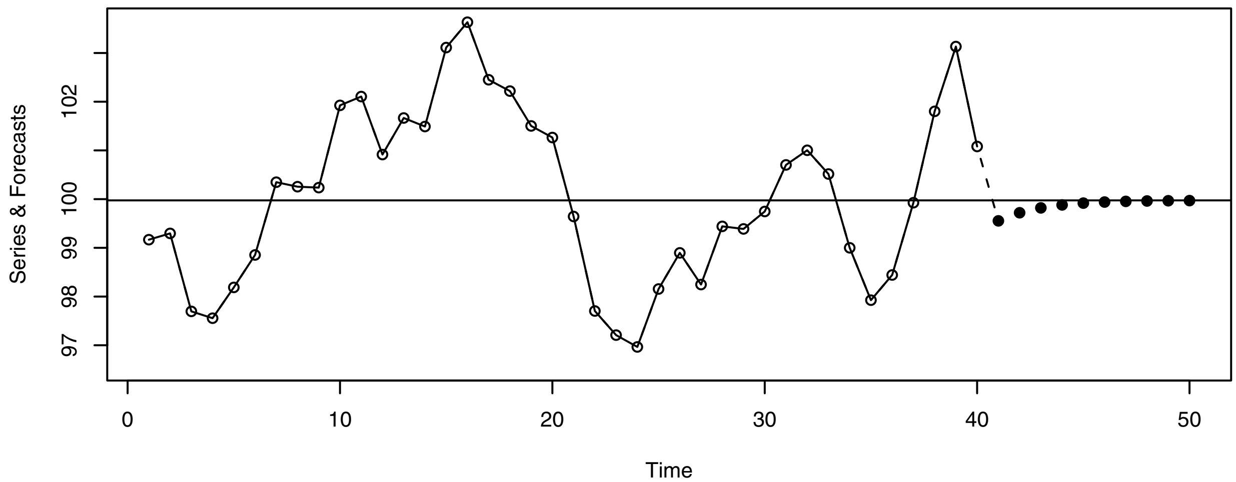
(a) Using the first 40 values of the series, find the values for the maximum likelihood estimates of ϕ, θ, and μ.

```
> model=arima(series,order=c(1,0,1)); model
```

```
Call:
arima(x = series, order = c(1, 0, 1))

Coefficients:
         ar1     ma1  intercept
      0.6048  0.6907    99.9745
s.e.  0.1585  0.2522     0.5846

sigma^2 estimated as 0.8162:  log likelihood = -53.6,  aic = 113.19
```

Taking the standard errors into account, the maximum likehood estimates are reasonably close to the true values in this simulation.

(b) Using the estimated model, forecast the next ten values of the series. Plot the series together with the ten forecasts. Place a horizontal line at the estimate of the process mean.




```
> result=plot(model,n.ahead=10,ylab='Series & Forecasts',col=NULL,pch=19)
> abline(h=coef(model)[names(coef(model))=='intercept'])
```

The forecfasts approach the series mean fairly quickly.

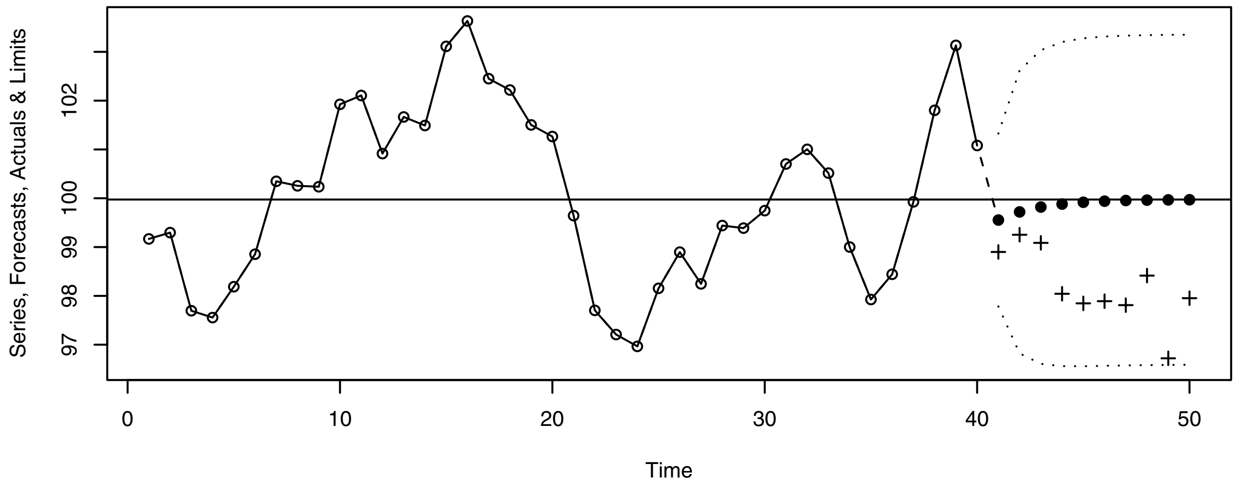
(c) Compare the ten forecasts with the actual values that you set aside.

```
> forecast=result$pred; cbind(actual,forecast)
```

```
Time Series:
Start = 41
End = 50
Frequency = 1
     actual forecast
41 98.90034 99.55443
42 99.25304 99.72043
43 99.08626 99.82082
44 98.04358 99.88154
45 97.84692 99.91826
46 97.89159 99.94047
47 97.81065 99.95391
48 98.41574 99.96203
49 96.72142 99.96694
50 97.95263 99.96992
```

See part (d) for a graphical comparison.

(d) Plot the forecasts together with 95% forecast limits. Do the actual values fall within the forecast limits?




```
> plot(model,n.ahead=10,ylab='Series, Forecasts, Actuals & Limits',pch=19)
> points(x=(41:50),y=actual,pch=3)
> abline(h=coef(model)[names(coef(model))=='intercept'])
```

This series is quite erratic but the actual series values are contained within the forecast limits. The forecasts decay to the estimated process mean rather quickly and the prediction limits are quite wide.

(e) Repeat parts (a) through (d) with a new simulated series using the same values of the parameters and same sample size.

Exercise 9.14 Simulate an IMA(1,1) process with $\theta = 0.8$ and $\theta_0 = 0$. Simulate 35 values, but set aside the last five values to compare forecasts with actual values.

```
> set.seed(127456); series=arima.sim(n=35,list(order=c(0,1,1),ma=-0.8))[-1]
> # delete initial term as it is always = 0
> actual=window(series,start=31); series=window(series,end=30)
```

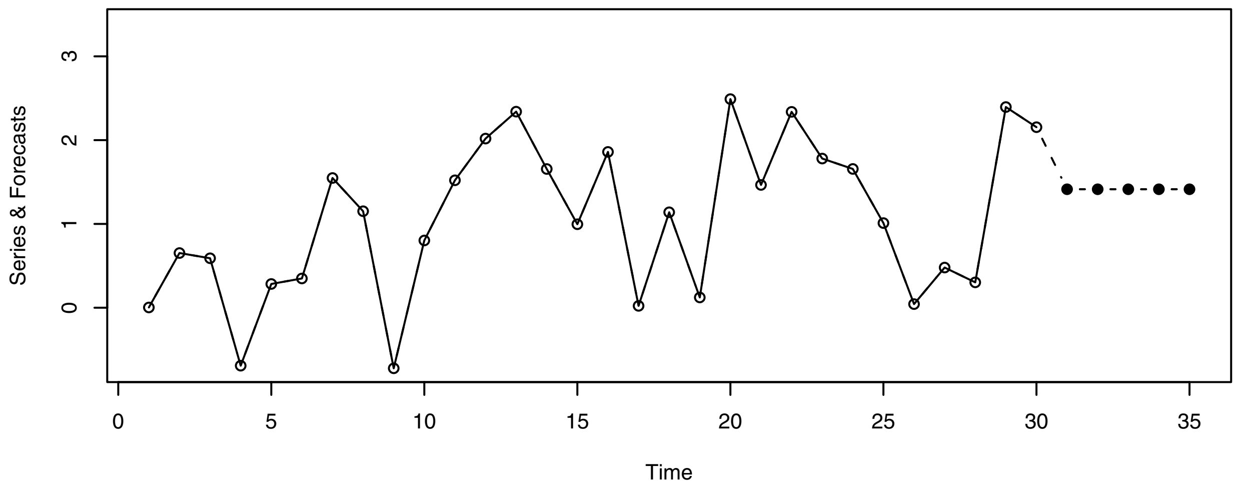
(a) Using the first 30 values of the series, find the value for the maximum likelihood estimate of θ.

```
> model=arima(series,order=c(0,1,1)); model
```

```
Call:
arima(x = series, order = c(0, 1, 1))

Coefficients:
          ma1
      -0.7696
s.e.   0.1832

sigma^2 estimated as 0.845:  log likelihood = -39.15,  aic = 80.31
```

Taking the standard errors into account, the maximum likehood estimate is quite close to the true value in this simulation.

(b) Using the estimated model, forecast the next five values of the series. Plot the series together with the five forecasts. What is special about the forecasts?




```
> result=plot(model,n.ahead=5,ylab='Series & Forecasts',col=NULL,pch=19)
```

For this model the forecasts are constant for all lead times.

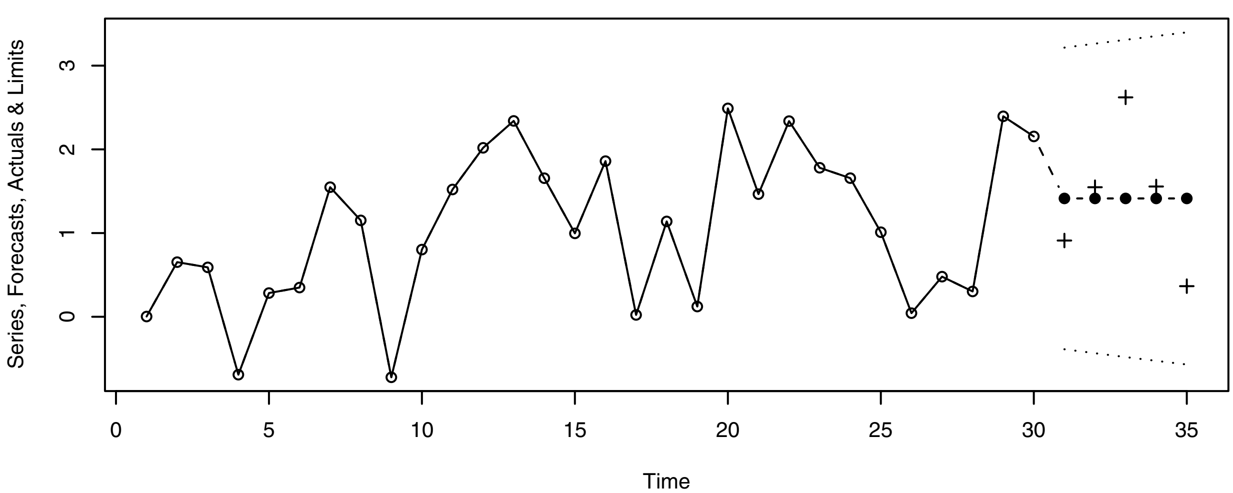
(c) Compare the five forecasts with the actual values that you set aside.

```
> forecast=result$pred; cbind(actual,forecast)
```

```
Time Series:
Start = 31
End = 35
Frequency = 1
      actual  forecast
31 0.9108642 1.413627
32 1.5476147 1.413627
33 2.6211930 1.413627
34 1.5560880 1.413627
35 0.3657792 1.413627
```

For this model the foecasts are the same at all lead times. See part (d) for a graphical comparison.

(d) Plot the forecasts together with 95% forecast limits. Do the actual values fall within the forecast limits?




```
> plot(model,n.ahead=5,ylab='Series, Forecasts, Actuals & Limits',pch=19)
> points(x=(31:35),y=actual,pch=3)
```

The forecat limits contain all of the actual values but they are quite wide.

(e) Repeat parts (a) through (d) with a new simulated series using the same values of the parameters and same sample size.

Exercise 9.15 Simulate an IMA(1,1) process with $\theta = 0.8$ and $\theta_0 = 10$. Simulate 35 values, but set aside the last five values to compare forecasts to actual values.

```
> set.seed(1245637); series=arima.sim(n=35,list(order=c(0,1,1),ma=-0.8))[-1]
> # delete initial term that is always zero
> series=series+10*(1:35) # add in intercept term that becomes a linear time trend
> actual=window(series,start=31); series=window(series,end=30)
```

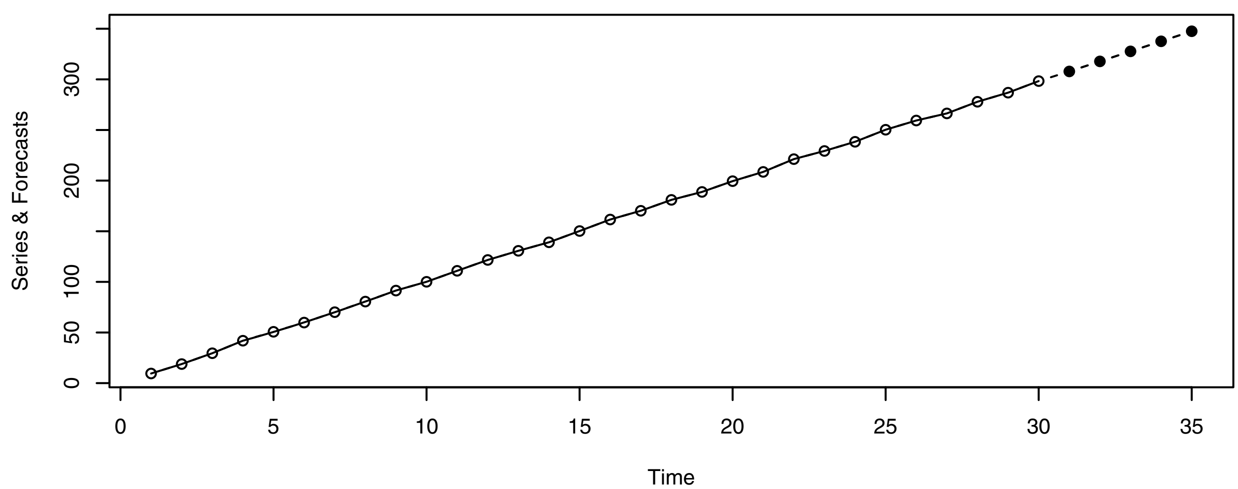
(a) Using the first 30 values of the series, find the values for the maximum likelihood estimates of θ and θ_0.

```
> model=arima(series,order=c(0,1,1),xreg=(1:30)); model
```

```
Call:
arima(x = series, order = c(0, 1, 1), xreg = (1:30))

Coefficients:
          ma1    xreg
      -0.7178  9.2545
s.e.   0.1389  0.6751

sigma^2 estimated as 135.7:  log likelihood = -112.71,  aic = 229.41
```

Both the ma and intercept parameters are estimated well in this simulation.

(b) Using the estimated model, forecast the next five values of the series. Plot the series together with the five forecasts. What is special about these forecasts?




```
> result=plot(model,n.ahead=5,newxreg=(31:35),
   ylab='Series & Forecasts',col=NULL,pch=19)
```

In this model the nonzero intercept term dominates the time series and the series and forecasts virually follow a straight line.

(c) Compare the five forecasts with the actual values that you set aside.

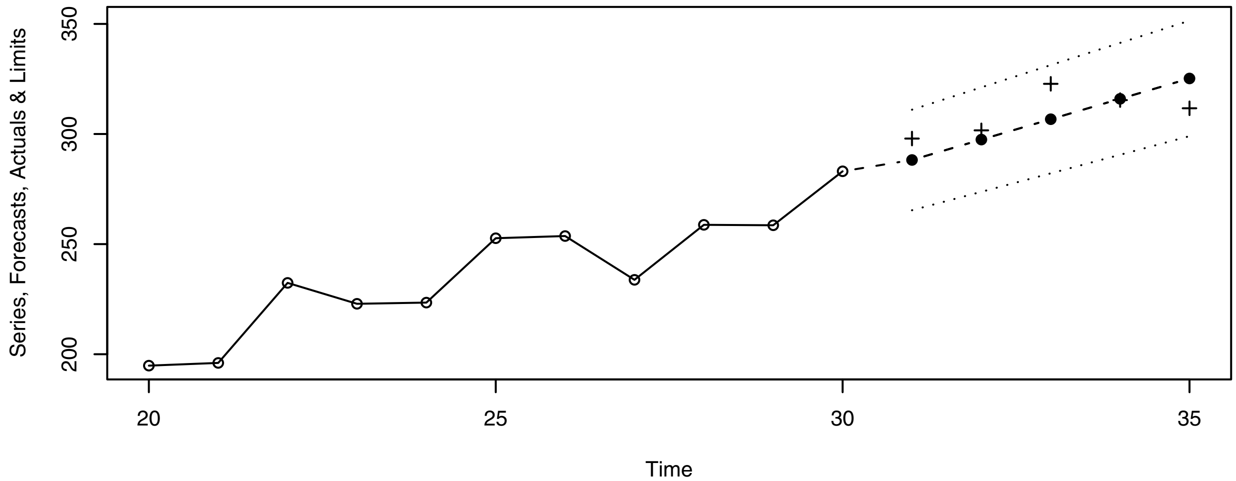
```
> forecast=result$pred; cbind(actual,forecast)
```

```
Time Series:
Start = 31
End = 35
Frequency = 1
     actual forecast
31 308.7936 307.8215
32 318.1666 317.7470
33 329.2790 327.6724
34 337.5403 337.5979
```

```
35 346.1691 347.5233
```

The comparison is seen best in the graph in part (d).

(d) Plot the forecasts together with 95% forecast limits. Do the actual values fall within the forecast limits?




```
> plot(model,n1=20,n.ahead=5,newxreg=(31:35),ylab='Series, Forecasts, Actuals &
   Limits',pch=19)
> points(x=seq(31,35),y=actual,pch=3) # Actuals plotted as plus signs
```

The actual values fall within the forecast limits. We plotted just the end of the series so we could see more detail.

(e) Repeat parts (a) through (d) with a new simulated series using the same values of the parameters and same sample size.

Exercise 9.16 Simulate an IMA(2,2) process with $\theta_1 = 1$, $\theta_2 = -0.75$, and $\theta_0 = 0$. Simulate 45 values, but set aside the last five values to compare forecasts with actual values.

```
> set.seed(432456)
> series=(arima.sim(n=45,list(order=c(0,2,2),ma=c(-1,0.75)))[-1])[-1]
> # Delete first two values as they are always zero
> actual=window(series,start=41); series=window(series,end=40)
```

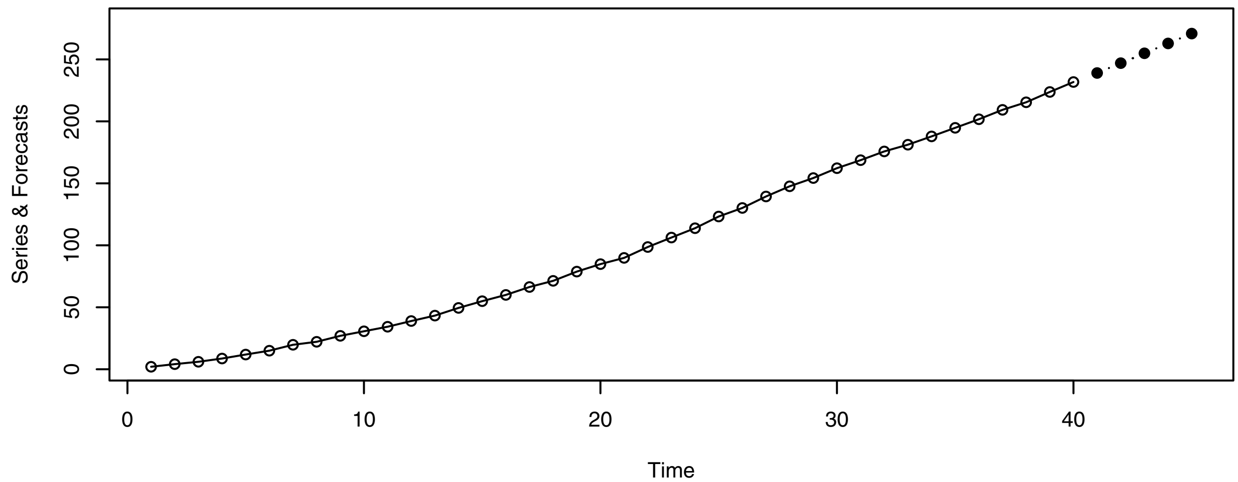
(a) Using the first 40 values of the series, find the value for the maximum likelihood estimate of θ_1 and θ_2.

```
> model=arima(series,order=c(0,2,2)); model
```

```
Call:
arima(x = series, order = c(0, 2, 2))

Coefficients:
          ma1     ma2
      -0.9821  0.7264
s.e.   0.1619  0.2254

sigma^2 estimated as 0.9944:  log likelihood = -54.76,  aic = 113.52
```

For this simulation the maximum likelihood estimates are excellent.

(b) Using the estimated model, forecast the next five values of the series. Plot the series together with the five forecasts. What is special about the forecasts?




```
> result=plot(model,n.ahead=5,ylab='Series & Forecasts',col=NULL,pch=19)
```

The forecasts seem to follow a straight line.

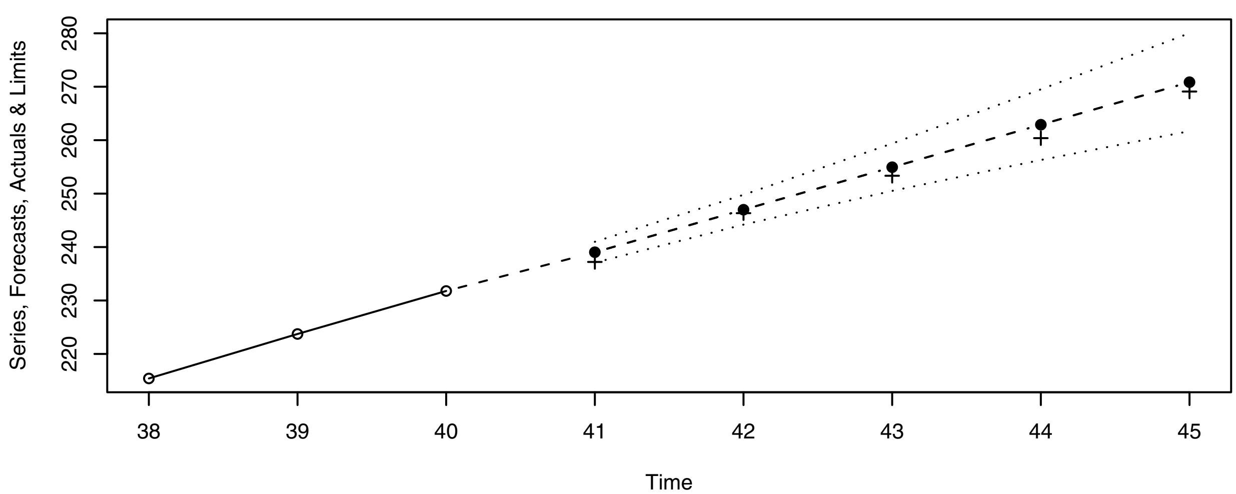
(c) Compare the five forecasts with the actual values that you set aside.

```
> forecast=result$pred; cbind(actual,forecast)
```

```
Time Series:
Start = 41
End = 45
Frequency = 1
     actual forecast
41 237.2138 239.0336
42 246.3410 246.9881
43 253.3401 254.9425
44 260.3783 262.8970
45 269.1054 270.8514
```

All of the forecasts are a bit higher than the actual values. See part (d) below for more comparisons.

(d) Plot the forecasts together with 95% forecast limits. Do the actual values fall within the forecast limits?




```
> plot(model,n1=38,n.ahead=5,ylab='Series, Forecasts, Actuals & Limits', pch=19)
> points(x=seq(41,45),y=actual,pch=3)
```

The forecast limits spread out as the lead time increases. This, of course, is characteristic of forecast limits for nonstationary models. The forecast at lead 1 is very close to the lower forecast limit. To check it more precisely, we display the numbers.

```
> lower=result$lpi; upper=result$upi; cbind(lower,actual,upper)
```

```
Time Series:
Start = 41
End = 45
Frequency = 1
      lower   actual    upper
41 237.0792 237.2138 240.9881
42 244.1991 246.3410 249.7770
43 250.5107 253.3401 259.3743
44 256.2908 260.3783 269.5031
45 261.6857 269.1054 280.0171
```

From the detailed numbers, we see that in fact the lead 1 forecast is a little above the lower forecast limit so that all of the forecasts are within the 95% limits in this simulation.

(e) Repeat parts (a) through (d) with a new simulated series using the same values of the parameters and same sample size.

Exercise 9.17 Simulate an IMA(2,2) process with $\theta_1 = 1$, $\theta_2 = -0.75$, and $\theta_0 = 10$. Simulate 45 values, but set aside the last five values to compare forecasts with actual values.

(Note: An intercept value of, say, $\theta_0 = 0.03$ provides a more reasonable looking series.)

```
> set.seed(13243546)
> series=arima.sim(n=45,list(order=c(0,2,2),ma=c(-1,0.75)))[-(1:2)]
> # Delete initial zero values
> series=series+10*(1:45)^2 # Add nonzero intercept term
> actual=window(series,start=41); series=window(series,end=40)
```

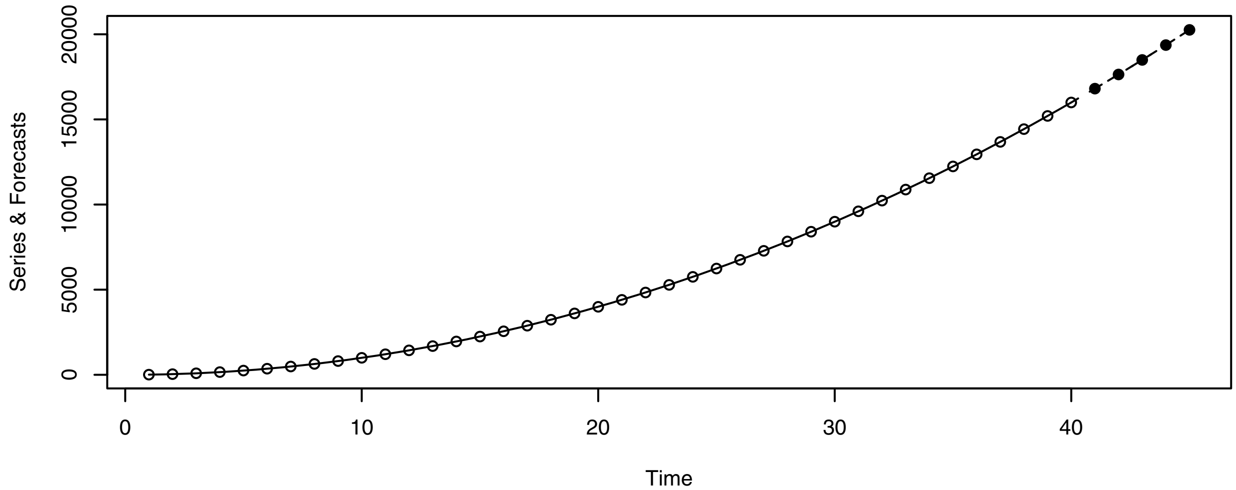
(a) Using the first 40 values of the series, find the values for the maximum likelihood estimates of θ_1, θ_2, and θ_0.

```
> model=arima(series,order=c(0,2,2),xreg=(1:40)^2); model
```

```
Call:
arima(x = series, order = c(0, 2, 2), xreg = (1:40)^2)

Coefficients:
          ma1     ma2     xreg
      -1.0382  0.6638  10.0338
s.e.   0.1123  0.1435   0.0551

sigma^2 estimated as 1.202:  log likelihood = -58.24,  aic = 122.48
```

Taking the standard errors into consideration, the estimates of the MA parameters and intercept are quite reasonable.

(b) Using the estimated model, forecast the next five values of the series. Plot the series together with the five forecasts. What is special about these forecasts?




```
> result=plot(model,n.ahead=5,newxreg=(41:45)^2,ylab='Series & Forecasts',
   type='o',col=NULL,pch=19)
```

They look like they fall on a straight line. In fact, we know that it is really a quadratic curve but over this range the slope is nearly constant.

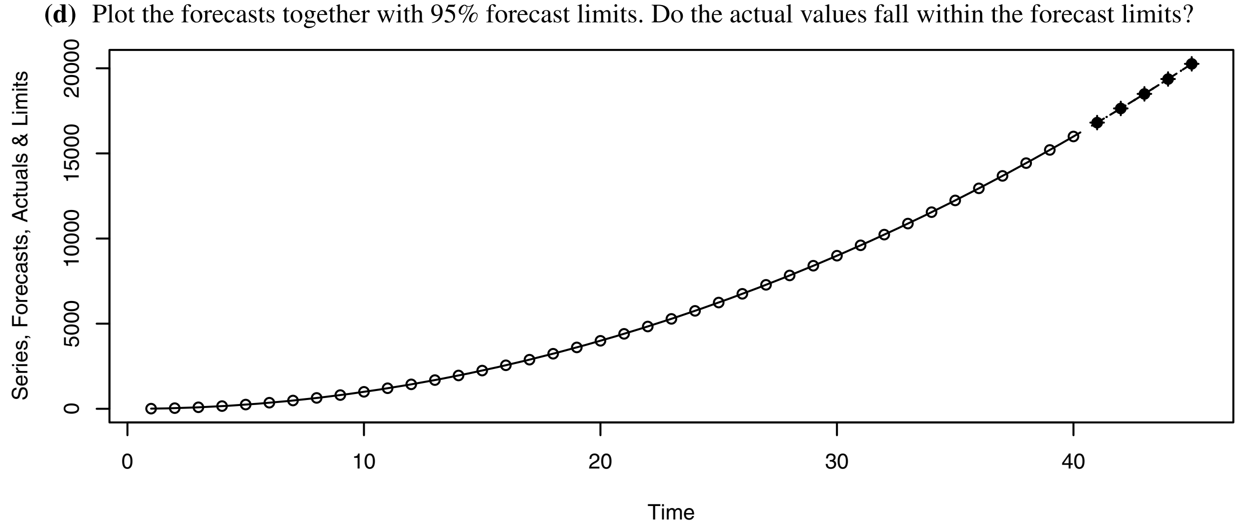
(c) Compare the five forecasts with the actual values that you set aside.

```
> forecast=result$pred; lower=result$lpi; upper=result$upi
> cbind(lower,forecast,actual,upper)
```

```
Time Series:
Start = 41
End = 45
Frequency = 1
      lower forecast   actual    upper
41 16804.96 16807.11 16808.72 16809.26
42 17636.78 17639.76 17640.59 17642.74
43 18487.95 18492.48 18495.70 18497.01
44 19358.70 19365.26 19368.23 19371.83
45 20249.16 20258.12 20259.99 20267.08
```

The comparison is best seen in the plot in part (d) below.

(d) Plot the forecasts together with 95% forecast limits. Do the actual values fall within the forecast limits?



```
> plot(model,n1=1,n.ahead=5,newxreg=(41:45)^2,type='o',pch=19,
   ylab='Series, Forecasts, Actuals & Limits')
> points(x=seq(41,45),y=actual,pch=3)
```

The forecast limits are so tight in this example that they cannot be seen on the plot. See part (c) for the actual numbers. The actual values are all within the forecast intervals for this simulation.

(e) Repeat parts (a) through (d) with a new simulated series using the same values of the parameters and same sample size.

Exercise 9.18 Consider the model $Y_t = \beta_0 + \beta_1 t + X_t$, where $X_t = \phi X_{t-1} + e_t$. We assume that β_0, β_1, and ϕ are known. Show that the minimum mean square error forecast ℓ steps ahead can be written as

$$\hat{Y}_t(\ell) = \beta_0 + \beta_1(t+\ell) + \phi^\ell(Y_t - \beta_0 - \beta_1 t).$$

Note: Since we assume all parameters are *known*, conditioning on $Y_1, Y_2, \ldots, Y_t$ is the same as conditioning on $X_1, X_2, \ldots, X_t$. So

$$\begin{aligned}\hat{Y}_t(\ell) &= E(\beta_0 + \beta_1(t+\ell) + X_{t+\ell} \mid Y_1, Y_2, \ldots, Y_t) \\ &= \beta_0 + \beta_1(t+\ell) + E(X_{t+\ell} \mid Y_1, Y_2, \ldots, Y_t) \\ &= \beta_0 + \beta_1(t+\ell) + E(X_{t+\ell} \mid X_1, X_2, \ldots, X_t)\end{aligned}$$

Now, since $\{X_t\}$ is an AR(1) process, $E(X_{t+\ell} \mid X_1, X_2, \ldots, X_t) = \phi^\ell X_t = \phi^\ell X_t = \phi^\ell(Y_t - \beta_0 - \beta_1 t)$ and the desired result is obtained.

Exercise 9.19 Verify Equation (9.3.12), page 196. $Y_t = e_t + \phi e_{t-1} + \phi^2 e_{t-2} + \cdots + \phi^{k-1} e_{t-k+1} + \phi^k Y_{t-k}$

$Y_t = e_t + \phi e_{t-1} + \phi^2 e_{t-2} + \phi^3 e_{t-3} + \cdots$

Exercise 9.20 Verify Equation (9.3.28), page 199.

From Equation (4.4.1), page 77, replacing t by $t + \ell$, we have

$$Y_{t+\ell} = \phi_1 Y_{t-1+\ell} + \phi_2 Y_{t-2+\ell} + \cdots + \phi_p Y_{t-p+\ell} + e_{t+\ell} - \theta_1 e_{t-1+\ell} - \theta_2 e_{t-2+\ell} - \cdots - \theta_q e_{t-q+\ell}. \text{ So}$$

$$\begin{aligned}\hat{Y}_t(\ell) = E(Y_{t+\ell} \mid Y_1, Y_2, \ldots, Y_t) &= \phi_1 E(Y_{t-1+\ell} \mid Y_1, Y_2, \ldots, Y_t) + \phi_2 E(Y_{t-2+\ell} \mid Y_1, Y_2, \ldots, Y_t) + \cdots \\ &+ \phi_p E(Y_{t-p+\ell} \mid Y_1, Y_2, \ldots, Y_t) + E(e_{t+\ell} \mid Y_1, Y_2, \ldots, Y_t) - \theta_1 E(e_{t-1+\ell} \mid Y_1, Y_2, \ldots, Y_t) - \cdots \\ &- \theta_q E(e_{t-q+\ell} \mid Y_1, Y_2, \ldots, Y_t)\end{aligned}$$

This becomes

$$\begin{aligned}\hat{Y}_t(\ell) &= \phi_1 \hat{Y}_t(\ell-1) + \phi_2 \hat{Y}_t(\ell-2) + \cdots + \phi_p \hat{Y}_t(\ell-p) + \theta_0 \\ &\quad - \theta_1 E(e_{t+\ell-1} | Y_1, Y_2, \ldots, Y_t) - \theta_2 E(e_{t+\ell-2} | Y_1, Y_2, \ldots, Y_t) \\ &\quad - \cdots - \theta_q E(e_{t+\ell-q} | Y_1, Y_2, \ldots, Y_t)\end{aligned}$$

where

$$E(e_{t+j} | Y_1, Y_2, \ldots, Y_t) = \begin{cases} 0 & \text{for } j > 0 \\ e_{t+j} & \text{for } j \le 0 \end{cases}$$

Exercise 9.21 The data file named deere3 contains 57 consecutive values from a complex machine tool process at Deere & Co. The values given are deviations from a target value in units of ten millionths of an inch. The process employs a control mechanism that resets some of the parameters of the machine tool depending on the magnitude of deviation from target of the last item produced.

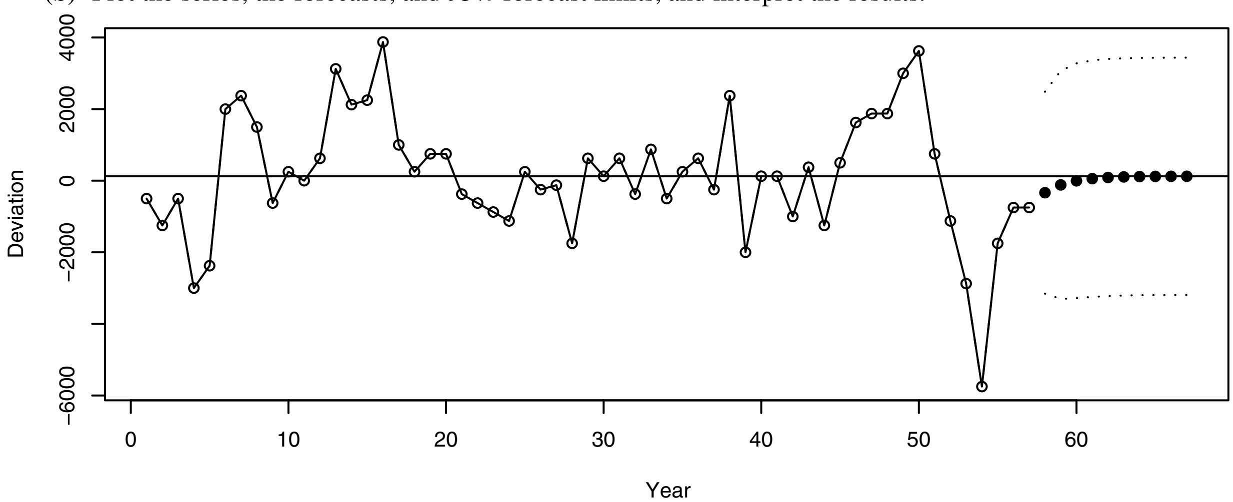
(a) Using an AR(1) model for this series, forecast the next ten values.

```
> data(deere3); model=arima(deere3,order=c(1,0,0)); plot(model,n.ahead=10)$pred
```

```
Time Series:
Start = 58
End = 67
Frequency = 1
 [1] -335.145915 -117.120755   -2.538371   57.680013   89.327581  105.959853
 [7]  114.700888  119.294709  121.708976  122.977786
```

The forecasts are reasonably constant from forecast lead 8 on.

(b) Plot the series, the forecasts, and 95% forecast limits, and interpret the results.




```
> win.graph(width=6.5,height=3,pointsize=8)
> plot(model,n.ahead=10,ylab='Deviation',xlab='Year',pch=19)
> abline(h=coef(model)[names(coef(model))=='intercept'])
```

Since the model does not contain a lot of autocorrelation or other structure, the forecasts, plotted as solid circles, quickly settle down to the mean of the series.

Exercise 9.22 The data file named days contains accounting data from the Winegard Co. of Burlington, Iowa. The data are the number of days until Winegard receives payment for 130 consecutive orders from a particular distributor of Winegard products. (The name of the distributor must remain anonymous for confidentiality reasons.) The time series contains outliers that are quite obvious in the time series plot. Replace each of the unusual values at "times" 63, 106, and 129 with the much more typical value of 35 days.

```
> data(days); daysmod=days; daysmod[63]=35; daysmod[106]=35; daysmod[129]=35
```

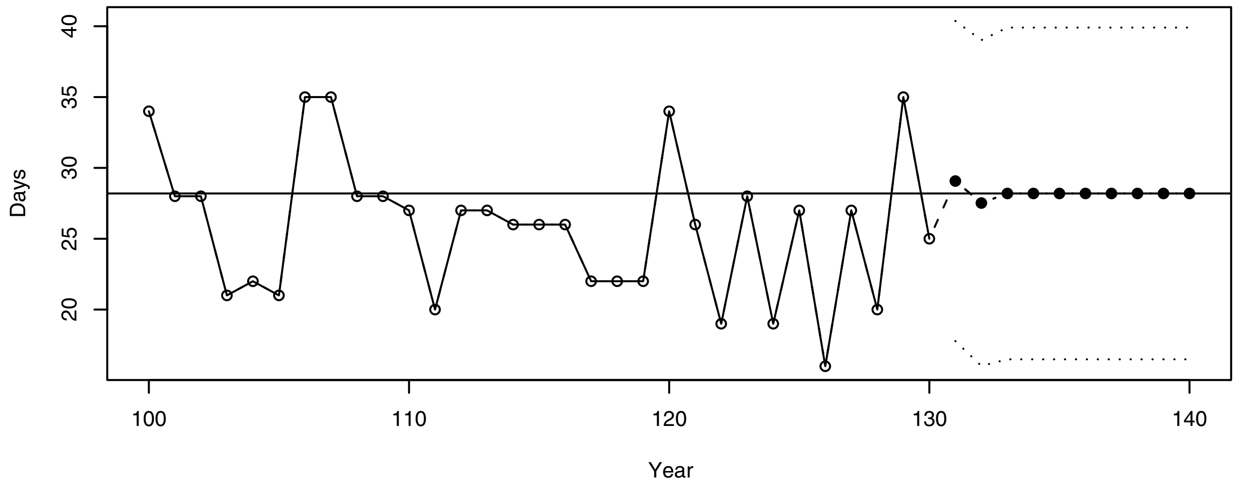
(a) Use an MA(2) model to forecast the next ten values of this modified series.

```
> model=arima(daysmod,order=c(0,0,2)); plot(model,n.ahead=10)$pred
```

```
Time Series:
Start = 131
End = 140
Frequency = 1
 [1] 29.07436 27.52056 28.19564 28.19564 28.19564 28.19564 28.19564 28.19564 28.19564
[10] 28.19564
```

Of course, these values would be rounded before reporting them. Notice that they are identical from lead 3 on.

(b) Plot the series, the forecasts, and 95% forecast limits, and interpret the results.




```
> plot(model,n1=100,n.ahead=10,ylab='Days',xlab='Year',pch=23)
> abline(h=coef(model)[names(coef(model))=='intercept'])
```

We start the series plot at $t = 100$ so that the detail in the forcasts can be seen more easily. The MA(2) model has autocorrelation only at lags 1 and 2 so the forecasts are just the process mean beyond lead 2.

Exercise 9.23 The time series in the data file robot gives the final position in the "*x*-direction" after an industrial robot has finished a planned set of exercises. The measurements are expressed as deviations from a target position. The robot is put through this planned set of exercises in the hope that its behavior is repeatable and thus predictable.

(a) Use an IMA(1,1) model to forecast five values ahead. Obtain 95% forecast limits also.

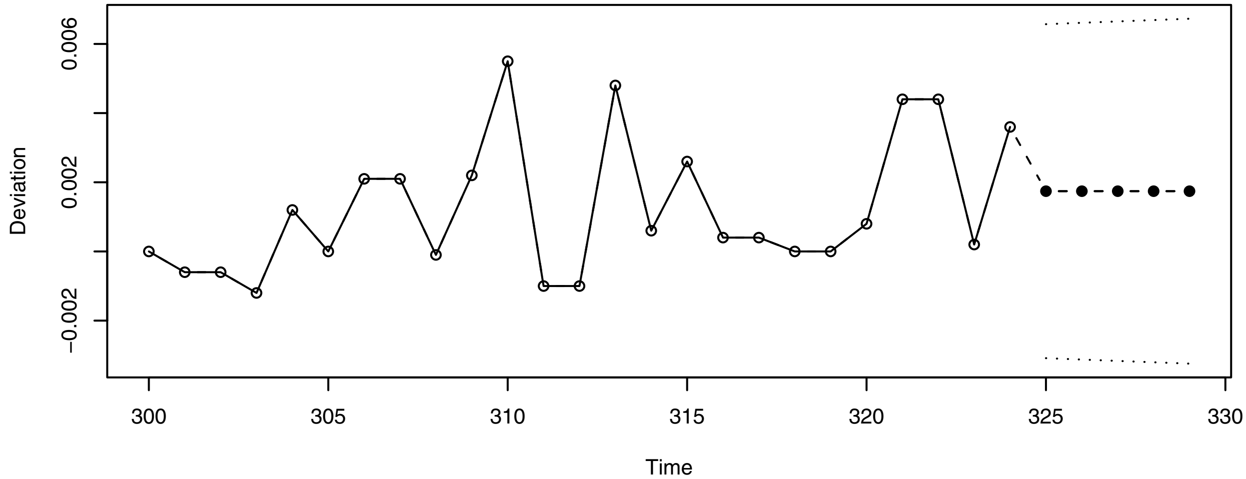
```
> data(robot); model=arima(robot,order=c(0,1,1)); model; plot(model,n.ahead=5)$pred
```

```
Time Series:
Start = 325
End = 329
Frequency = 1
[1] 0.001742672 0.001742672 0.001742672 0.001742672 0.001742672
```

```
> plot(model,n.ahead=5)$upi; plot(model,n.ahead=5)$lpi
```

```
Time Series:
Start = 325
End = 329
Frequency = 1
[1] 0.006669889 0.006710540 0.006750862 0.006790862 0.006830548
Time Series:
Start = 325
End = 329
Frequency = 1
[1] -0.003184545 -0.003225197 -0.003265519 -0.003305518 -0.003345204
```

(b) Display the forecasts, forecast limits, and actual values in a graph and interpret the results.




```
> win.graph(width=6.5,height=3,pointsize=8)
> plot(model,n1=300,n.ahead=5,ylab='Deviation',pch=19)
```

Thew forecast limits are quite wide in this fitted model and the forecasts are relatively constant.

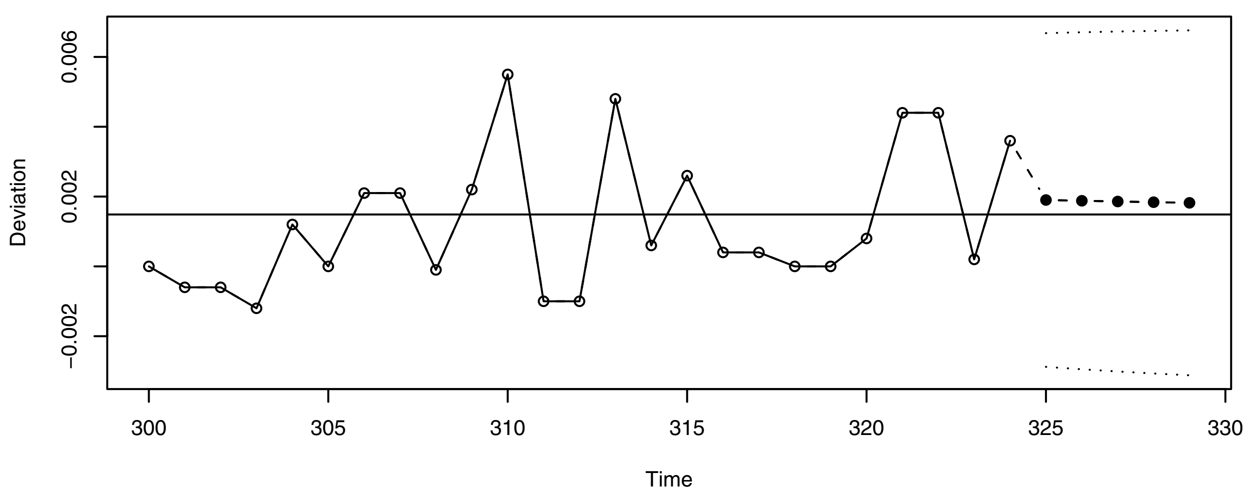
(c) Now use an ARMA(1,1) model to forecast five values ahead and obtain 95% forecast limits. Compare these results with those obtained in part (a).

```
> model=arima(robot,order=c(1,0,1)); plot(model,n.ahead=5)$pred
```

```
Time Series:
Start = 325
End = 329
Frequency = 1
[1] 0.001901348 0.001879444 0.001858695 0.001839041 0.001820424
```

```
> plot(model,n.ahead=5)$upi; plot(model,n.ahead=5)$lpi
```

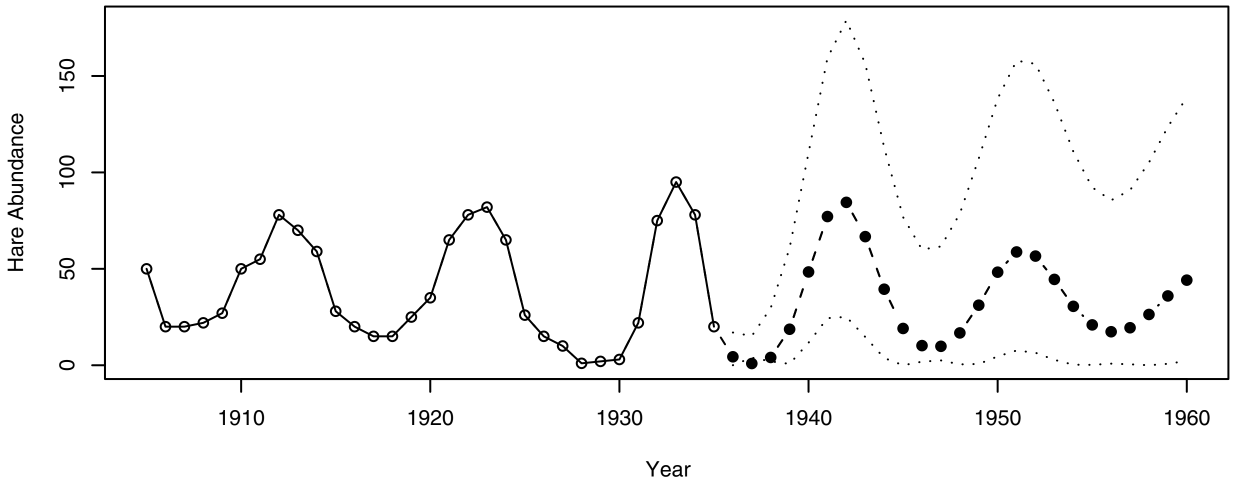
```
Time Series:
Start = 325
End = 329
Frequency = 1
[1] 0.006571344 0.006611183 0.006650699 0.006689898 0.006728790
Time Series:
Start = 325
End = 329
Frequency = 1
[1] -0.003086000 -0.003125839 -0.003165355 -0.003204555 -0.003243446
```



```
> plot(model,n1=300,n.ahead=5,ylab='Deviation',pch=19)
> abline(h=coef(model)[names(coef(model))=='intercept'])
```

Both of these models give quite similar forecasts and forecast limits.

Exercise 9.24 Exhibit (9.4), page 206, displayed the forecasts and 95% forecast limits for the square root of the Canadian hare abundance. The data are in the file named hare. Produce a similar plot in original terms. That is, plot the original abundance values together with the squares of the forecasts and squares of the forecast limits.




```
> data(hare); m1.hare=arima(sqrt(hare),order=c(3,0,0))
> square=function(x){y=x^2} # Define the square function
> plot(m1.hare,n.ahead=25,xlab='Year',ylab='Hare Abundance',pch=19,transform=square)
```

In original terms, the prediction intervals are not symmetric around the predictions.

Exercise 9.25 Consider the seasonal means plus linear time trend model for the logarithms of the monthly electricity generation time series in Exercise 9.8. (The data are in the file named electricity.)

(a) Find the two-year forecasts and forecast limits in original terms. That is, exponentiate (antilog) the results obtained in Exercise 9.8.

```
> data(electricity); win.graph(width=6.5,height=3,pointsize=8)
> month.=season(electricity); trend=time(electricity) # From Exercise 9.8
> model=arima(log(electricity),order=c(0,0,0),
   xreg=as.matrix(model.matrix(~month.+trend-1))[,-1])
> newmonth.=season(ts(rep(1,24),start=c(2006,1),freq=12))
> newtrend=time(electricity)[length(electricity)]+(1:24)*deltat(electricity)
> result=plot(model,n.ahead=24,n1=c(2001,1),ylab='Electricity',xlab='Year',
   newxreg=as.matrix(model.matrix(~newmonth.+ newtrend-1))[,-1],pch=19,transform=exp)
```

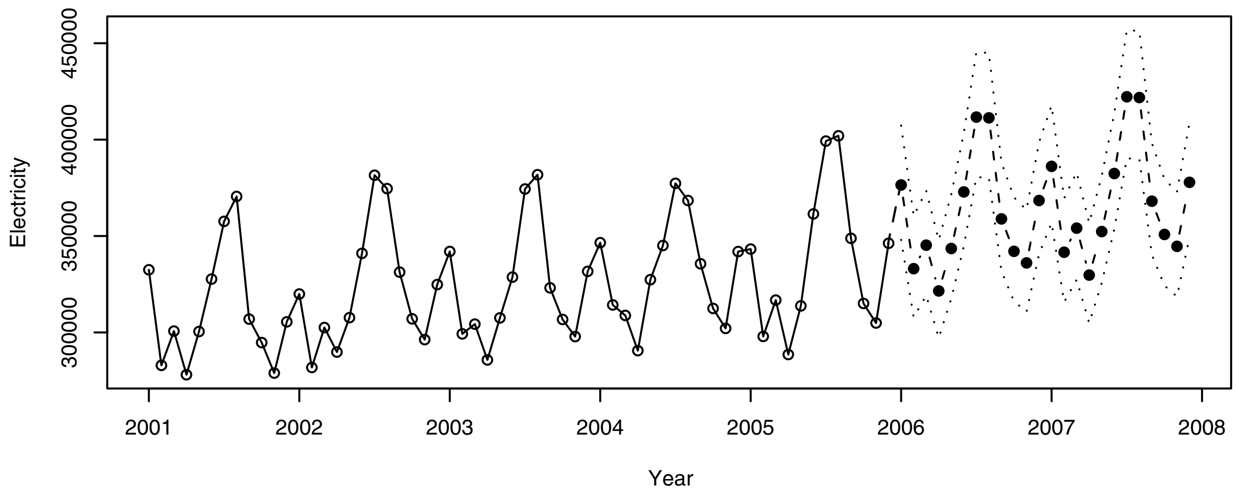
The statement `transform=exp` in the last line is the only thing new compared to the R code in Exercise 9.8. This last line will produce the plot needed for part (b) below. To see the forecasts and prediction limits use the following additional R code.

```
> forecast=result$pred; upper=result$upi; lower=result$lpi; cbind(lower,forecast,upper)
```

```
            lower forecast    upper
Jan 2006 348024.4 376498.6 407302.4
Feb 2006 307893.1 333083.9 360335.8
Mar 2006 319157.2 345269.6 373518.4
Apr 2006 297206.9 321523.4 347829.4
May 2006 317557.9 343539.4 371646.7
Jun 2006 344687.8 372889.0 403397.6
Jul 2006 380563.5 411699.9 445383.9
Aug 2006 380196.3 411302.7 444954.1
```

```
Sep 2006 331748.7 358891.2 388254.5
Oct 2006 316204.6 342075.4 370062.9
Nov 2006 310663.6 336081.1 363578.1
Dec 2006 340569.1 368433.3 398577.3
Jan 2007 356926.9 386129.5 417721.3
Feb 2007 315769.1 341604.3 369553.2
Mar 2007 327321.3 354101.7 383073.1
Apr 2007 304809.5 329748.0 356726.9
May 2007 325681.1 352327.2 381153.4
Jun 2007 353505.0 382427.6 413716.6
Jul 2007 390298.4 422231.3 456776.9
Aug 2007 389921.8 421823.9 456336.1
Sep 2007 340234.9 368071.8 398186.2
Oct 2007 324293.2 350825.8 379529.2
Nov 2007 318610.4 344678.1 372878.5
Dec 2007 349280.9 377857.9 408773.0
```

(b) Plot the last five years of the original time series together with two years of forecasts and the 95% forecast limits, all in original terms. Interpret the plot.



The forecasts follow the general upward trend and seasonal pattern of the series quite well and the prediction limits are reasonably tight from this model.

CHAPTER 10

Exercise 10.1 Based on quarterly data, a seasonal model of the form

$$Y_t = Y_{t-4} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$

has been fit to a certain time series.

(a) Find the first four ψ-weights for this model.

Iterating back in time, we have

$$Y_t = Y_{t-4} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} = (Y_{t-8} + e_{t-4} - \theta_1 e_{t-5} - \theta_2 e_{t-6}) + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$
$$= e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} + e_{t-4} - \theta_1 e_{t-5} - \theta_2 e_{t-6} + Y_{t-8}$$

From this we may read off $\psi_0 = 1$, $\psi_1 = -\theta_1$, $\psi_2 = -\theta_2$, $\psi_3 = 0$, and $\psi_4 = 1$.

(b) Suppose that $\theta_1 = 0.5$, $\theta_2 = -0.25$, and $\sigma_e = 1$. Find forecasts for the next four quarters if data for the last four quarters are

Quarter	I	II	III	IV
Series	25	20	25	40
Residual	2	1	2	3

$$\hat{Y}_t(1) = Y_{t-3} - \theta_1 e_t - \theta_2 e_{t-1} = 25 - (0.5)(3) - (-0.25)(2) = 24$$
$$\hat{Y}_t(2) = Y_{t-2} - \theta_2 e_t = 20 - (-0.25)(3) = 20.75$$
$$\hat{Y}_t(3) = Y_{t-1} = 25$$
$$\hat{Y}_t(4) = Y_t = 40$$

(c) Find 95% prediction intervals for the forecasts in part (b).

Quarter	Forecast Interval	
I	$24 \pm 2\sqrt{1} = 24 \pm 2$	or 20 to 24
II	$20.75 \pm 2\sqrt{1 + (0.5)^2} = 20.75 \pm 2.24$	or 18.51 to 22.99
III	$25 \pm 2\sqrt{1 + (0.5)^2 + (0.25)^2} = 25 \pm 2.29$	or 22.71 to 27.29
IV	$40 \pm 2\sqrt{1 + (0.5)^2 + (0.25)^2 + 0^2} = 40 \pm 2.29$	or 37.71 to 42.29

Exercise 10.2 An AR model has AR characteristic polynomial

$$(1 - 1.6x + 0.7x^2)(1 - 0.8x^{12})$$

(a) Is the model stationary?

Since $\Phi = 0.8$, the seasonal part of the model is staionary. In the nonseasonal part, $\phi_1 = 1.6$ and $\phi_2 = -0.7$. These parameter values satisfy Equations (4.3.11) on page 72. Therefore the complete model is stationary.

(b) Identify the model as a certain seasonal ARIMA model.

The model is a stationary, seasonal ARIMA$(2,0,0)\times(1,0,0)_{12}$ with $\phi_1 = 1.6$, $\phi_2 = -0.7$, and $\Phi = 0.8$. Expanded out it would be

$$\begin{aligned} Y_t &= 1.6Y_{t-1} - 0.7Y_{t-2} + 0.8Y_{t-12} - 1.6(0.8)Y_{t-13} + 0.7(0.8)Y_{t-14} + e_t \\ &= 1.6Y_{t-1} - 0.7Y_{t-2} + 0.8Y_{t-12} - 1.28Y_{t-13} + 0.56Y_{t-14} + e_t \end{aligned}$$

Exercise 10.3 Suppose that $\{Y_t\}$ satisfies

$$Y_t = a + bt + S_t + X_t$$

where S_t is deterministic and periodic with period s and $\{X_t\}$ is a seasonal ARIMA$(p,0,q)\times(P,1,Q)_s$ series. What is the model for $W_t = Y_t - Y_{t-s}$?

$$W_t = Y_t - Y_{t-s} = (a + bt + S_t + X_t) - [a + b(t-s) + S_{t-s} + X_{t-s}] = bs + S_t - S_{t-s} + \nabla_s X_t = bs + \nabla_s X_t$$

Therefore $\{W_t\}$ is ARMA$(p,q)\times(P,Q)_s$ with constant term bs.

Exercise 10.4 For the seasonal model $Y_t = \Phi Y_{t-4} + e_t - \theta e_{t-1}$ with $|\Phi| < 1$, find γ_0 and ρ_k.

$$\gamma_0 = Var(Y_t) = \Phi^2 Var(Y_{t-4}) + (1+\theta^2)\sigma_e^2 = \Phi^2\gamma_0 + (1+\theta^2)\sigma_e^2$$

So $\gamma_0 = \dfrac{1+\theta^2}{1+\Phi^2}\sigma_e^2$. Furthermore, since $E(e_t Y_{t-1}) = E[e_{t-1}(\Phi Y_{t-5} + e_{t-1} - \theta e_{t-2})] = \sigma_e^2$,

we have $\gamma_1 = E[\Phi Y_{t-4}Y_{t-1} + e_t Y_{t-1} - \theta e_{t-1} Y_{t-1}] = \Phi\gamma_3 - \theta\sigma_e^2$.

For $k > 1$, we have $\gamma_k = E[\Phi Y_{t-4} Y_{t-k} + e_t Y_{t-k} - \theta e_{t-1} Y_{t-k}] = \Phi\gamma_{k-4}$. Setting $k = 3$, we obtain

$\gamma_3 = \Phi\gamma_{-1} = \Phi\gamma_1$. This may be solved simultaneously with $\gamma_1 = \Phi\gamma_3 - \theta\sigma_e^2$ to yield $\gamma_1 = \frac{-\theta}{1-\Phi^2}\sigma_e^2$ which, in turn, gives $\rho_1 = \frac{-\theta}{1-\Phi^2}$.

For general k, we use the recursion $\gamma_k = \Phi\gamma_{k-4}$. For $k = 2$, we have $\rho_2 = \Phi\rho_{-2} = \Phi\rho_2$ and so $\rho_2 = 0$. We have further $\rho_3 = \Phi\rho_{-1} = \Phi\rho_1$, $\rho_4 = \Phi\rho_0 = \Phi$, and so on.

In summary, $\gamma_0 = \frac{1+\theta^2}{1+\Phi^2}\sigma_e^2$, $\rho_{4k+1} = \rho_{4k-1} = \frac{-\theta}{1-\Phi^2}\phi^k$, $\rho_{4k} = \Phi^k$, and $\rho_{4k+2} = 0$ for $k = 0, 1, 2, ...$

Exercise 10.5 Identify the following as certain multiplicative seasonal ARIMA models:

(a) $Y_t = 0.5Y_{t-1} + Y_{t-4} - 0.5Y_{t-5} + e_t - 0.3e_{t-1}$.

Rewriting as $Y_t - Y_{t-4} = 0.5(Y_{t-1} - Y_{t-5}) + e_t - 0.3e_{t-1}$ we see that it is an ARIMA$(1,0,1)\times(0,1,0)_4$ model with $\Phi_1 = 0.5$ and $\theta_1 = 0.3$.

Alternatively, we could factor the AR characteristic polynomial as $(1 - 0.5x - x^4 + 0.5x^5) = (1-0.5x)(1-x)^4$ and pick off the model and coefficients.

(b) $Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + e_t - 0.5e_{t-1} - 0.5e_{t-12} + 0.25e_{t-13}$. **[Typos in first printing of the book!]**

Rewriting $(Y_t - Y_{t-1}) - (Y_{t-12} - Y_{t-13}) = \nabla\nabla_{12}Y_t = e_t - 0.5e_{t-1} - 0.5e_{t-12} + (0.5)(0.5)e_{t-13}$ we see that the model is an ARIMA$(0,1,1)\times(0,1,1)_{12}$ with $\theta_1 = 0.5$, $\Theta_1 = 0.5$, and seasonal period 12.

Exercise 10.6 Verify Equations (10.2.11) on page 232.

The model is $Y_t = \Phi Y_{t-12} + e_t - \theta e_{t-1}$ with $0 < |\Phi| < 1$, so we have $Var(Y_t) = \Phi^2 Var(Y_{t-12}) + (1+\theta^2)\sigma_e^2$ or $\gamma_0 = \Phi^2\gamma_0 + (1+\theta^2)\sigma_e^2$ which gives $\gamma_0 = \left[\frac{1+\theta^2}{1-\Phi^2}\right]\sigma_e^2$.

Now $\gamma_1 = E(\Phi Y_{t-12}Y_{t-1} + e_tY_{t-1} - \theta e_{t-1}Y_{t-1}) = \Phi\gamma_{11} - \theta\sigma_e^2$ and, for $k > 1$, $\gamma_k = E(\Phi Y_{t-12}Y_{t-k} + e_tY_{t-k} - \theta e_{t-1}Y_{t-k}) = \Phi\gamma_{k-12}$. Putting $k = 11$, we have $\gamma_{11} = \Phi\gamma_1$ which we may solve simultaneously with $\gamma_1 = \Phi\gamma_{11} - \theta\sigma_e^2$ to obtain $\gamma_1 = \frac{-\theta}{1-\Phi^2}\sigma_e^2$ and then $\rho_1 = \frac{-\theta}{1-\theta^2}$. With $k = 2$ and $k = 10$, we get the pair $\gamma_2 = \Phi\gamma_{10}$ and $\gamma_{10} = \Phi\gamma_2$ which imply $\gamma_2 = \gamma_{10} = 0$. Similarly, the pair $k = 3$ and $k = 9$ yield $\gamma_3 = \gamma_9 = 0$ and eventually $\rho_2 = \rho_3 = \rho_4 = \cdots = \rho_{10} = 0$. However, $\rho_{11} = \Phi\rho_1$ and $\rho_{12} = \Phi\rho_0 = \Phi$. Furthermore, $\rho_{13} = \Phi\rho_1$ and $\rho_{13} = \Phi\rho_1 = \rho_{11}$. Similarly, $\rho_{14} = \rho_{15} = \cdots = \rho_{23} = 0$ but $\rho_{24} = \Phi\rho_{12} = \Phi^2$. In summary $\rho_{12k} = \Phi^k$ for $k \geq 1$, $\rho_{12k-1} = \rho_{12k+1} = -\frac{\theta}{1+\theta^2}\Phi^k$ for $k = 1, 2, 3, ...$, and all other autocorrelations are zero.

Exercise 10.7 Suppose that the process $\{Y_t\}$ develops according to $Y_t = Y_{t-4} + e_t$ with $Y_t = e_t$ for $t = 1, 2, 3$, and 4.

(a) Find the variance function for $\{Y_t\}$.

First note that $E\{Y_t\} = 0$. Now write $t = 4k + r$ where $r = 1, 2, 3$, or 4 and $k = 0, 1, 2, 3, ...$. (r is the quarter indicator and k is the year.) Now iterating back on t we have

$$\begin{aligned} Y_t &= Y_{t-4} + e_t = (Y_{t-8} + e_{t-4}) + e_t = (Y_{t-12} + e_{t-8}) + e_{t-4} + e_t \\ &\vdots \\ &= e_t + e_{t-4} + e_{t-8} + \cdots + e_{r+4} + e_r \end{aligned}$$

Since there are exactly $k + 1$ e's in the sum we have $Var(Y_t) = (k+1)\sigma_e^2$.

(b) Find the autocorrelation function for $\{Y_t\}$.

Let $s > t$ and express $s = 4j + i$ where $i = 1, 2, 3$, or 4 and $j = 0, 1, 2, 3, ...$. Then

$Cov(Y_t, Y_s) = Cov(e_t + e_{t-4} + \cdots + e_r, e_s + e_{s-4} + \cdots + e_i)$. If $r \neq i$, then there will be no overlap among te two sets of e's and the covariance will be zero. If $r = i$, then that is, s and t correspond to the same quarter of the year, then the e's will be the same from t down to r and $Cov(Y_t, Y_s) = Var(Y_t)$. In summary, for $s > t$

$$Cov(Y_t, Y_s) = \begin{cases} (k+1)\sigma_e^2 & \text{if } r = i \\ 0 & \text{otherwise} \end{cases}$$ Note: $r = i$ is equivalent to $s - t$ being divisible by 4. Finally,

$$Corr(Y_t, Y_s) = \begin{cases} \sqrt{\dfrac{k+1}{j+1}} & \text{if } s = 4j + r, t = 4k + r \quad (r = 1, 2, 3, 4) \\ 0 & \text{otherwise} \end{cases}$$

For large t, observations for the same quarter in neighboring years will be strongly positively correlated. Observations in different quarters will always be uncorrelated.

(c) Identify the model for $\{Y_t\}$ as a certain seasonal ARIMA model.

Since $Y_t - Y_{t-4} = e_t$, this is a seasonal ARIMA(0,0,0)×$(0,1,0)_4$. That is, the quarterly seasonal difference is white noise.

Exercise 10.8 Consider the Alert, Canada, monthly carbon dioxide time series shown in Exhibit (10.1), page 227. The data are in the file named co2.

(a) Fit a deterministic seasonal means plus linear time trend model to these data. Are any of the regression coefficients “statistically significant”?

```
> data(co2); month.=season(co2); trend=time(co2)
> model=lm(co2~month.+trend); summary(model)
```

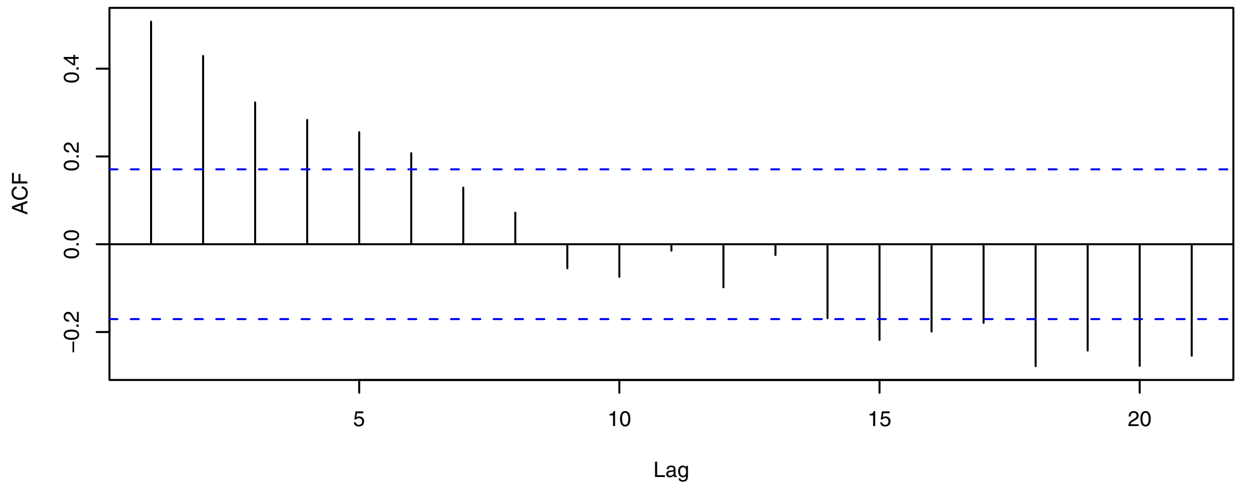
```
Call:
lm(formula = co2 ~ month. + trend)

Residuals:
     Min       1Q   Median       3Q      Max
-1.73874 -0.59689 -0.06947  0.54086  2.15539

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)     -3290.5412    44.1790 -74.482  < 2e-16 ***
month.February      0.6682     0.3424   1.952 0.053320 .
month.March         0.9637     0.3424   2.815 0.005715 **
month.April         1.2311     0.3424   3.595 0.000473 ***
month.May           1.5275     0.3424   4.460 1.87e-05 ***
month.June         -0.6761     0.3425  -1.974 0.050696 .
month.July         -7.2851     0.3426 -21.267  < 2e-16 ***
month.August      -13.4414     0.3426 -39.232  < 2e-16 ***
month.September   -12.8205     0.3427 -37.411  < 2e-16 ***
month.October      -8.2604     0.3428 -24.099  < 2e-16 ***
month.November     -3.9277     0.3429 -11.455  < 2e-16 ***
month.December     -1.3367     0.3430  -3.897 0.000161 ***
trend               1.8321     0.0221  82.899  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.8029 on 119 degrees of freedom
Multiple R-squared: 0.9902,     Adjusted R-squared: 0.9892
F-statistic: 997.7 on 12 and 119 DF,  p-value: < 2.2e-16
```

All of the regression coefficients are statistically significant except for the seasonal effects for February and June. Those two have p-values just above 0.05.

(b) What is the multiple R-squared for this model?

The multiple R-squared can be read off the output as 99.02%.

(c) Now calculate the sample autocorrelation of the residuals from this model. Interpret the results.

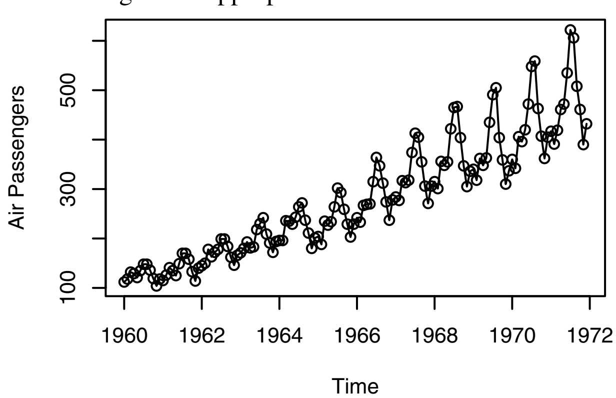


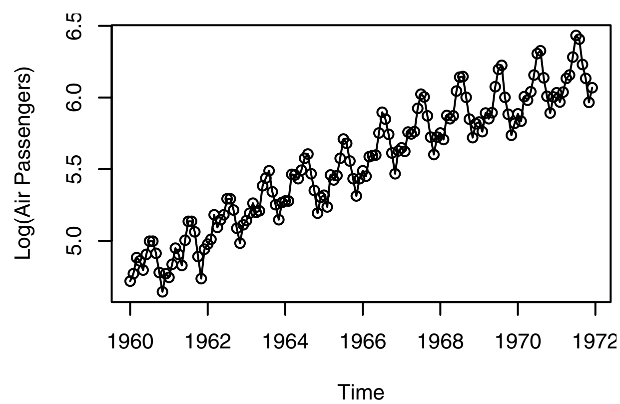

```
> acf(residuals(model))
```

Clearly, this deterministic trend model has not captured the autocorrelation in this time series. The seasonal ARIMA model illustrated in Chapter 10 is a much better model for these data.

Exercise 10.9 The monthly airline passenger time series, first investigated in Box and Jenkins (1976), is considered a classic time series. The data are in the file named airline. **[Typo: The filename is airpass.]**

(a) Display the time series plots of both the original series and the logarithms of the series. Argue that taking logs is an appropriate transformation.

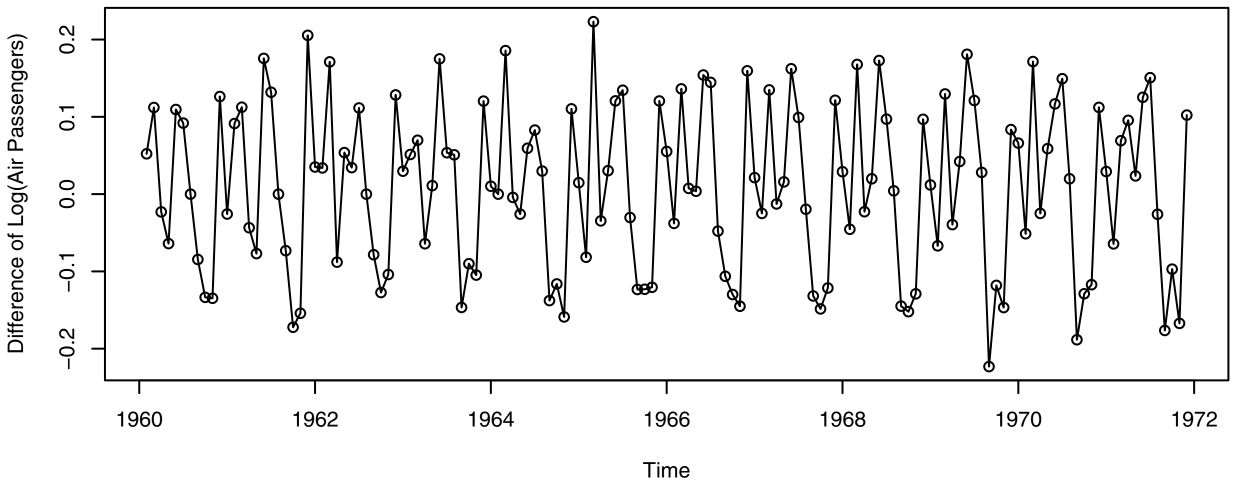





```
> win.graph(width=3.25,height=2.5,pointsize=8)
> data(airpass); plot(airpass, type='o',ylab='Air Passengers')
> plot(log(airpass), type='o',ylab='Log(Air Passengers)')
```

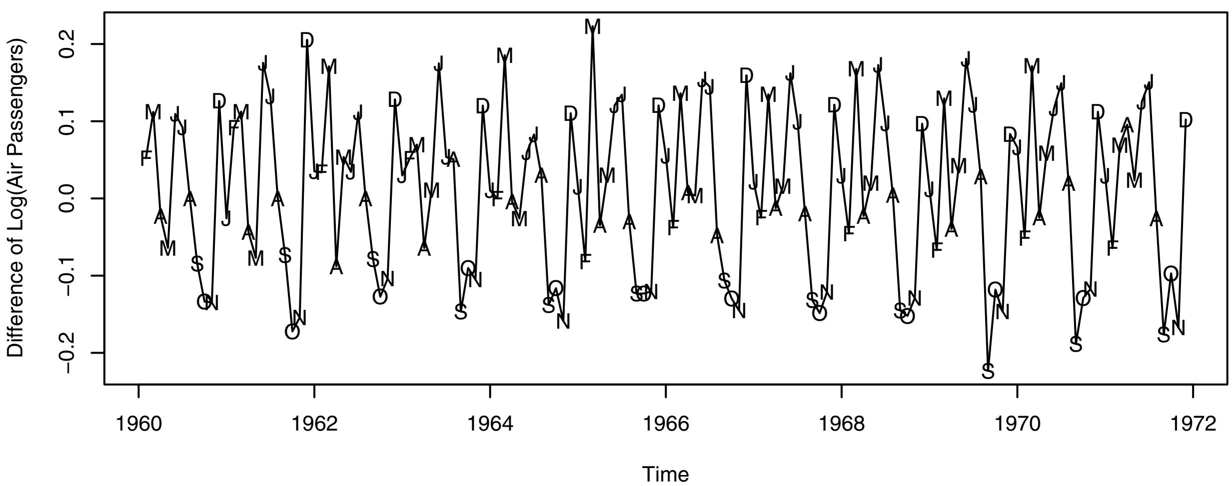
The graph of the logarithms displays a much more constant variation around the upward “trend.”

(b) Display and interpret the time series plots of the first difference of the logged series.




```
> win.graph(width=6.5,height=3,pointsize=8)
> plot(diff(log(airpass)),type='o',ylab='Difference of Log(Air Passengers)')
```

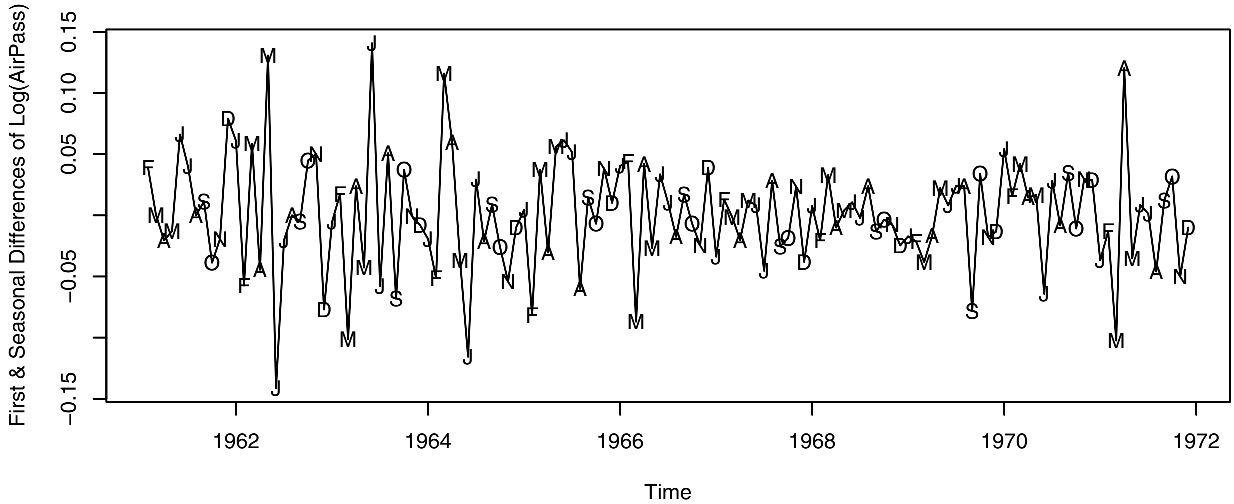
This series appears to be stationary. However, seasonality, if any, could be hiding in this plot. Perhaps a plot using seasonal plotting symbols would reveal the seasonality.




```
> plot(diff(log(airpass)),type='l',ylab='Difference of Log(Air Passengers)')
> points(diff(log(airpass)),x=time(diff(log(airpass))),
   pch=as.vector(season(diff(log(airpass))))
```

The seasonality can be observed by looking at the plotting symbols carefully. Septembers, Octobers, and Novembers are mostly at the low points and Decembers mostly at the high points.

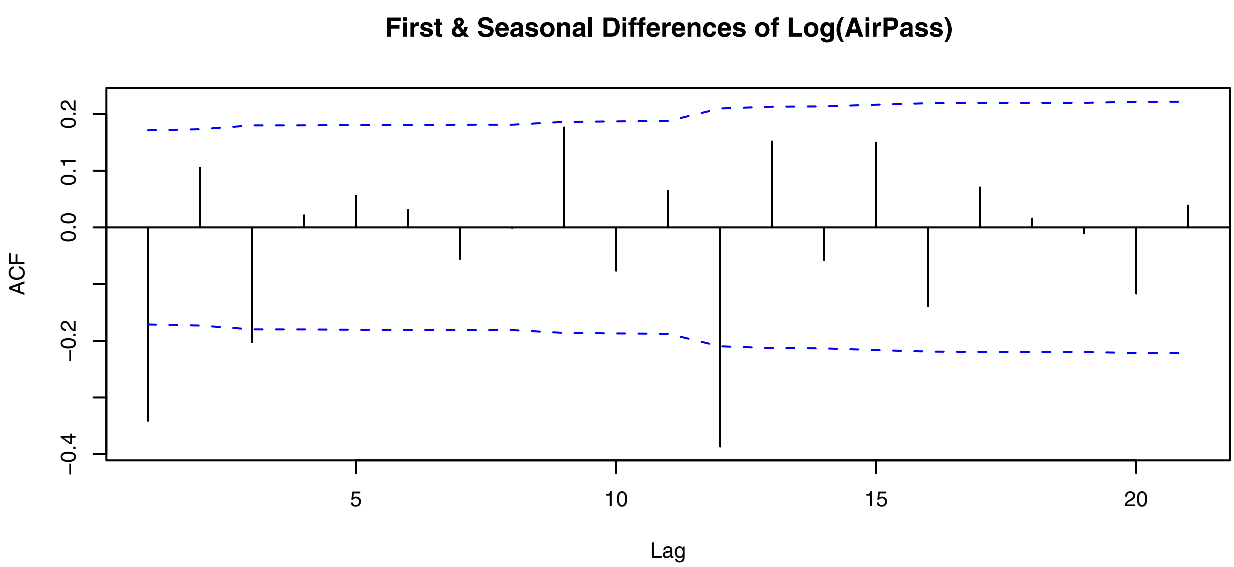
(c) Display and interpret the time series plot of the seasonal difference of the first difference of the logged series.




```
> plot(diff(diff(log(airpass)),lag=12),type='l',
   ylab='First & Seasonal Differences of Log(AirPass)')
> points(diff(diff(log(airpass)),lag=12),x=time(diff(diff(log(airpass)),lag=12)),
   pch=as.vector(season(diff(diff(log(airpass)),lag=12))))
```

We chose to do the plot with seasonal plotting symbols. The seasonality is much less obvious now. Some Decembers are high and some low. Similarly, some Octobers are high and some low.

(d) Calculate and interpret the sample ACF of the seasonal difference of the first difference of the logged series.




```
> acf(as.vector(diff(diff(log(airpass)),lag=12)),ci.type='ma',
   main='First & Seasonal Differences of Log(AirPass)')
```

Although there is a “significant” autocorrelation at lag 3, the most prominent autocorrelations are at lags 1 and 12 and the “airline” model seems like a reasonable choice to investigate.

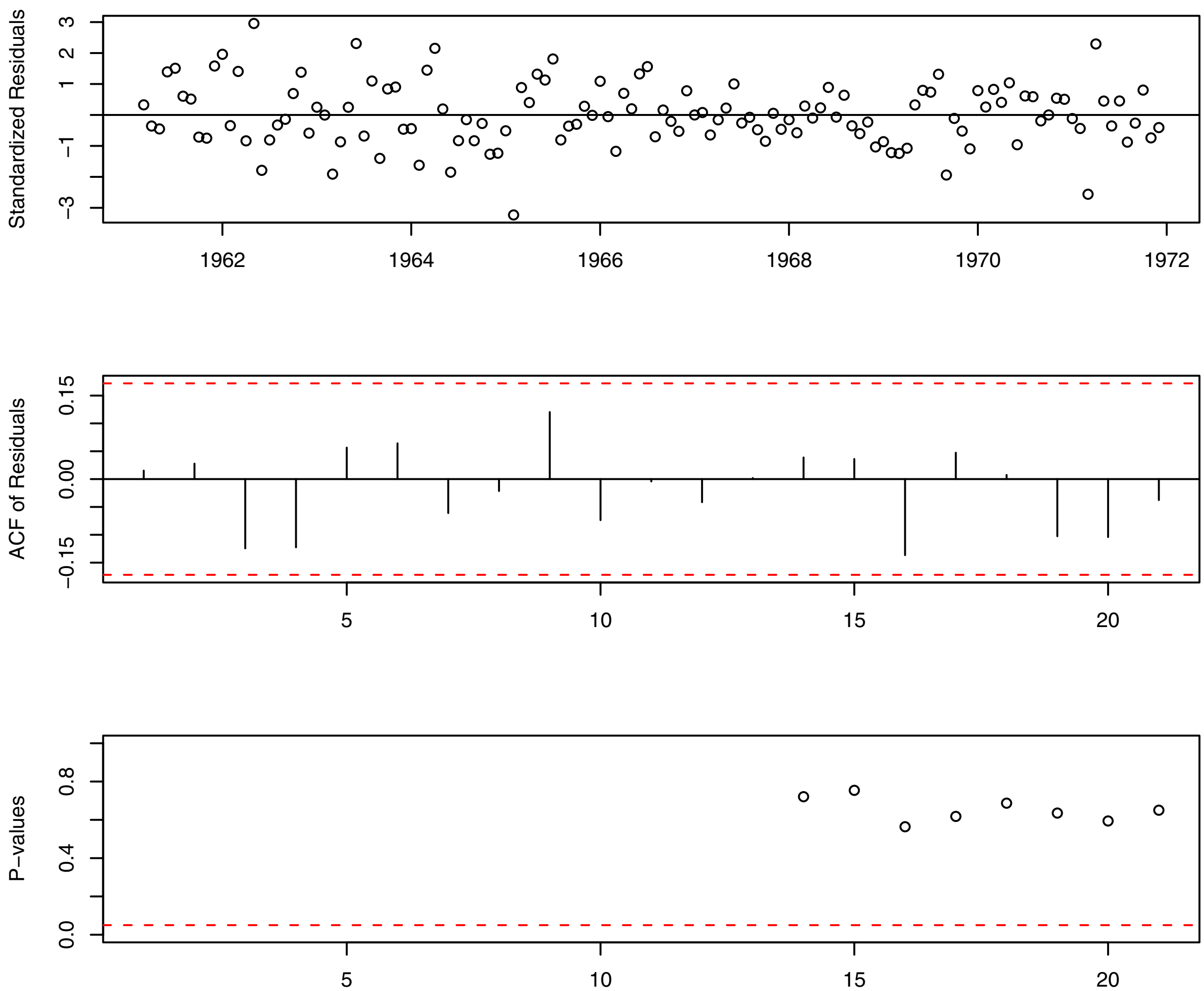
(e) Fit the “airline model” (ARIMA(0,1,1)×(0,1,1)$_{12}$) to the logged series.

```
> model=arima(log(airpass),order=c(0,1,1),seasonal=list(order=c(0,1,1),period=12))
> model
```

```
Call:
arima(x = log(airpass), order = c(0, 1, 1), seasonal = list(order = c(0, 1,
    1), period = 12))

Coefficients:
          ma1     sma1
      -0.4018  -0.5569
s.e.   0.0896   0.0731

sigma^2 estimated as 0.001348:  log likelihood = 244.7,  aic = -485.4
```

Notice that both the seasonal and nonsrasonal ma parameters are significant.

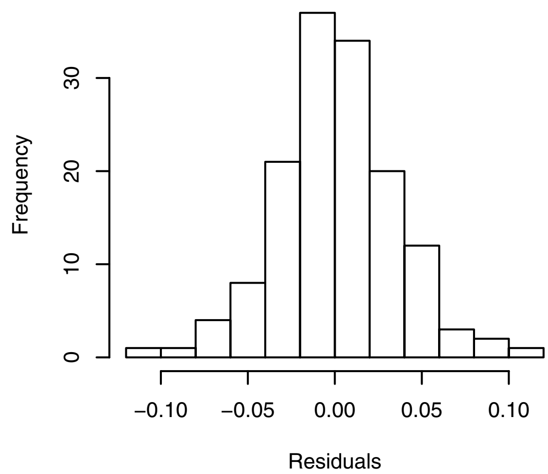
(f) Investigate diagnostics for this model, including autocorrelation and normality of the residuals.

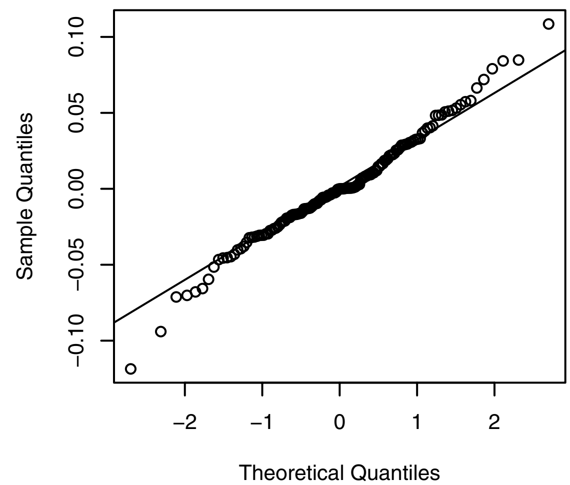



```
> win.graph(width=6.5,height=6); tsdiag(model)
```

None of these three plots indicate difficulties with the model. There are no outliers and little autocorrelation in the residuals, both indiviually and jointly.

Let's look at normality.






```
> win.graph(width=4,height=3,pointsize=8); hist(residuals(model),xlab='Residuals')
> qqplot(residuals(model)); qqline(residuals(model))
```

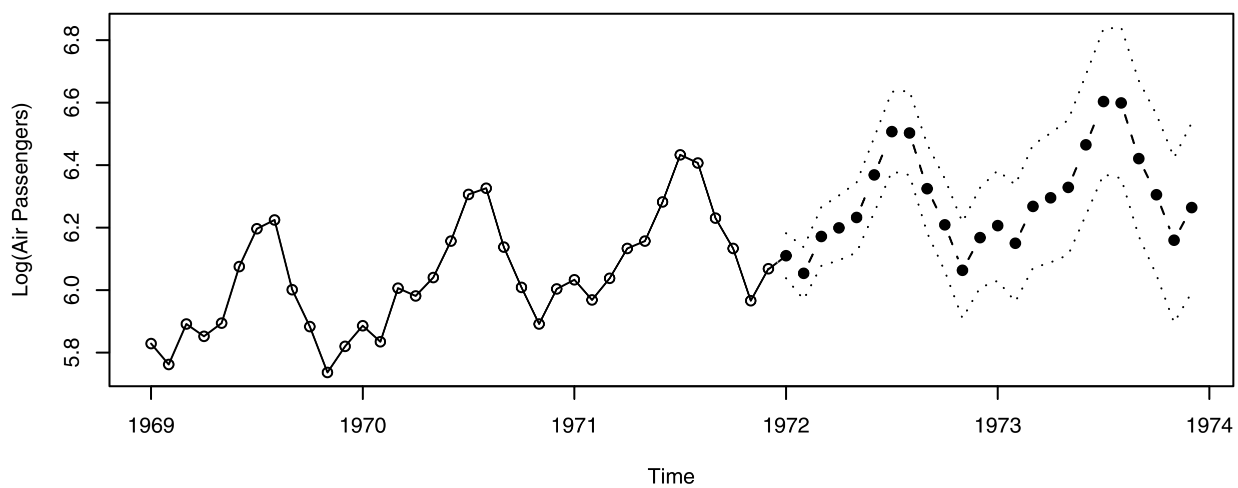
The distribution of the residuals is quite symmetric but the Q-Q plot indicates that the tails are lighter than a normal distribution. Let's do a Shapiro-Wilk test.

```
> shapiro.test(residuals(model))
```

```
        Shapiro-Wilk normality test

data:  residuals(model)
W = 0.9864, p-value = 0.1674
```

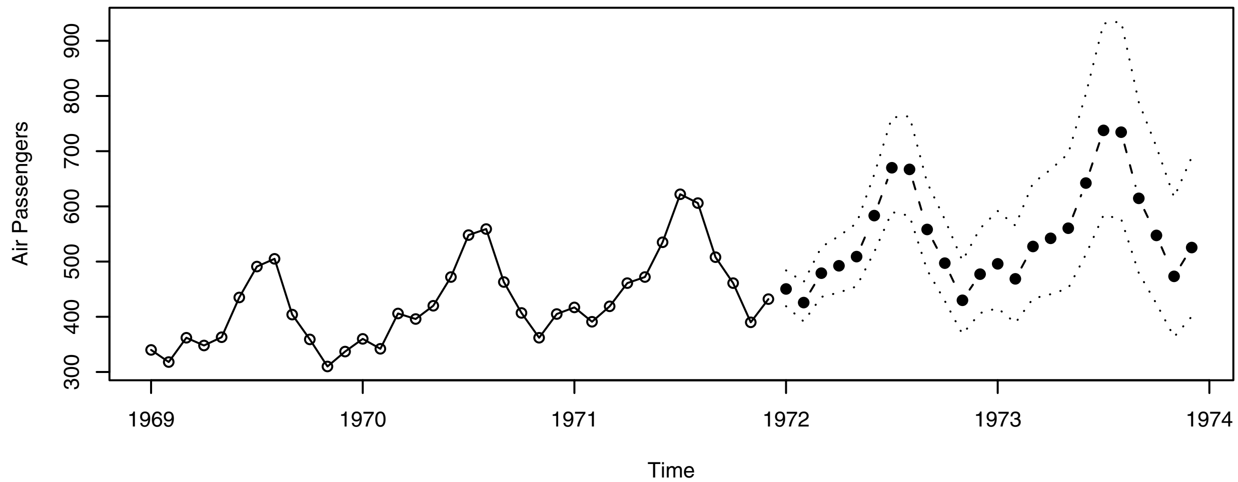
The Shapiro-Wilk test does not reject normality of the error terms at any of the usual significance levels and we procede to use the model for forecasting.

(g) Produce forecasts for this series with a lead time of two years. Be sure to include forecast limits.




```
> win.graph(width=6.5,height=3,pointsize=8)
> plot(model,n1=c(1969,1),n.ahead=24,pch=19,ylab='Log(Air Passengers)')
```

The forecasts follow the seasonal and upward “trend” of the time series nicely. The forecast limits provide us with a clear measure ofthe uncertainty in the forecasts. For completeness, we also plot the forecasts and limits in orginal terms.

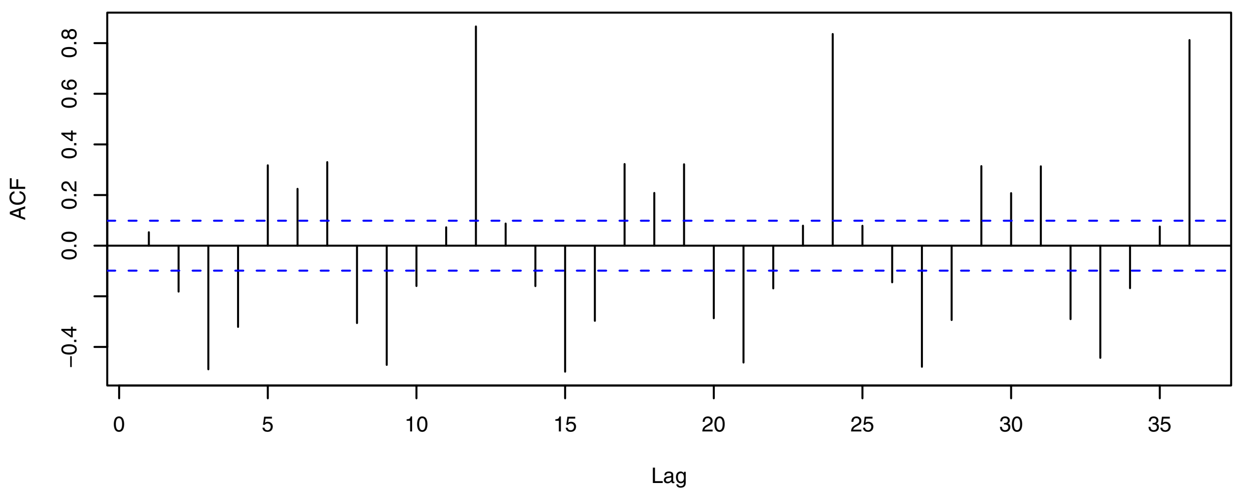



```
> plot(model,n1=c(1969,1),n.ahead=24,pch=19,ylab='Air Passengers',transform=exp)
```

In original terms it is easier to see that the forecast limits spread out as we get further into the future.

Exercise 10.10 Exhibit (5.8), page 99 displayed the monthly electricity generated in the United States. We argued there that taking logarithms was appropriate for modeling. Exhibit (5.10), page 100 showed the time series plot of the first differences for this series. The filename is electricity.

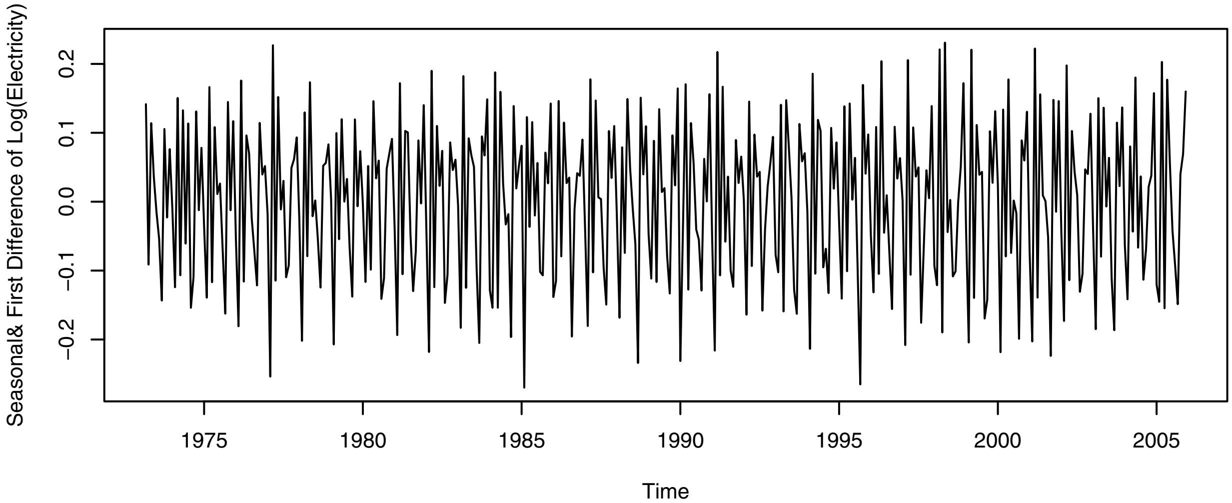
(a) Calculate the sample ACF of the first difference of the logged series. Is the seasonality visible in this display?




```
> data(electricity); acf(diff(log(as.vector(electricity))),lag.max=36)
```

The very strong autocorrelations at lags 12, 24, and 36 point out the substantial seasonality in this time series.

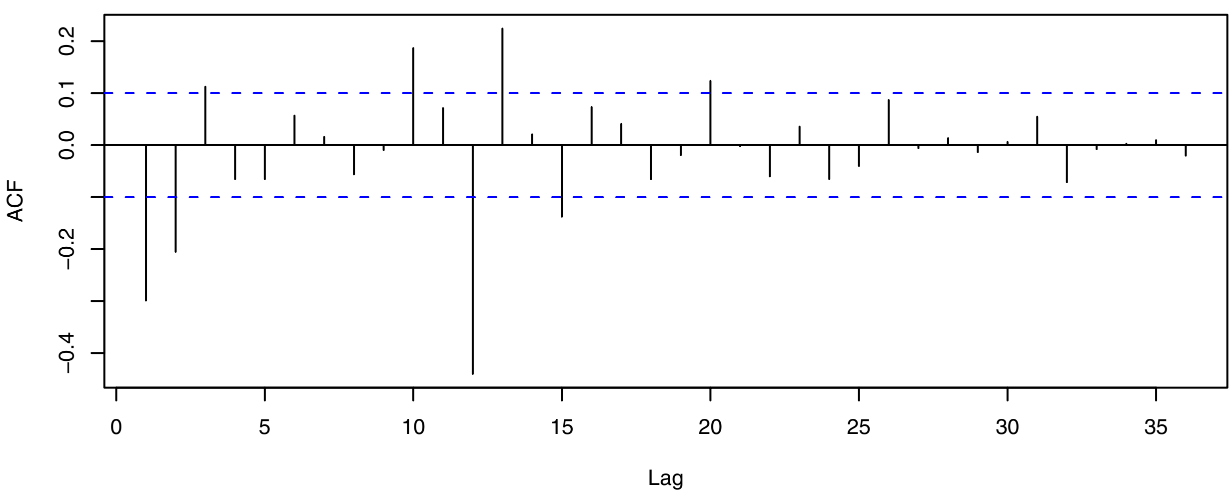
(b) Plot the time series of seasonal difference and first difference of the logged series. Does a stationary model seem appropriate now?




```
> plot(diff(diff(log(electricity))),ylab='Seasonal & First Difference of
   Log(Electricity)')
```

The time series plot appears stationary but the seasonality will still have to be investigated further and modeled.

(c) Display the sample ACF of the series after a seasonal difference and a first difference have been taken of the logged series. What model(s) might you consider for the electricity series?

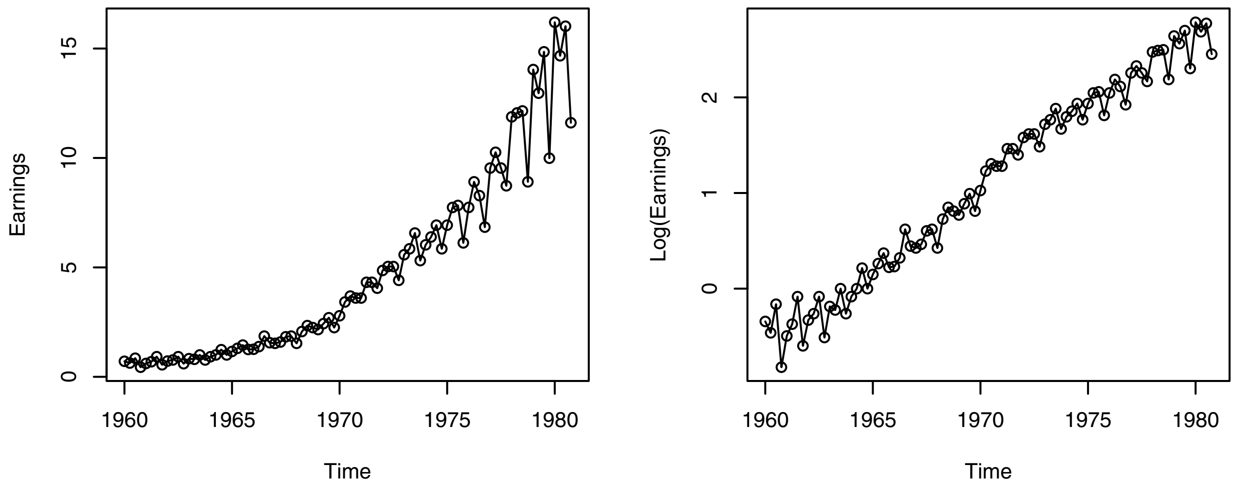



```
> acf(diff(diff(log(as.vector(electricity)),lag=12)),lag.max=36)
```

After seasonal differencing, the strong autocorrelations at lags 24 and 36 have disappeared. Perhaps a stationay model could now be entertained. The "airline" model, ARIMA(0,1,1)×$(0,1,1)_{12}$ for the logarithms, might capture most of the autocorrelation structure in this series.

Exercise 10.11 The quarterly earnings per share for 1960–1980 of the U.S. company Johnson & Johnson, are saved in the file named JJ.

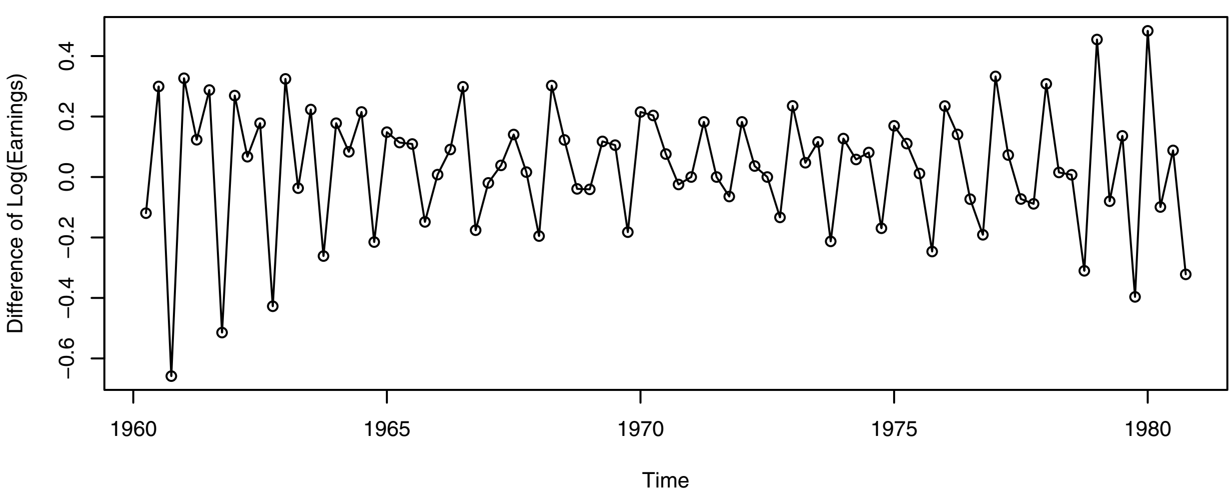
(a) Plot the time series and also the logarithm of the series. Argue that we should transform by logs to model this series.




```
> data(JJ); win.graph(width=6.5,height=3,pointsize=8); oldpar=par; par(mfrow=c(1,2))
> plot(JJ,ylab='Earnings',type='o'); plot(log(JJ),ylab='Log(Earnings)',type='o')
> par=oldpar
```

In the plot at the left it is clear that at the higher values of the series there is also much more variation. The plot opf the logs at the right shows much more equal variation at all levels of the series. We use logs for all of the remaining modeling.

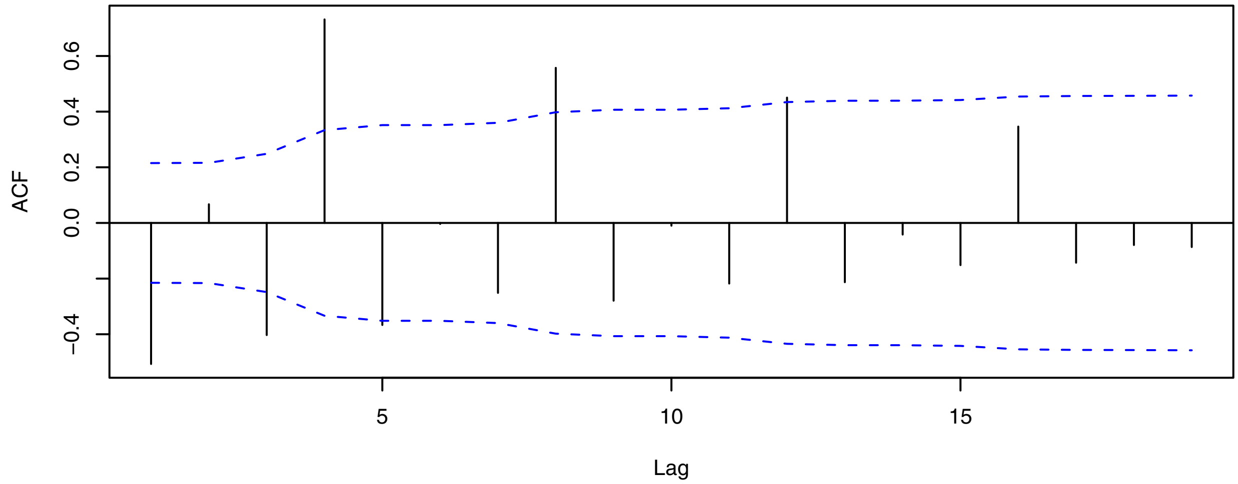
(b) The series is clearly not stationary. Take first differences and plot that series. Does stationarity now seem reasonable?




```
> plot(diff(log(JJ)),ylab='Difference of Log(Earnings)',type='o')
```

We do not expect stationay series to have less variability in the middle of the series as this one does but we might entertain a stationary model and see where it leads us.

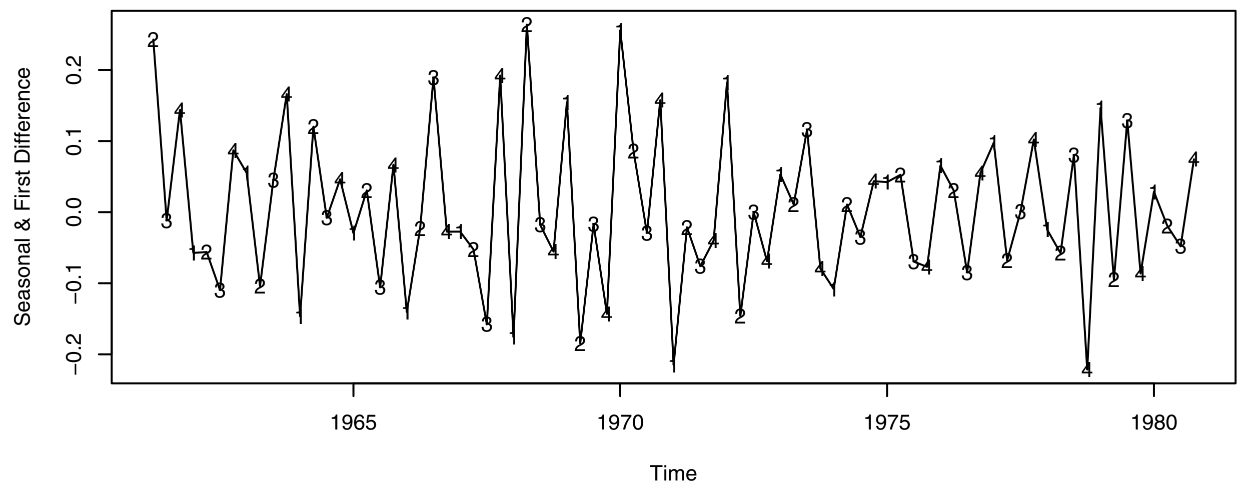
(c) Calculate and graph the sample ACF of the first differences. Interpret the results.




```
> acf(diff(log(as.vector(JJ))),ci.type='ma')
```

In this quarterly series, the strongest autocorrelations are at the seasonal lags of 4, 8, 12, and 16. Clearly, we need to address the seasonality in this series.

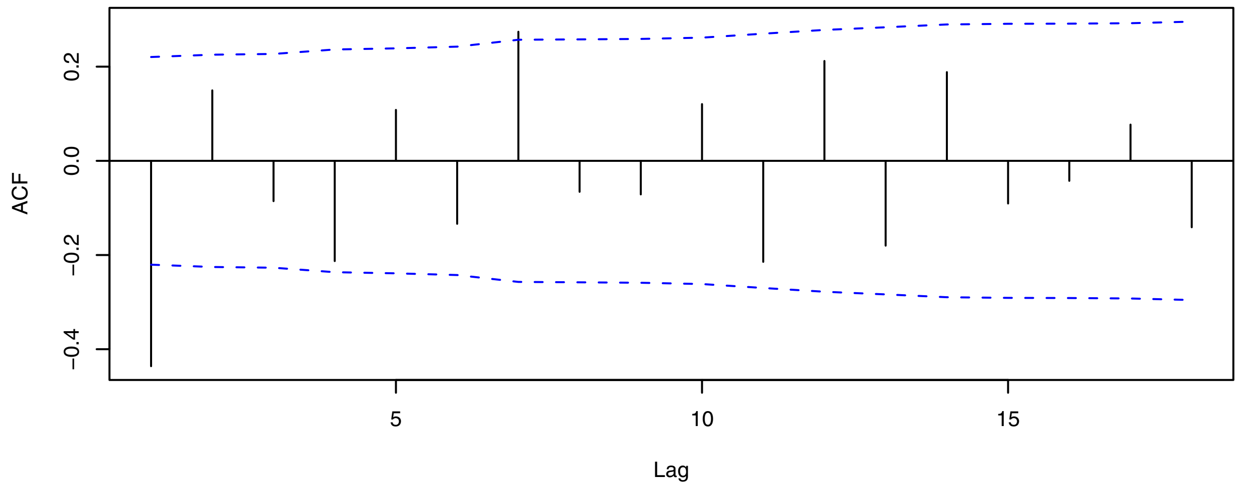
(d) Display the plot of seasonal differences and the first differences. Interpret the plot. Recall that for quarterly data, a season is of length 4.




```
> series=diff(diff(log(JJ),lag=4))
> plot(series,ylab='Seasonal & First Difference',type='l')
> points(y=series,x=time(series),pch=as.vector(season(series)))
```

The various quarters seem to be quite randomly distributede among high, middle, and low values, so that most of the seasonality is accounted for in the seasonal difference.

(e) Graph and interpret the sample ACF of seasonal differences with the first differences.




```
> acf(as.vector(series),ci.type='ma')
```

They only significant autocorrelations are at lags 1 and 7. Lag 4 (the quarterly lag) is nearly significant.

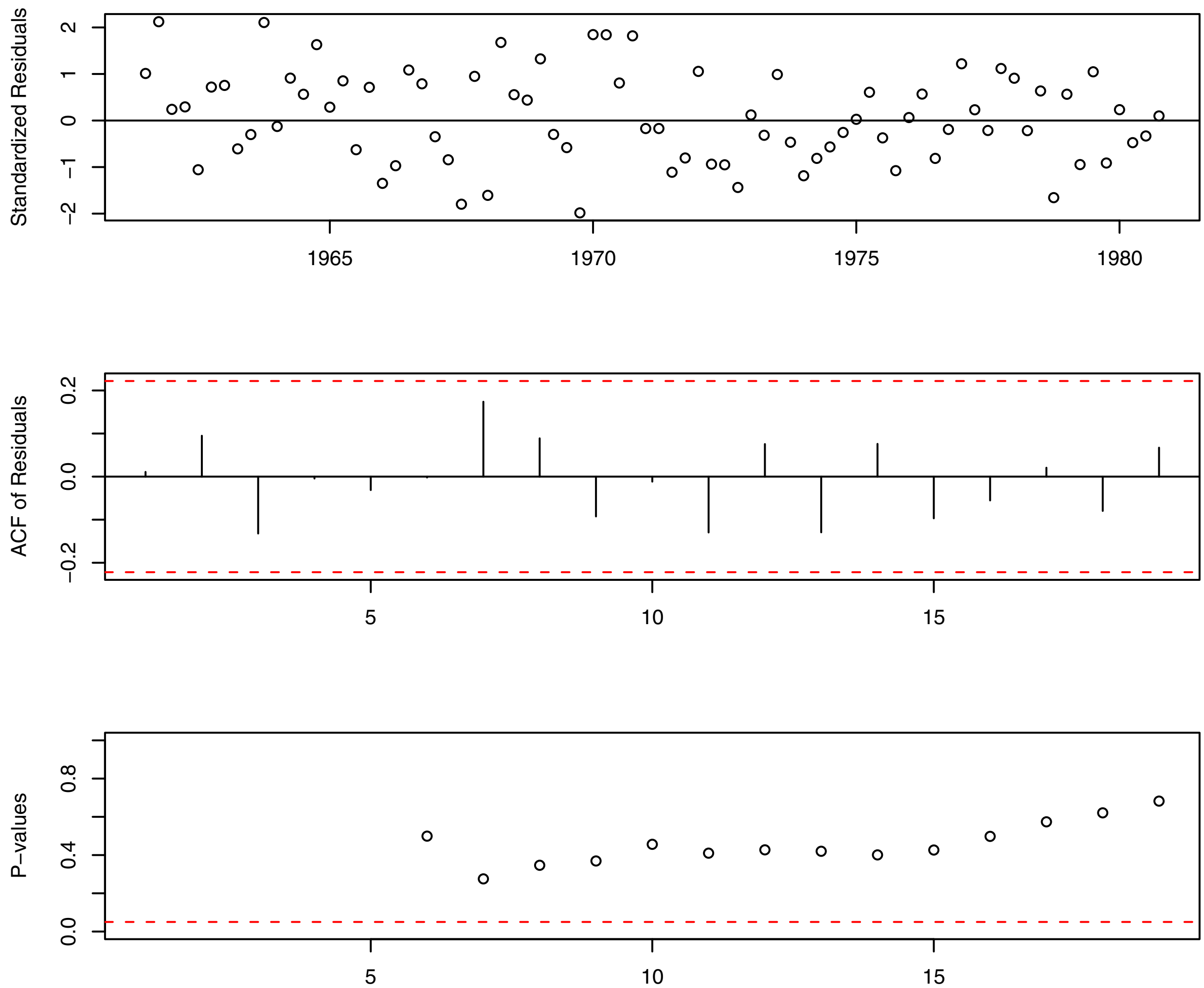
(f) Fit the model ARIMA(0,1,1)×(0,1,1)$_4$, and assess the significance of the estimated coefficients.

```
> model=arima(log(JJ),order=c(0,1,1),seasonal=list(order=c(0,1,1),period=4)); model
```

```
Call:
arima(x = log(JJ), order = c(0, 1, 1), seasonal = list(order = c(0, 1, 1), period
= 4))

Coefficients:
          ma1     sma1
      -0.6809  -0.3146
s.e.   0.0982   0.1070

sigma^2 estimated as 0.00793:  log likelihood = 78.38,  aic = -152.75
```

Both the seasonal and nonseasonal ma parameters are significant in this model.

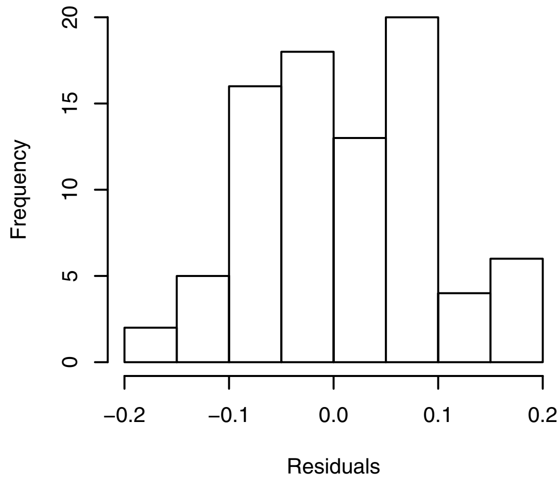
(g) Perform all of the diagnostic tests on the residuals.

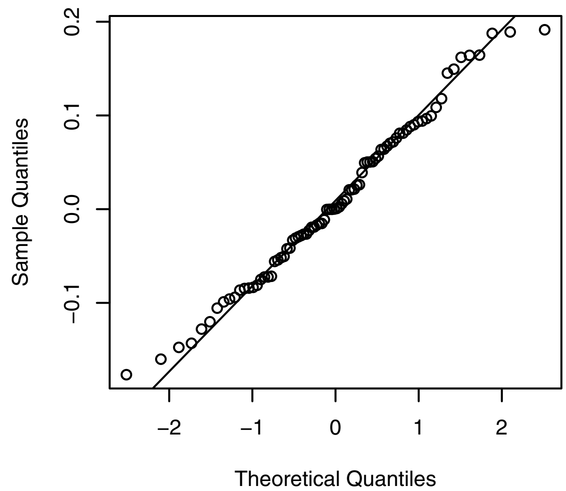



```
> win.graph(width=6.5,height=6); tsdiag(model)
```

These diagnostic plots do not show any inadequacies with the model. No outliers are detected and there is little autocorrelation in the residuals.

On to normality.



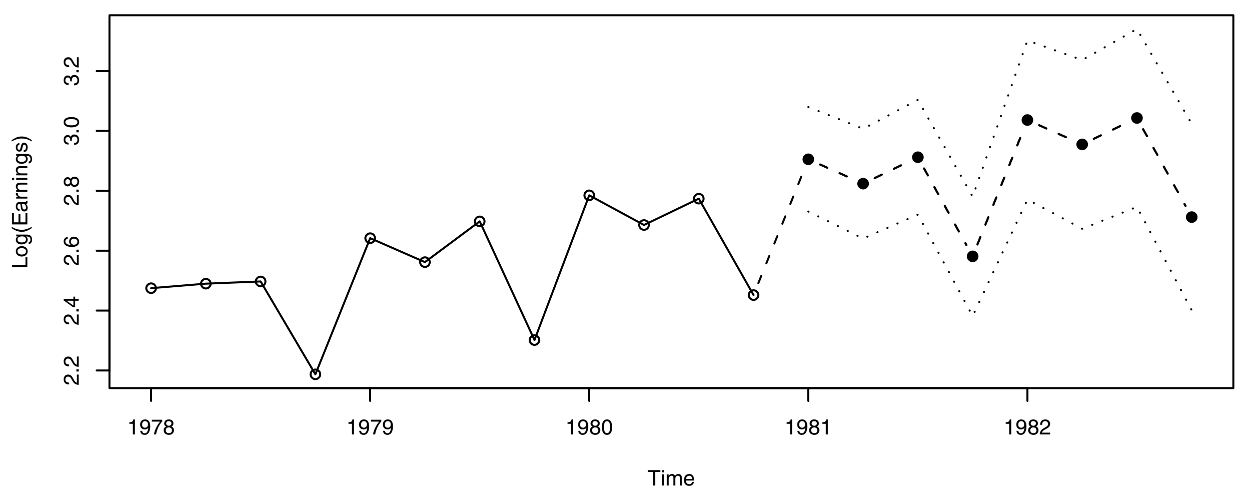



```
> win.graph(width=3,height=3,pointsize=8); hist(residuals(model),xlab='Residuals')
> qqnorm(residuals(model)); qqline(residuals(model)); shapiro.test(residuals(model))
```

```
        Shapiro-Wilk normality test

data:  residuals(model)
W = 0.9858, p-value = 0.489
```

Normality of the error terms looks like a very good assumption.

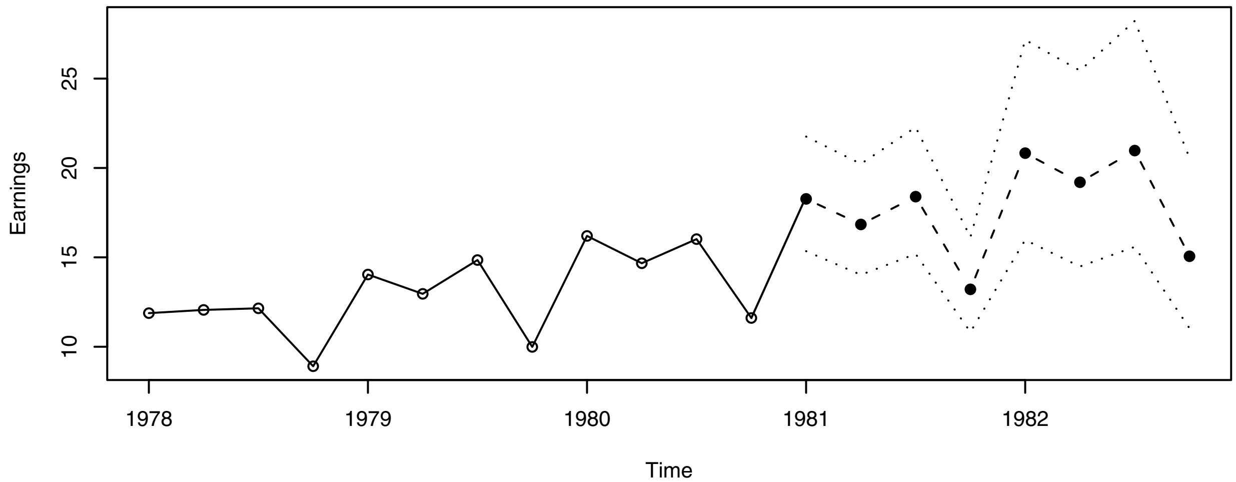
(h) Calculate and plot forecasts for the next two years of the series. Be sure to include forecast limits.




```
> win.graph(width=6.5,height=3,pointsize=8)
> plot(model,n1=c(1978,1),n.ahead=8,pch=19,ylab='Log(Earnings)')
```

The forecasts follow the general pattern of seasonality and "trend" in the earnings series and the forecast limits give a good indication of the confidence in these forecasts.

Lastly, we display the forecasts in orginal terms.

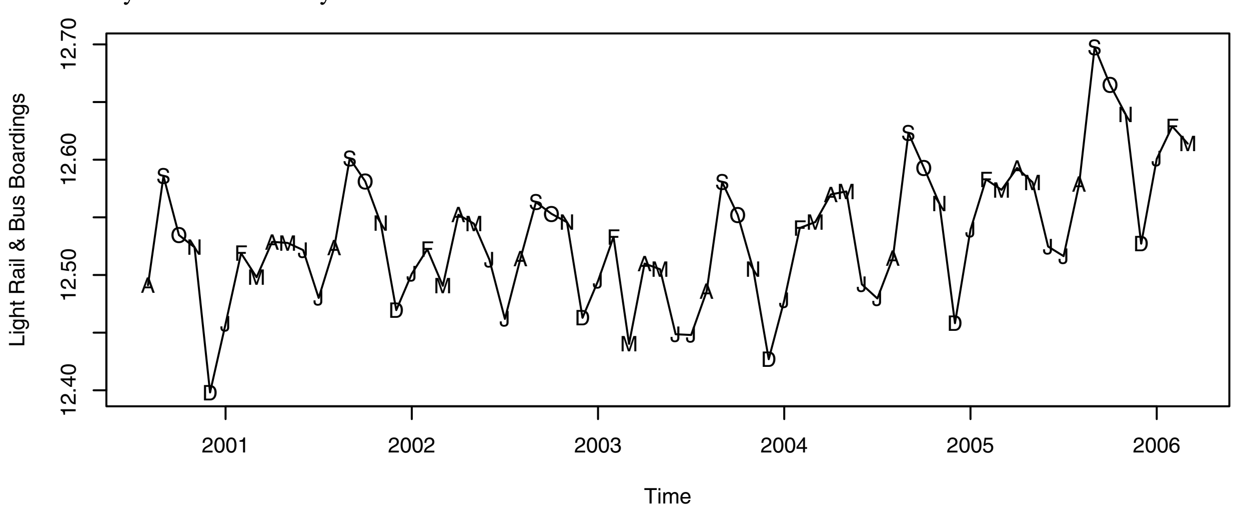



```
> plot(model,n1=c(1978,1),n.ahead=8,pch=19,ylab='Earnings',transform=exp)
```

In original terms, the uncertainty in the forecasts is easier to understand.

Exercise 10.12 The file named boardings contains monthly data on the number of people who boarded transit vehicles (mostly light rail trains and city buses) in the Denver, Colorado, region for August 2000 through December 2005.

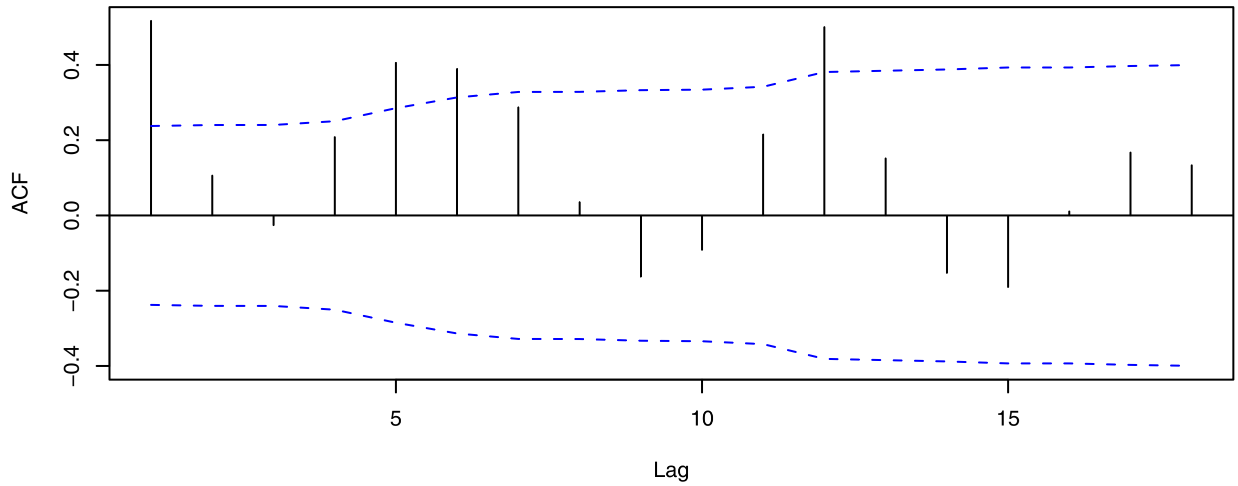
(a) Produce the time series plot for these data. Be sure to use plotting symbols that will help you assess seasonality. Does a stationary model seem reasonable?




```
> data(boardings); series=boardings[,1]
> plot(series,type='l',ylab='Light Rail & Bus Boardings')
> points(series,x=time(series),pch=as.vector(season(series)))
```

As one would expect, there is substantial seasonality in this series. Decembers are generaly low due to the holidays and Septembers are usually quite high due to the start up of school. There may also be a gradual upward “trend” that may need to be modeled with some kind of nonstationarity.

(b) Calculate and plot the sample ACF for this series. At which lags do you have significant autocorrelation?




```
> acf(as.vector(series),ci.type='ma')
```

There are significant autocorrelations at lags 1, 5, 6, and 12. Perhaps the following models will incorporate these effects.

(c) Fit an ARMA(0,3)×$(1,0)_{12}$ model to these data. Assess the significance of the estimated coefficients.

```
> model=arima(series,order=c(0,0,3),seasonal=list(order=c(1,0,0),period=12)); model
```

```
Call:
arima(x = series, order = c(0, 0, 3), seasonal = list(order = c(1, 0, 0), period
= 12))

Coefficients:
         ma1     ma2     ma3    sar1  intercept
      0.7290  0.6116  0.2950  0.8776    12.5455
s.e.  0.1186  0.1172  0.1118  0.0507     0.0354

sigma^2 estimated as 0.0006542:  log likelihood = 143.54,  aic = -277.09
```

All of these coefficients are statistically significant at the usual significance levels.

(d) Overfit with an ARMA(0,4)×$(1,0)_{12}$ model. Interpret the results.

```
> model2=arima(series,order=c(0,0,4),seasonal=list(order=c(1,0,0),period=12)); model2
```

```
Call:
arima(x = series, order = c(0, 0, 4), seasonal = list(order = c(1, 0, 0), period
= 12))

Coefficients:
         ma1     ma2     ma3     ma4    sar1  intercept
      0.7277  0.6686  0.4244  0.1414  0.8918    12.5459
s.e.  0.1212  0.1327  0.1681  0.1228  0.0445     0.0419

sigma^2 estimated as 0.0006279:  log likelihood = 144.22,  aic = -276.45
```

In this model, the added coefficient, ma4, is not statistically significant. Furthemore, the coefficients in common have changed very little—especially in light of the sizes of their standard errors. Finally, the AIC value is slightly better for the simpler model.

Chapter 11

Exercise 11.1

Let us draw a time series plot of the logarithms of data from January 1996 to May 2005 using different symbols to denote different months.

```
> ( )
> (( ))
> ( =10, =6,o =8)
> o ( (( ), = (( )), = )
> =c(" "," "," "," "," "," "," "," "," "," "," "," ")
> o ( (( ),o = )
```

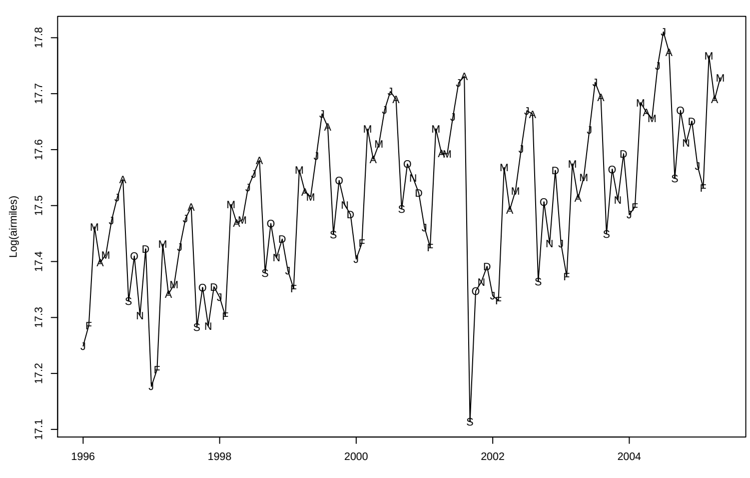


Figure 1: Plot of Log(airmiles) with Monthly Symbols

We can easily see that the air traffic is generally higher during July, August and December, and otherwise lower in January and Febrary. So the seasonality is obvious.

Exercise 11.2

Equation (11.1.7) gives

$$m_t = \begin{cases} \omega\frac{1-\delta^{t-T}}{1-\delta}, \text{ for } t > T \\ 0, \text{ otherwise.} \end{cases}$$

If $t \leq T$, then $m_t = m_{t-1} = 0$ and $S_{t-1}^{(T)} = 0$. Hence, for $t \leq T$ the "AR(1)" type recursion

$$m_t = \delta m_{t-1} + \omega S_{t-1}^{(T)} \tag{1}$$

holds.

When $t = T + 1$, we have $m_{T+1} = \omega, m_T = 0$ and $S_T^{(T)} = 1$. Recursion (1) still holds.

For $t \geq T + 2$, from (11.1.7) we get $m_t = \omega\frac{1-\delta^{t-T}}{1-\delta}, m_{t-1} = \omega\frac{1-\delta^{t-T-1}}{1-\delta}$ and $S_{t-1}^{(T)} = 1$. It's easy to check that (1) holds also in this case.

Exercise 11.3

When $\delta = 0.7$, the half-life for the intervention effect specified in Equation (11.1.6) is $\log(0.5) / \log(0.7) = 1.94$.

Exercise 11.4

The half-life for the intervention effect specified in Equation (11.1.6) is $\log(0.5) / \log(\delta)$, which satisfies $\lim_{\delta \nearrow 1} \log(0.5) / \log(\delta) = \infty$.

Exercise 11.5

For $t > T$, we have

$$m_t = \omega\frac{1-\delta^{t-T}}{1-\delta} = \omega\left(1 + \delta + \cdots + \delta^{t-T-1}\right) \to \omega(t-T), \quad \text{as } \delta \to 1.$$

For $t \leq T$, m_t is always 0. So

$$\lim_{\delta \to 1} m_t = \begin{cases} \omega(t-T), \text{ for } t \geq T \\ 0, \text{ otherwise.} \end{cases}$$

Exercise 11.6

(a) The intervention model is given by

$$\begin{aligned} & m_t = \frac{\omega B}{1-\delta B} S_t^{(T)} \\ \Longleftrightarrow\quad & m_t = \delta m_{t-1} + \omega S_{t-1}^{(T)}, \end{aligned}$$

which is the model given by Equation (11.1.6). With the initial condition $m_0 = 0$, we know $m_0 = m_1 = \cdots = m_T = 0$ and $m_{T+1} = \omega$. So the jump at time $T+1$ is of height ω.

(b) By Exercise 11.2 we know that

$$m_t = \begin{cases} \omega\frac{1-\delta^{t-T}}{1-\delta}, & \text{for } t > T \\ 0, & \text{otherwise.} \end{cases}$$

Since $0 < \delta < 1$, we obtain

$$\lim_{t\to\infty} m_t = \omega \lim_{t\to\infty} \frac{1-\delta^{t-T}}{1-\delta} = \frac{\omega}{1-\delta}.$$

Exercise 11.7

The intervention model is given by

$$\begin{aligned} & m_t = \frac{\omega B}{1-B} S_t^{(T)} \\ \iff\ & m_t = m_{t-1} + \omega S_{t-1}^{(T)}, \end{aligned}$$

with the initial condition $m_0 = 0$. Then $m_0 = m_1 = \cdots = m_T = 0$ and $m_t = m_{t-1} + \omega$ for $t \geq T+1$. So the effect increases linearly starting at time $T+1$ with slope ω.

Exercise 11.8

(a) The intervention model is given by

$$\begin{aligned} & m_t = \frac{\omega B}{1-\delta B} P_t^{(T)} \\ \iff\ & m_t = \delta m_{t-1} + \omega P_{t-1}^{(T)}, \end{aligned}$$

with the initial condition $m_0 = 0$. Then $m_0 = m_1 = \cdots = m_T = 0$ and $m_{T+1} = m_T + \omega = \omega$.

(b) For $t \geq T+2$, since $P_{t-1}^{(T)} = 0$, we have

$$\begin{aligned} & m_t = \delta m_{t-1} \\ \implies\ & m_t = \delta^{t-T-1}\omega \to 0, \quad \text{as } t \to \infty. \end{aligned}$$

Exercise 11.9

(a) The intervention model is given by

$$\begin{aligned} & m_t = \left[\frac{\omega_1 B}{1-\delta B} + \frac{\omega_2 B}{1-B}\right] P_t^{(T)} \\ \iff\ & m_t = (\delta+1)\, m_{t-1} - \delta m_{t-2} + (\omega_1+\omega_2)\, P_{t-1}^{(T)} - (\omega_1 + \delta\omega_2)\, P_{t-2}^{(T)}, \end{aligned} \tag{2}$$

with the initial condition $m_0 = m_1 = 0$. Then $m_0 = m_1 = \cdots = m_T = 0$ and $m_{T+1} = (\omega_1 + \omega_2) \cdot 1 = \omega_1 + \omega_2$.

(b) By (2), we have $m_{T+2} = (\delta + 1)(\omega_1 + \omega_2) - (\omega_1 + \delta\omega_2) = \delta\omega_1 + \omega_2$. It's easy by using the inductive method to prove that for all $t \geq T + 1$,

$$m_t = \delta^{t-T-1}\omega_1 + \omega_2 \to \omega_2, \quad \text{as } t \to \infty.$$

Exercise 11.10

(a) The intervention model is given by

$$\begin{aligned} & m_t = \left[\omega_0 + \frac{\omega_1 B}{1 - \delta B} + \frac{\omega_2 B}{1 - B}\right] P_t^{(T)} \\ \iff & m_t = (\delta + 1) m_{t-1} - \delta m_{t-2} + \omega_0 P_t^{(T)} + [\omega_1 + \omega_2 - (\delta + 1)\omega_0] P_{t-1}^{(T)} \\ & \qquad + (\delta\omega_0 - \omega_1 - \delta\omega_2) P_{t-2}^{(T)}, \end{aligned} \tag{3}$$

with the initial condition $m_0 = m_1 = 0$. Then $m_0 = m_1 = \cdots = m_{T-1} = 0$ and $m_T = \omega_0 \cdot 1 = \omega_0$.

(b) By (3), we have $m_{T+1} = (\delta + 1)\omega_0 + [\omega_1 + \omega_2 - (\delta + 1)\omega_0] = \omega_1 + \omega_2$.

(c) Again by (3), we have $m_{T+2} = (\delta + 1)(\omega_1 + \omega_2) - \delta\omega_0 + (\delta\omega_0 - \omega_1 - \delta\omega_2) = \delta\omega_1 + \omega_2$. It's easy by using the inductive method to prove that for all $t \geq T + 1$,

$$m_t = \delta^{t-T-1}\omega_1 + \omega_2 \to \omega_2, \quad \text{as } t \to \infty.$$

Exercise 11.11

The sample CCF between the generated 100 pairs (X_t, Y_t) is shown below.

```
> .  . (12345)
> =2*     (105)
> = - ( ,3)+     (105)
> = ( -(1:5)],  ' =1   =1)
> = ( -(1:5)],  ' =1   =1)
> 'o (  =4.875,    =2.5,o     =8)
> ( , , '=' ')
```

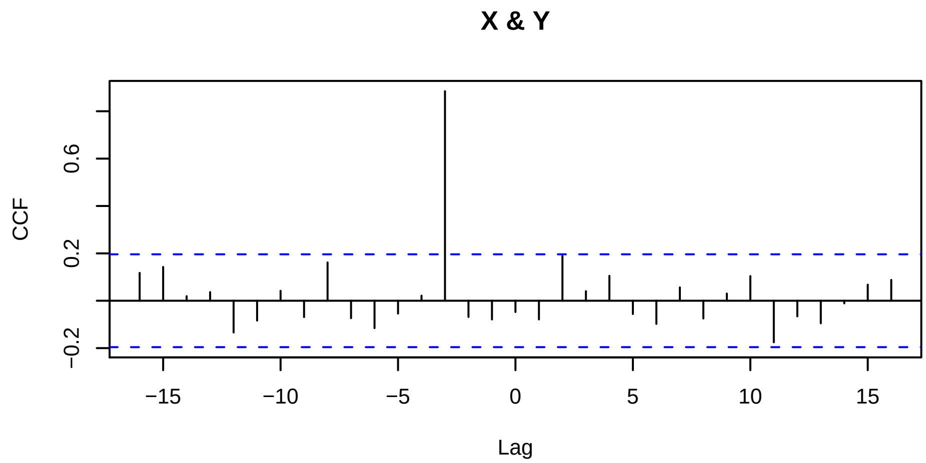


Figure 2: Sample CCF From Equation (11.3.1) with $d = 3$

The sample CCF of the simulated data is only significant at lag -3. This result coincides with the theoretical analysis. Theoretically, the CCF should be zero except at lag -3 where it equals $\rho_{-3}(X,Y) = 2/\sqrt{4+1} = 0.89$.

Exercise 11.12

Suppose X and Y are two independent and stationary AR(1) time series with parameters ϕ_X and ϕ_Y. We know that the autocorrelations of X and Y are $\rho_k(X) = \phi_X^k$ and $\rho_k(Y) = \phi_Y^k$ respectively. Equation (11.3.5) tells us the variance of $\sqrt{n}r_k(X,Y)$ is approximately

$$\begin{aligned} & 1 + 2\sum_{k=1}^{\infty} \rho_k(X)\rho_k(Y) \\ = & 1 + 2\sum_{k=1}^{\infty} \phi_X^k \phi_Y^k = \frac{1+\phi_X\phi_Y}{1-\phi_X\phi_Y}. \end{aligned}$$

So the variance of $r_k(X,Y)$ is approximately

$$\frac{1+\phi_X\phi_Y}{n(1-\phi_X\phi_Y)},$$

which is given by Equation (11.3.6).

Exercise 11.13

Since

$$\widetilde{Y}_t = \sum_{k=-\infty}^{\infty} \beta_k \widetilde{X}_{t-k} + \widetilde{Z}_t,$$

where $\widetilde{X}$ is a white noise sequence independent of $\widetilde{Z}$, we have

$$\begin{aligned}\rho_k\left(\widetilde{X}, \widetilde{Y}\right) &= \frac{cov\left(\widetilde{Y}_t, \widetilde{X}_{t+k}\right)}{\sigma_{\widetilde{X}}\sigma_{\widetilde{Y}}} \\ &= \frac{\beta_{-k}\sigma^2_{\widetilde{X}}}{\sigma_{\widetilde{X}}\sigma_{\widetilde{Y}}} = \beta_{-k}\frac{\sigma_{\widetilde{X}}}{\sigma_{\widetilde{Y}}}.\end{aligned}$$

Exercise 11.14

The following three graphs are a simulated AR time series and its ACF and PACF plots.

```
>  =‘  ‘.  (   =  (‘ =.7), =48)
> o  ( ,  o=’ ’)
> ‘ ( )
> o‘ ( )
```

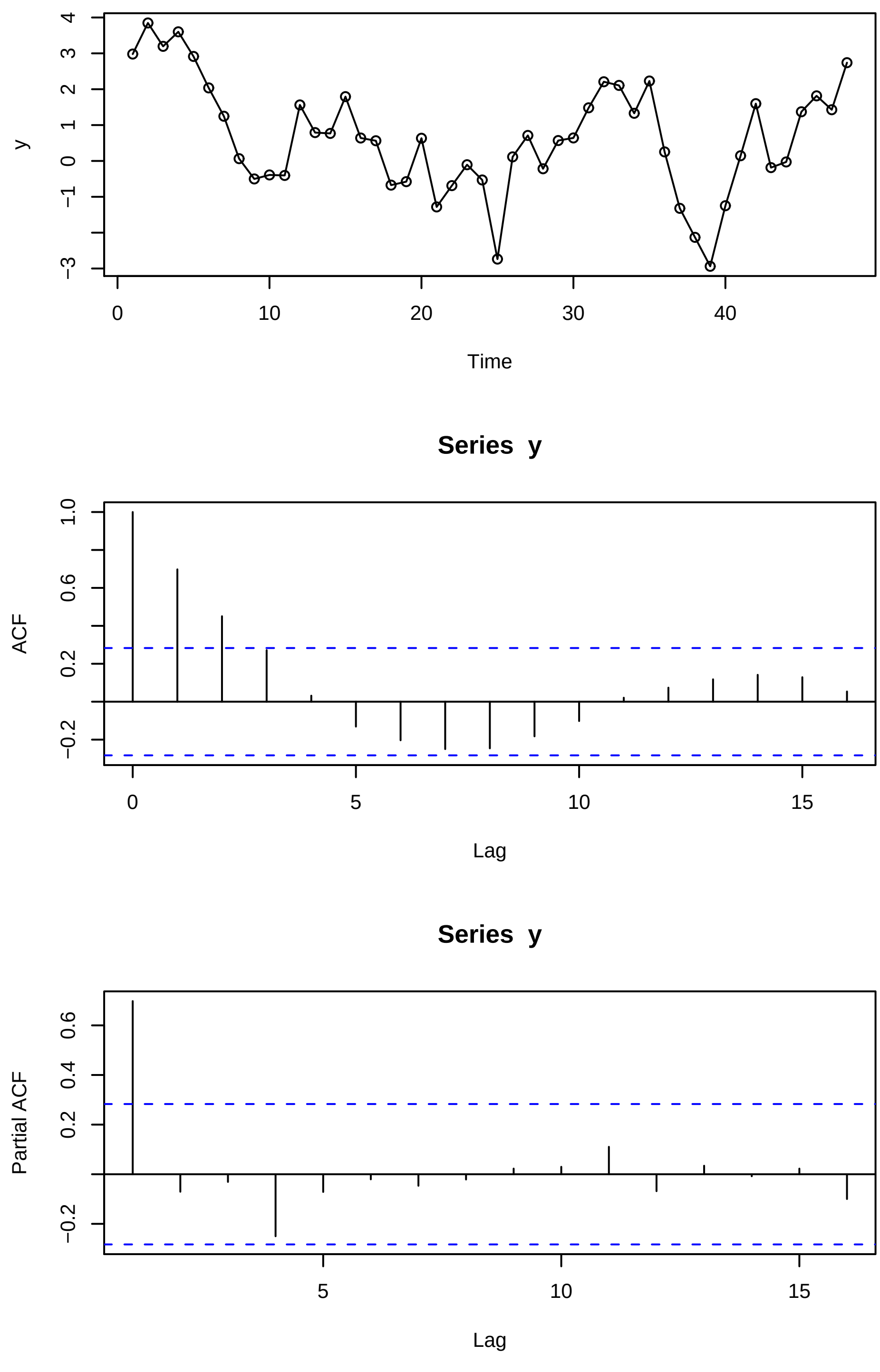


Figure 3: Simulated AR(1) Time Series

(a) If we add a step function response of $\omega = 1$ at time $t = 36$ to the simulated AR(1) time series, then the plots of the new time series and its ACF and PACF are given below.

```
> ‘=o( o(0,35), o(1,13))
>  1= +‘
> o   ( 1,  o ='  ')
> ‘ ( 1)
> o‘ ( 1)
```

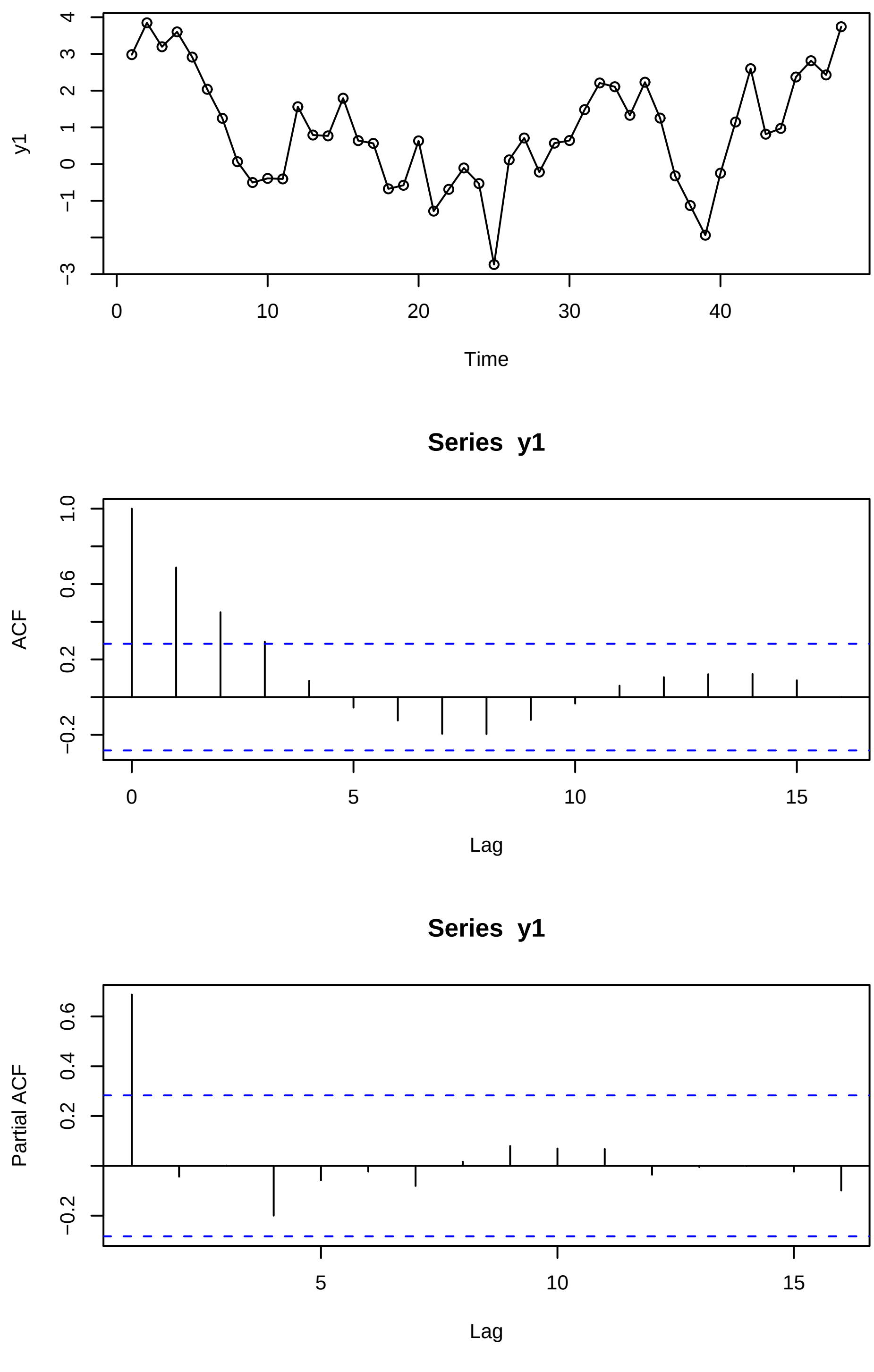


Figure 4: Simulated AR(1) Time Series Plus a Step Response

(b) If we add a pulse function response of $\omega = 1$ at time $t = 36$ to the simulated AR(1) time series, then the plots of the new time series and its ACF and PACF are given below.

```
> =D( o(0,35),1, o(0,12))
>  2= +
> o   ( 2,  o =' ')
> ' ( 2)
> o' ( 2)
```

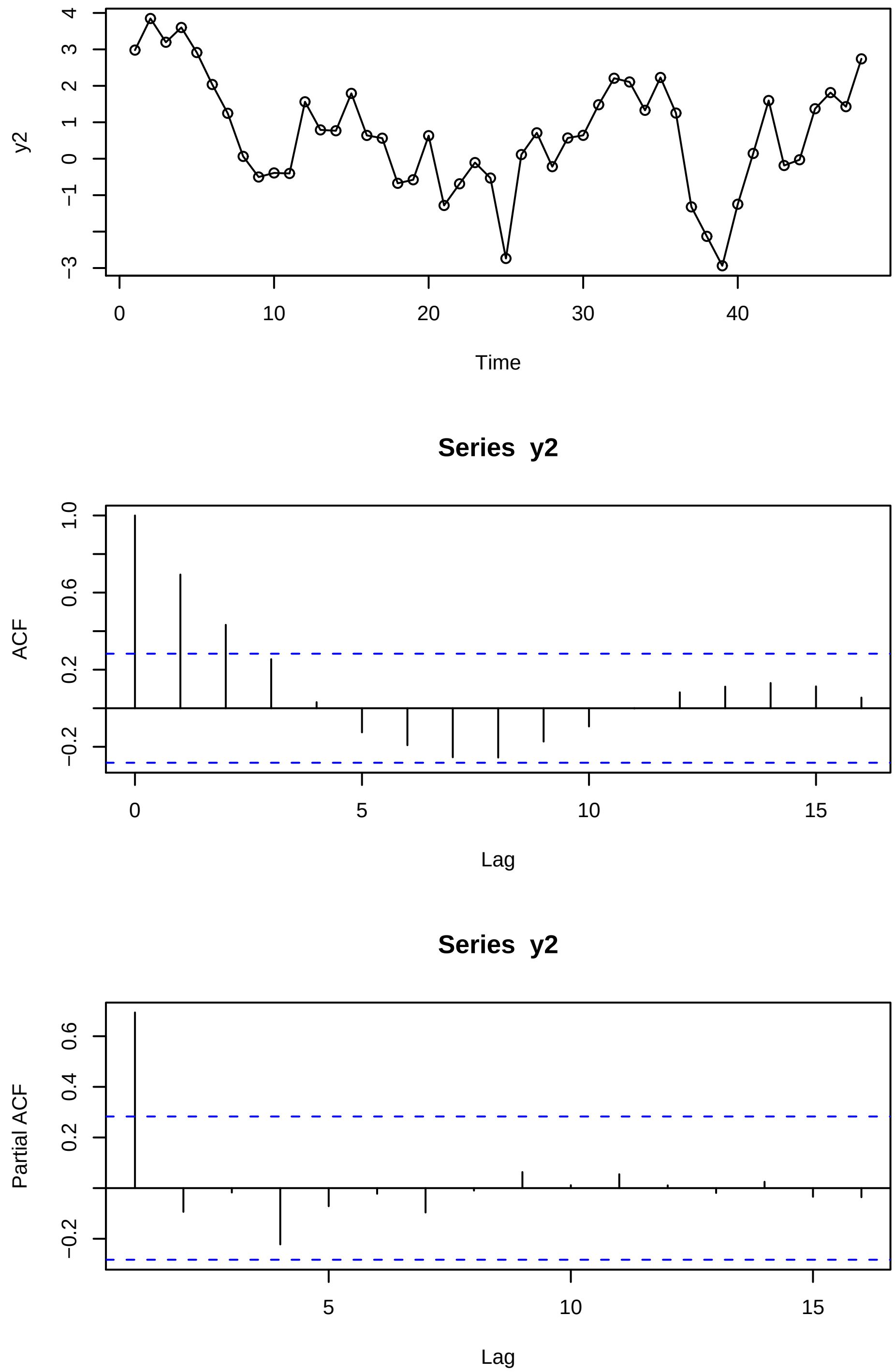


Figure 5: Simulated AR(1) Time Series Plus a Pulse Response

Exercise 11.15

(a) To verify Exhibit 11.5:

```
> acf(diff(diff(window(log(airmiles),end=c(2001,8)),12)),lag.max=48)
```

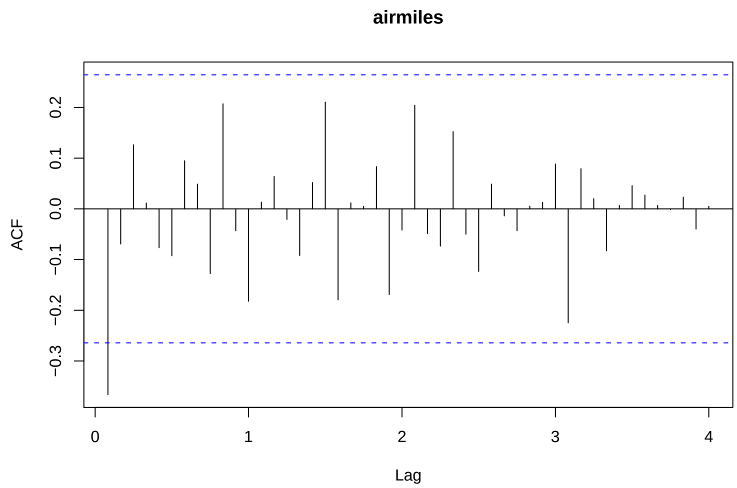


Figure 6: ACF of log(airmiles)

(b) The plot above tells us that the ACF for the twice differenced series of logarithms of the airmiles data is significantly away from 0 only in the case that the lag is 1/12 (1 month). It suggests an ARIMA$(0,1,1)\times(0,1,0)_{12}$ model.

```
> m1.airmiles=arima(window(log(airmiles),end=c(2001,8)),order=c(0,1,1),
+ seasonal=list(order=c(0,1,0),period=12))
> m1.airmiles

Call:
arima(x = window(log(airmiles), end = c(2001,8)), order = c(0,1,1),
    seasonal = list(order = c(0, 1, 0), period = 12))

Coefficients:
          ma1
      -0.4391 **
s.e.   0.1173

sigma^2 estimated as 0.00143:  log likelihood = 101.98,  aic=-199.96
```

The coefficient estimate is significant. Then let's look at the time series plot of the residuals from this model.

```
> o   (   ' c ' c ( 1.'      ),  '='   ' c ' c        '   ',   ='')
> '     ( =0)
```

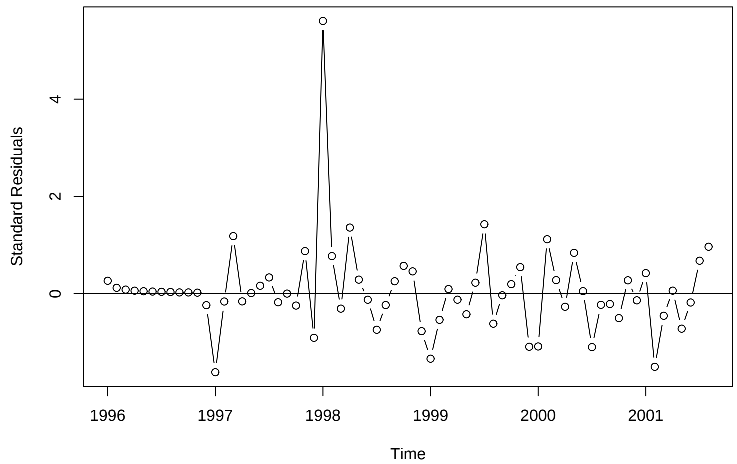


Figure 7: Residuals of model 1

From the residuals plot, we see that except for January 1998 the residuals are distributed around 0 with no pattern. It leads us to check for possible outliers.

```
> c    o    ( 1.'          )
             ,1
  c       25.000000  '  c'2  8.114302 ***
> c    o    ( 1.'          )
             ,1
  c       25.000000  '  c'1  8.434749 ***
```

We see that both $\widehat{\lambda}_1$ and $\widehat{\lambda}_2$ are highly significant (much larger than the critical value with $\alpha = 5\%$ and $n = 68$). So next we are going to add the innovative outlier to the previous model.

(c) Let's fit an ARIMA$(0,1,1) \times (0,1,0)_{12}$+an outlier model to the logarithms of the preintervention data.

```
> 2.‘      =‘    ‘(        (     (‘‘      ),    =(2001,8)),    =(0,1,1),
+   ‘   ‘ =    (    =(0,1,0),    =12),    =(25))
> 2.‘
 ‘ : ‘   ‘(  =      (    (‘‘      ),    =(2001,8)),    =(0,1,1),
      ‘   ‘ =    (     = (0, 1, 0),     = 12),     = (25))

           :
          ‘1          -25
      -0.3894 **    0.2132 ***
 .     0.0888       0.0248

   ‘ 2    ‘   ‘   0.0006092:          = 125.47,‘ =-244.94
```

In this model, the IO effect is highly significant. Compared with the previous model without outlier, the AIC is much better and $\widehat{\theta}$ has changed significantly. The residuals plot of this model is given by

```
>    (   ‘ ‘ ( 2.‘      ),  ‘ =’   ‘ ‘          ‘    2’,    =’ ’)
```

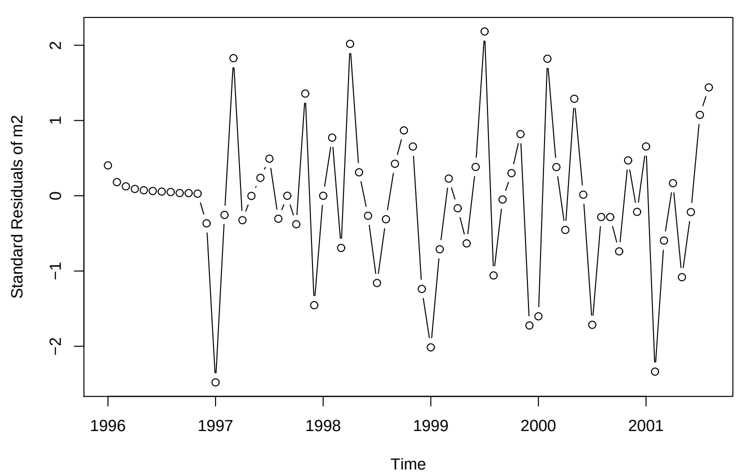


Figure 8: Residuals of model 2

The residuals look distributed much better. Furthermore, let's look at the sample ACF of the residuals.

```
> ‘ (‘ . (  ‘ ‘ ( 2.‘  )), ‘ . ‘ =68)
```

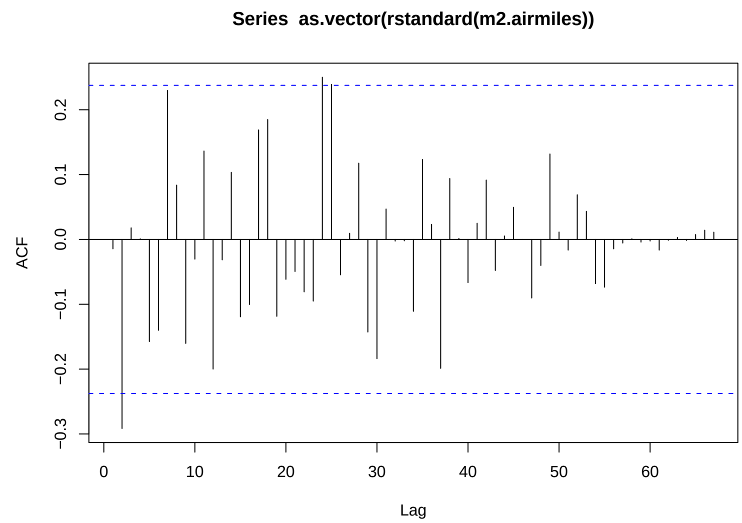


Figure 9: ACF of residuals of model 2

We see from the graph above that the ACF of the residuals are significantly different from 0, at lag 2, 7 , 24, and 25. This means that there is still some autocorrelation between the residuals. One more coefficient is needed. Next we'll fit the model ARIMA$(0,1,1) \times (0,1,1)_{12}$+an outlier.

(d) Let's fit an ARIMA$(0,1,1) \times (0,1,1)_{12}$+an outlier model to the logarithms of the preintervention data.

```
> 3.‘  =‘ ‘ ( ( (‘  ), =(2001,8)), =(0,1,1),
+ ‘ ‘ = ( =(0,1,1), =12), =(25))
> 3.‘
‘ : ‘ ‘ ( = ( (‘  ), =(2001,8)), =(0,1,1),
    ‘ ‘ = ( = (0, 1, 1), = 12), = (25))

 :
        ‘1          ‘1       -25
     -0.4030***  -0.2762**  0.2113***
 .    0.0801      0.0894    0.0235
```

```
sigma^2 estimated as 0.0005116:  log likelihood = 129.79,  aic=-251.58
```

All coefficient estimates are significantly different from 0. Compared with model 2, all of the estimates have not changed too much. But the autocorrelation between residuals amostly disappears as shown below.

```
> acf(as.vector(rstandard(m3.airmiles)),lag.max=68)
```

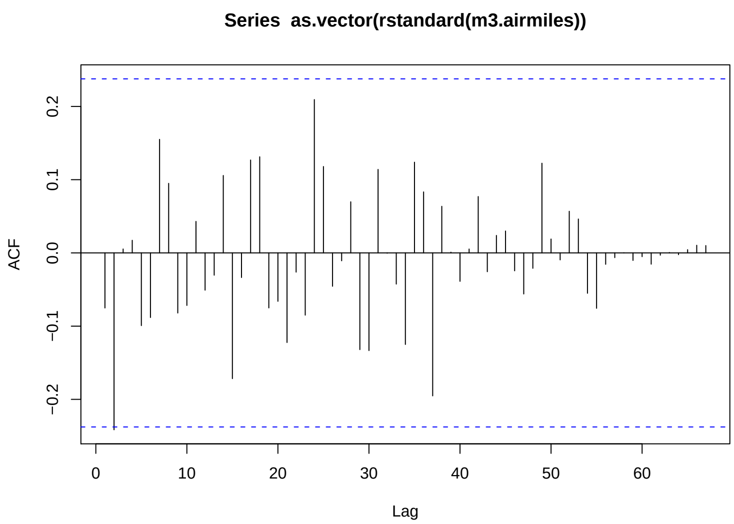


Figure 10: ACF of residuals of model 3

We see that the ACF is only significantly different from 0 at lag 2. This could easily happen by chance alone. So under this model there is no autocorrelation between residuals. Let's also check the normality of the residuals.

```
>     'o (     =2.5,       =2.5,o        =8)
>        (    ' ' ( 3.'       ))
>        (    ' ' ( 3.'       ))
```

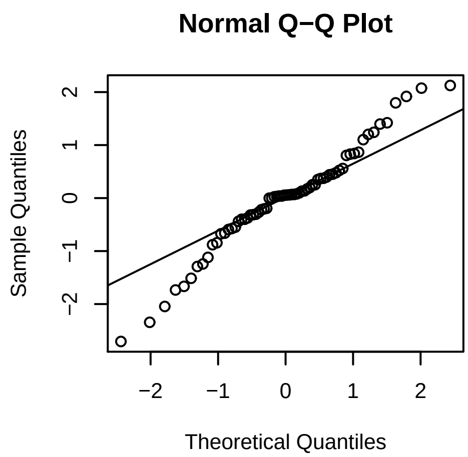


Figure 11: Q-Q norm

Here the normal Q-Q plot looks that the residuals are not samples from a normal distribution, while the Shapiro-Wilk test of normality has a test statistic $W = 0.969$ and $p - value = 0.088$ and normality is not rejected at significant level of 5%. We conclude that the ARIMA$(0,1,1) \times (0,1,1)_{12}$+an outlier model fits the preintervention data well.

Exercise 11.16

(a) The plot of monthly boardings using seasonal symbols is given below.

```
>      '  (    )
>   ' (  '       )
> o   (  '        ,1 ,   ' ='   '        '    ' =' '  ')
>        =  ("  ","  ,"  ","  ,"  ","  ","  ","  ","  ","  ","  ","  ")
> o      (  '        ,1 ,o  =        )
```

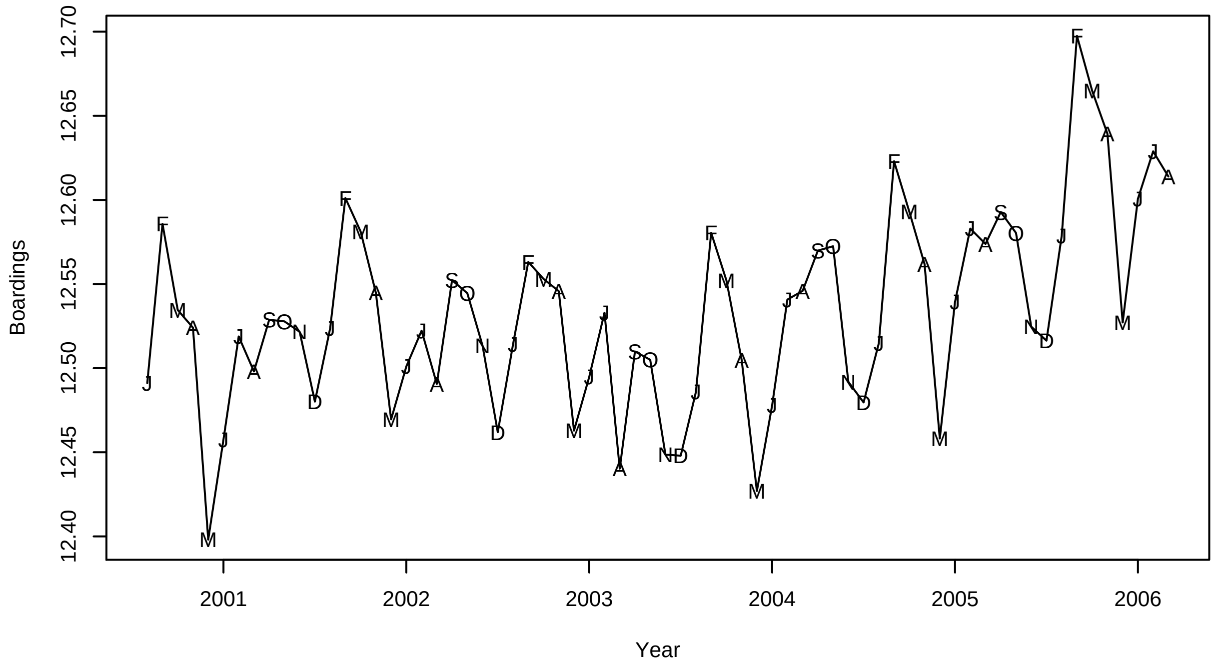


Figure 12: Logarithm of Monthly Boardings

We notice that during the period between August 2000 and March 2006, the peak public transportation boarding appears in February, September and October, while in May and December it comes to the bottom.

(b) The plot of monthly average gasoline prices using seasonal symbols is given below.

```
> o    (  ‘      ,2 ,   ‘ =’ ‘        ’,  ‘ =’  ‘  ’)
>        = ("  ","  ","  ","  ","  ","  ","  ","  ","  ","  ","  ","  ")
> o      (  ‘      ,2 ,o  =        )
```

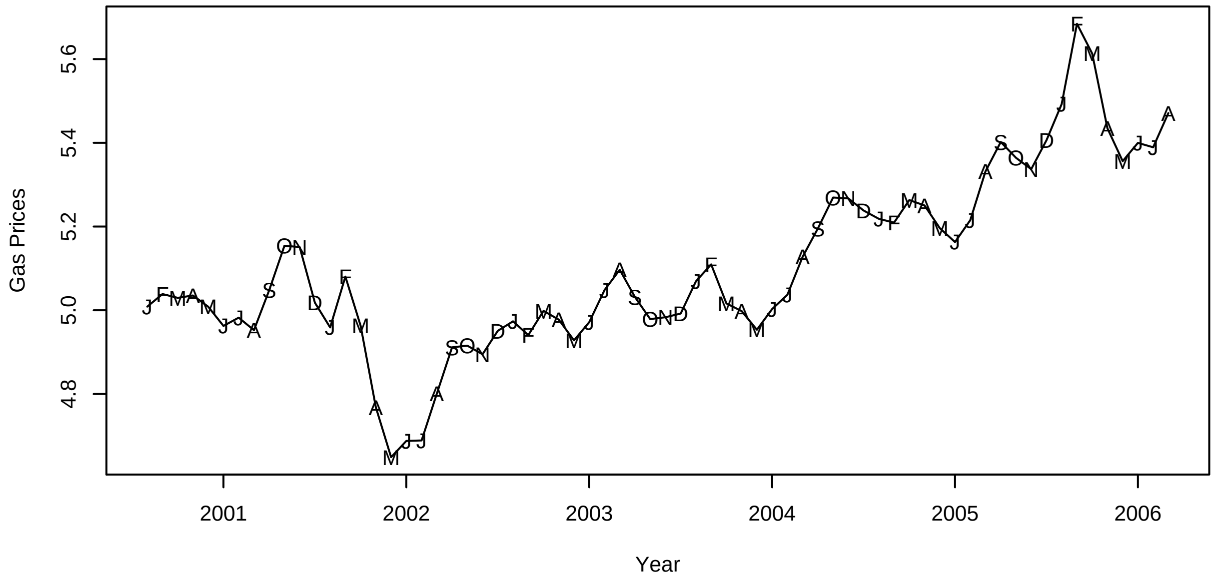


Figure 13: Logarithm of Monthly Average Gasoline Prices

From the graph we see that the average gasoline price is generally higher during Febrary, September and October, while lower during May.

Exercise 11.17

(a) Fit an AR(2) model to the original data including the outlier.

```
>     (    )
>     (    1)
>   1<-    (    1,     = (2,0,0))
>   1

   :     (  =    1,    = (2, 0, 0))

          :
       1      2
     0.0269  0.2392    1.4135
 .   0.1062  0.1061    0.6275

    2             17.68:            = -234.19,    = 476.38
```

(b) Test the possible AO and IO outliers. Both tests tell us at time $t = 27$ there is an outlier.

```
> ( 1)
         ,1
          27.000000
  '2   8.668582
> ( 1)
         ,1
          27.00000
  '1   8.81655
```

(c) Since $\widehat{\lambda}_1 > \widehat{\lambda}_2$, let us refit the AR(2) model incorporating an IO outlier at time $t = 27$.

```
>  2<-'  ( 1,  =(2,0,0),  =(27))
>  2


 :
  ( = 1,   = (2, 0, 0),   = (27))

   :
        1     2        -27
    -0.0018  0.2143    0.9754  28.5946
 .   0.0713  0.0713    0.3960   2.8263

  2       7.866:          = -200.96,   = 409.93
```

The IO is found to be highly significant. The other coefficient estimates are slightly changed, although the intercept term (the mean) is affected more.

(d) Model diagnostics can be carried out by the following command

```
>    ( 2)
```

which yields Figure 14 shown on the next page.

Now, the fitted model passes all model diagnostics. In particular, there are no more outliers as also confirmed by the formal tests.

```
> ( 2)
 1           ''
> ( 2)
 1           ''
```

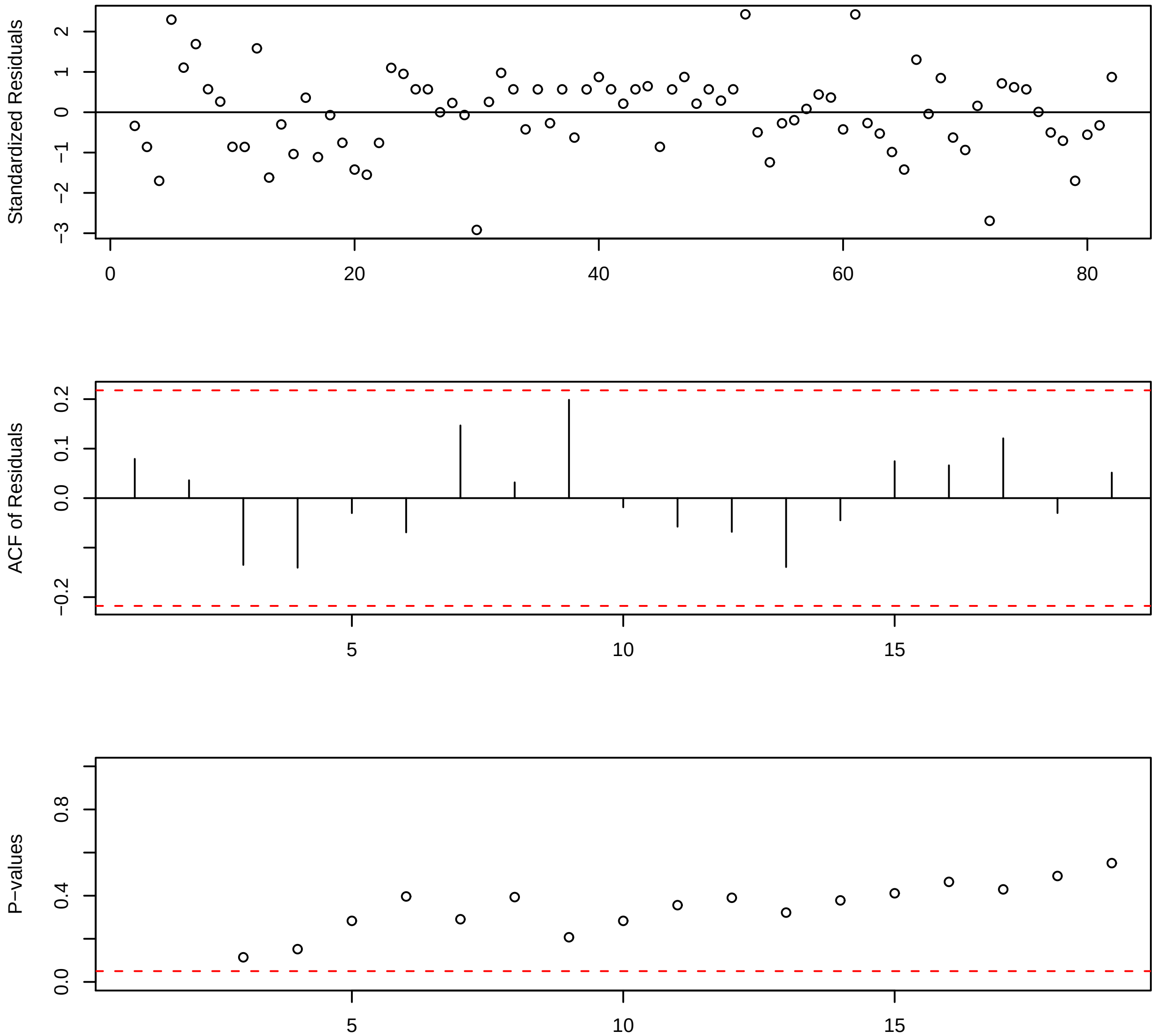


Figure 14: Model diagnostics

Exercise 11.18

(a) Fit the MA(2) model and test it for both AO and IO outliers.

```
> ‘ ‘( ‘   )
>   1<-‘  ‘( ‘   ,    =(0,0,2))
```

```
> ( 1)
        ,1      ,2
     63.000000 129.000000
  '2  4.009568   5.344322
> ( 1)
        ,1      ,2
     63.000000 129.000000
  '1  4.081066   5.268322
```

(b) Fit the MA(2) incorporating the outliers into the model.

```
> 2=  (  ,  =(0,0,2), =  (  =  ( )==129))
> ( 2)
        ,1      ,2
     63.000000 106.000000
  '2  4.369851   3.563718
> ( 2)
        ,1      ,2
     63.000000 106.000000
  '1  4.566191   3.654841
> 3=  (  ,  =(0,0,2), =  (  =  ( )==
129), =(63))
> 3

 :
  ( = ,  = (0, 0, 2), =  (  =  ( )
== 129),
   = (63))

      :
      '1      '2                        -63
     0.2283  0.1631     28.0764  36.6490  28.4814
 .   0.0823  0.0738      0.7259   5.8214   5.9752

  2    35.10:      = -415.8,  = 841.6
```

(c) The fit of the above model is then assessed.

```
> ( 3)
```

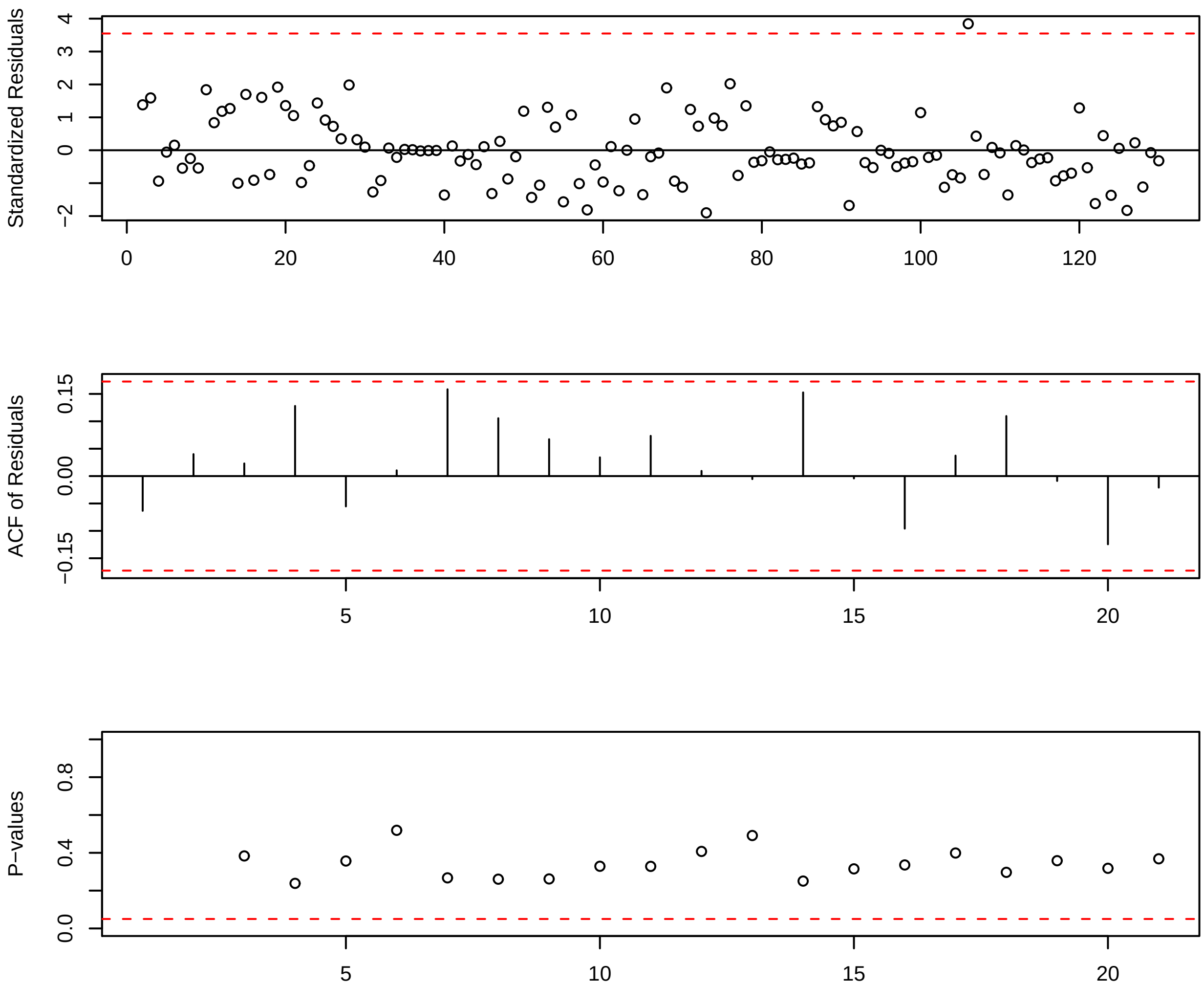


Figure 15: Model diagnostics

An outlier is apparent from the residual plot, with formal tests reported below suggesting an IO at time point 106.

```
>           ( 3)
                  ,1
           106.000000
             3.928688
          12              (               (                 )
>           ( 3)
```

```
                 ,1
        106.000000
   '2    3.805427
)
```

(d) We now incorporate all three outliers in the following.

```
>  4=       (    ,      =(0,0,2),       =             (   =       )==
129),   =(63,106))
>  4


    :
                   =       ,       = (0, 0, 2),       =              (     =          )
==    129),
          = (63, 106))

             :
        '1     '2                             -63      -106
      0.2472  0.1770    27.8145  36.8519  28.7753  23.0643
 .  . 0.0768  0.0683     0.7027   5.4496   5.6200   5.6065

       2              31.06:                    = -407.85,      = 827.7
>        ( 4)
```

All outliers are found to be significant.

No more outliers are found from the time plot of the residuals, which is also confirmed by formal tests (not reported). However, the Ljung-Box test and the residual ACF plot suggest that there is remaining serial autocorrelation. The residual ACF is significant at lag 7 suggesting perhaps a seasonal MA(1) pattern with period 7. We subsequently fitted an enlarged model that adds a seasonal MA(1) coefficient to the above model. However, the seasonal MA(1) coefficient is not significant, see below. Hence, we conclude that the MA(2) plus three outliers model provides a marginally adequate fit to the data.

```
>  5=       (    ,      =(0,0,2),        =     (      =(0,0,1),
       =7),      =          (   =       )== 129),   =(63,106))
>  5


    :
```

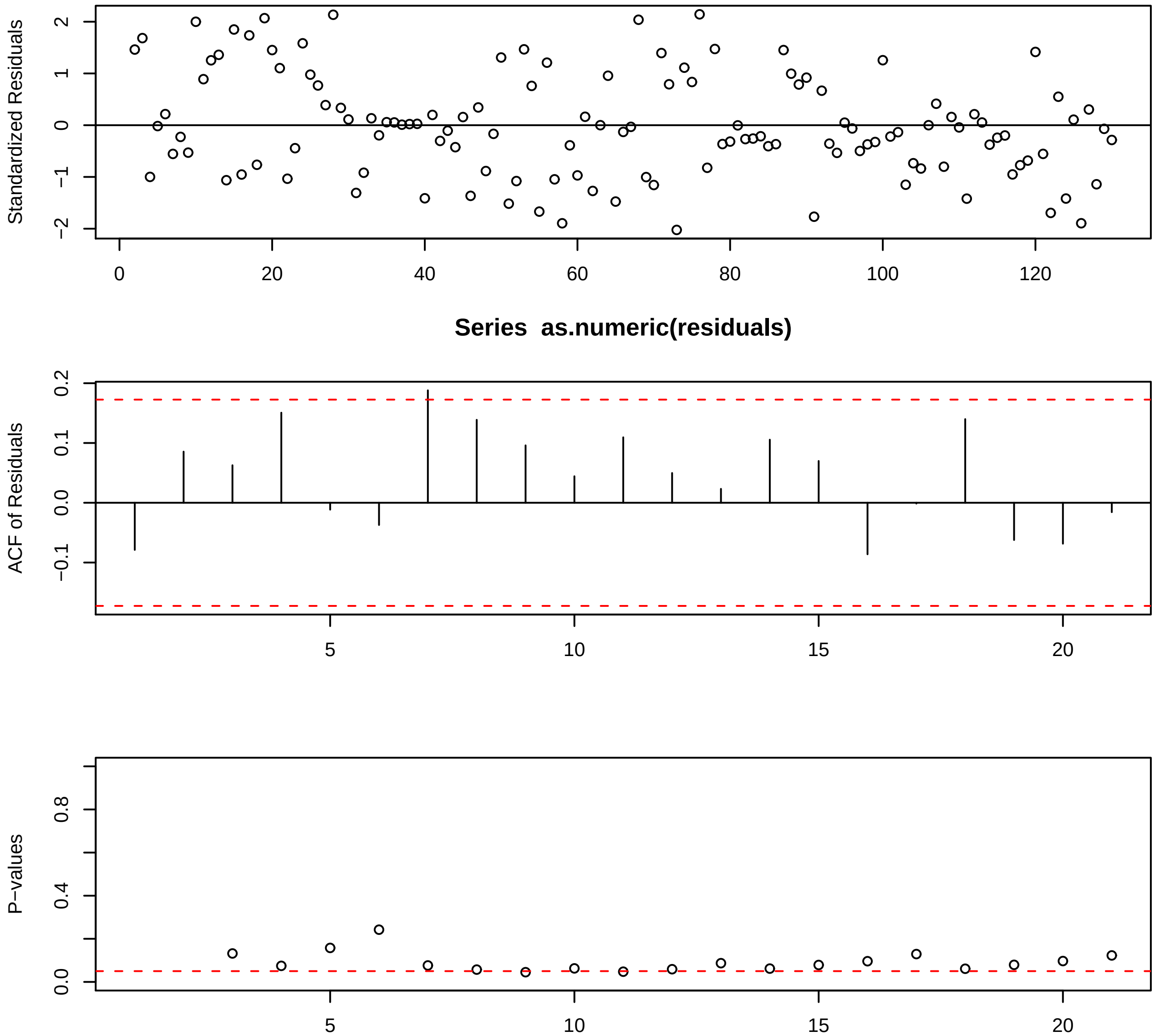


Figure 16: Model diagnostics

```
  (   =    ,      = (0, 0, 2),        =    (    = (0, 0,
1),      = 7),
        =       (   =   (   ) == 129),    = (63, 106))

        :
```

	‘1	‘2	‘1			-63	-106
	0.2432	0.1729	0.0899	27.7658	37.7247	28.1777	23.2698
.	0.0767	0.0698	0.0713	0.7544	5.4619	5.5982	5.5740

‘ 2 ‘ ‘ 30.67: = -407.06, ‘ = 828.13

Exercise 11.19

Let us see the plot of the log-transformed weekly unit sales of lite potato chips and the weekly average price over a period of 104 weeks.

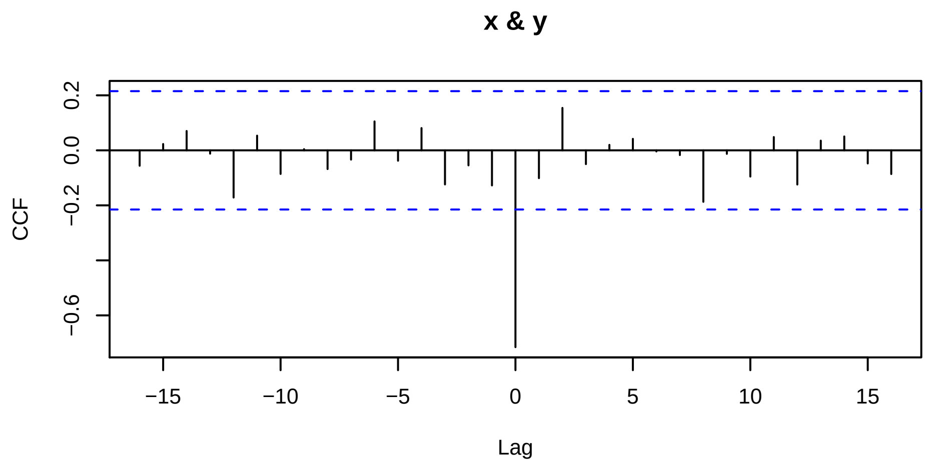


Figure 18: Sample Cross Correlation Between Prewhitened Differenced Log(Sales) and Price of Lite Potato Chips

We estimate the coefficients from the OLS regression of log(sales) on price and get the following estimates.

```
>     =                  ,1 ;o   =              ,2
>  1=   (       o      ;      =                )
>          ( 1)

      :

                =          o    ,            =              )

             :
                   1                         3
-0.47884 -0.13992   0.01661   0.11243   0.60085

               :
                                            P (>   )
(          o )  13.7894      0.2345   58.81    <2  -16 ***
o                -2.1000      0.1348  -15.57    <2  -16 ***
---
                  :   0   ***' 0.001   **' 0.01   *' 0.05   .' 0.1    ' 1

                        : 0.1963     102
```

```
Multiple R-squared: 0.7039,    Adjusted R-squared: 0.701
  F-statistic: 242.5 on 1 and 102 DF,  p-value: < 2.2e-16
```

The residuals are, however, autocorrelated as can be seen from their sample ACF and PACF displayed below respectively. Indeed, the sample autocorrelations of the residuals are significant for the first 8 lags whereas the sample partial autocorrelations are significant at lags $1, 2, 4$ and 16.

```
> acf(residuals(m1),ci.type='ma')
```

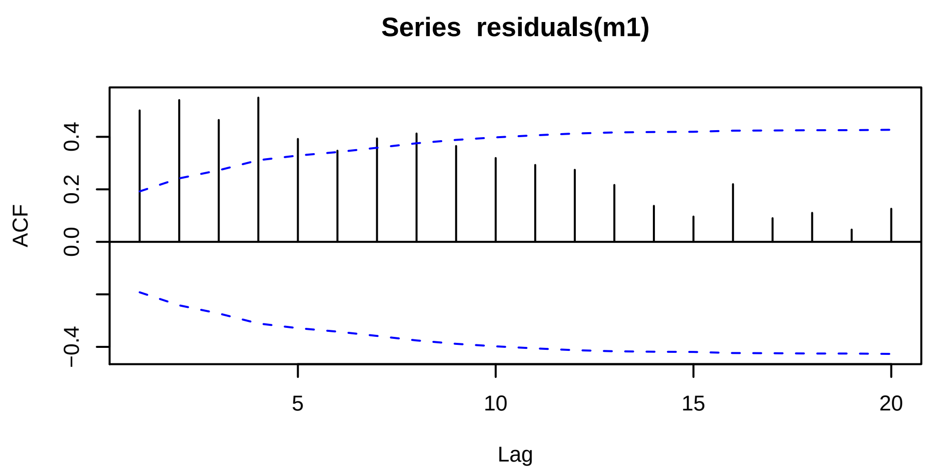


Figure 19: Sample ACF of Residuals from OLS Regression of Log(Sales) on Price

```
> pacf(residuals(m1))
```

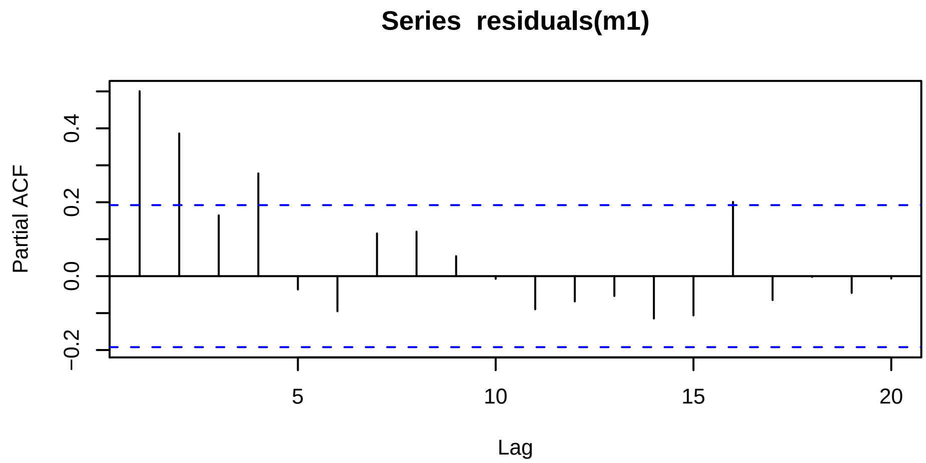


Figure 20: Sample PACF of Residuals from OLS Regression of Log(Sales) on Price

The sample EACF of the residuals, shown below, contains a triangle of zeroes with a vertex at (1,5), thereby suggesting an ARMA(1,5) model. Hence, we fit a regression model of log(sales) on price with an ARMA(1,5) error.

```
> ( ( 1))

  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0
1
2
3
4
5
6
7
```

It turns out that the estimates of the π_1, θ_1 and θ_5 coefficients are not significant and hence a model fixing these coefficient to be zero was subsequently fitted and reported below.

```
>  2=  ((  ,  = (1,0,5),  =  )
>  3=  ((  ,  = (0,0,4),  =  ,  = (0,  ,  ,  ,  ,  ))
>  4=  ((  ,  = (0,0,4),  =  ,  = (0,  ,0,  ,  ,  ))
>  4
```

```
  :        =        ,        = (0, 0, 4),        =      ,         =
(0,    , 0,    ,    ,    ))

          :
          1        2    3        4
          0  0.4354     0  0.5451     13.5306  -1.9505
  .       0  0.0827     0  0.0926      0.1785   0.1016

      2               0.02360:                 = 46.45,       = -82.9
```

Note that the regression coefficient estimate on Price is similar to that from the OLS regression fit earlier but the standard error of the estimete is about 25% lower than that from the simple OLS regression.

Exercise 11.20

(a) The data shows an increasing trend with quasi-periodic behavior.

```
>      (     )
>      (     ,    ='  '     ='    ')
```

(b)

```
>    1=  (                (     ))
>          (   1)


    :
            =                (     ))


          :
                       1                    3
-43.184 -23.096    6.212   19.289   37.258


          :
                                              (>   )
(          ) -6532.1466   1466.7183   -4.454 0.000199 ***
     (     )     3.3449      0.7357    4.546 0.000159 ***
---
```

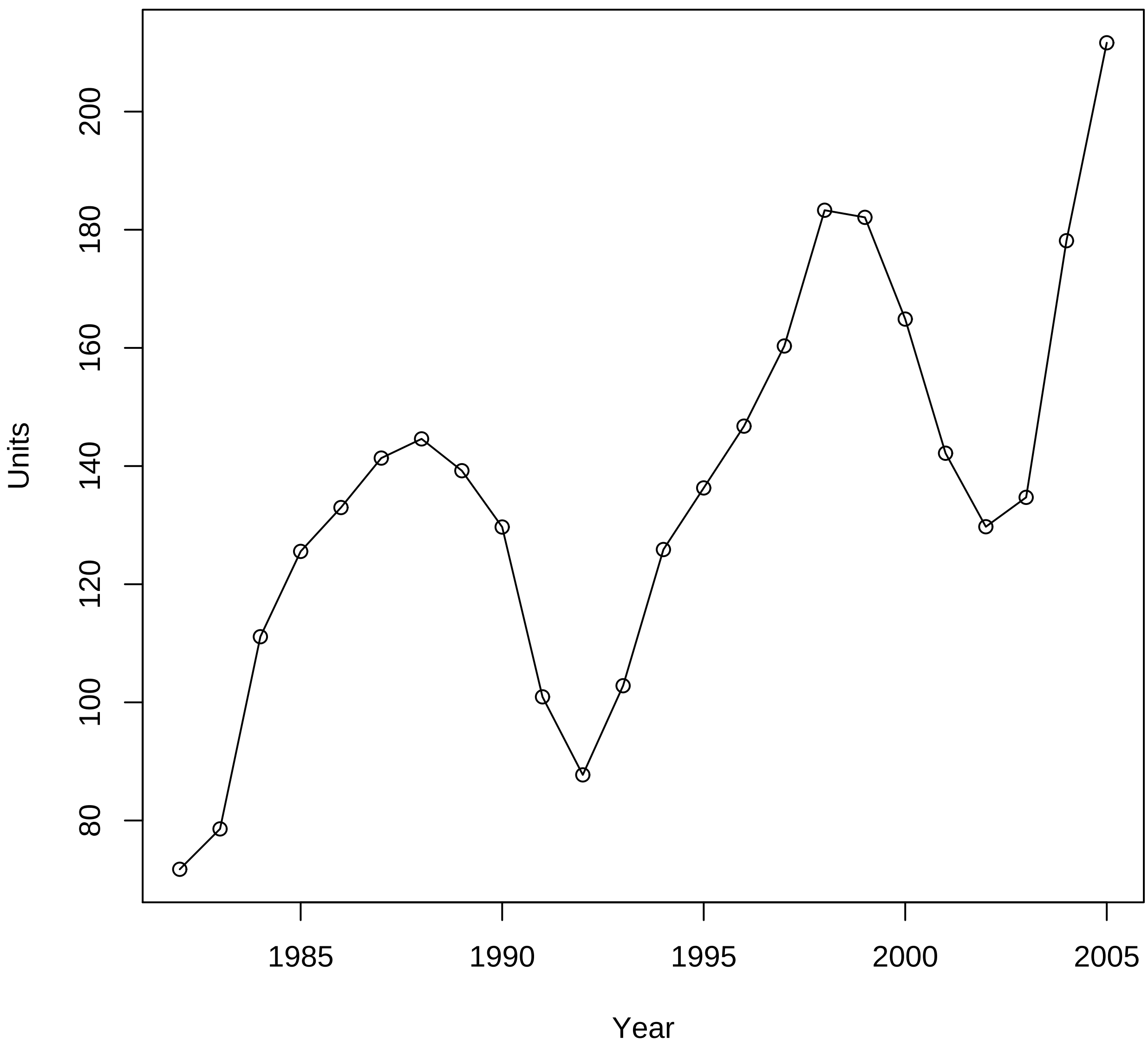


Figure 21: Time plot of the annual units sale.

```
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 24.95 on 22 degrees of freedom
Multiple R-squared: 0.4844,     Adjusted R-squared: 0.461
F-statistic: 20.67 on 1 and 22 DF,  p-value: 0.0001589
```

The fitted straight line is found to have a significant, positive slope.

(c)

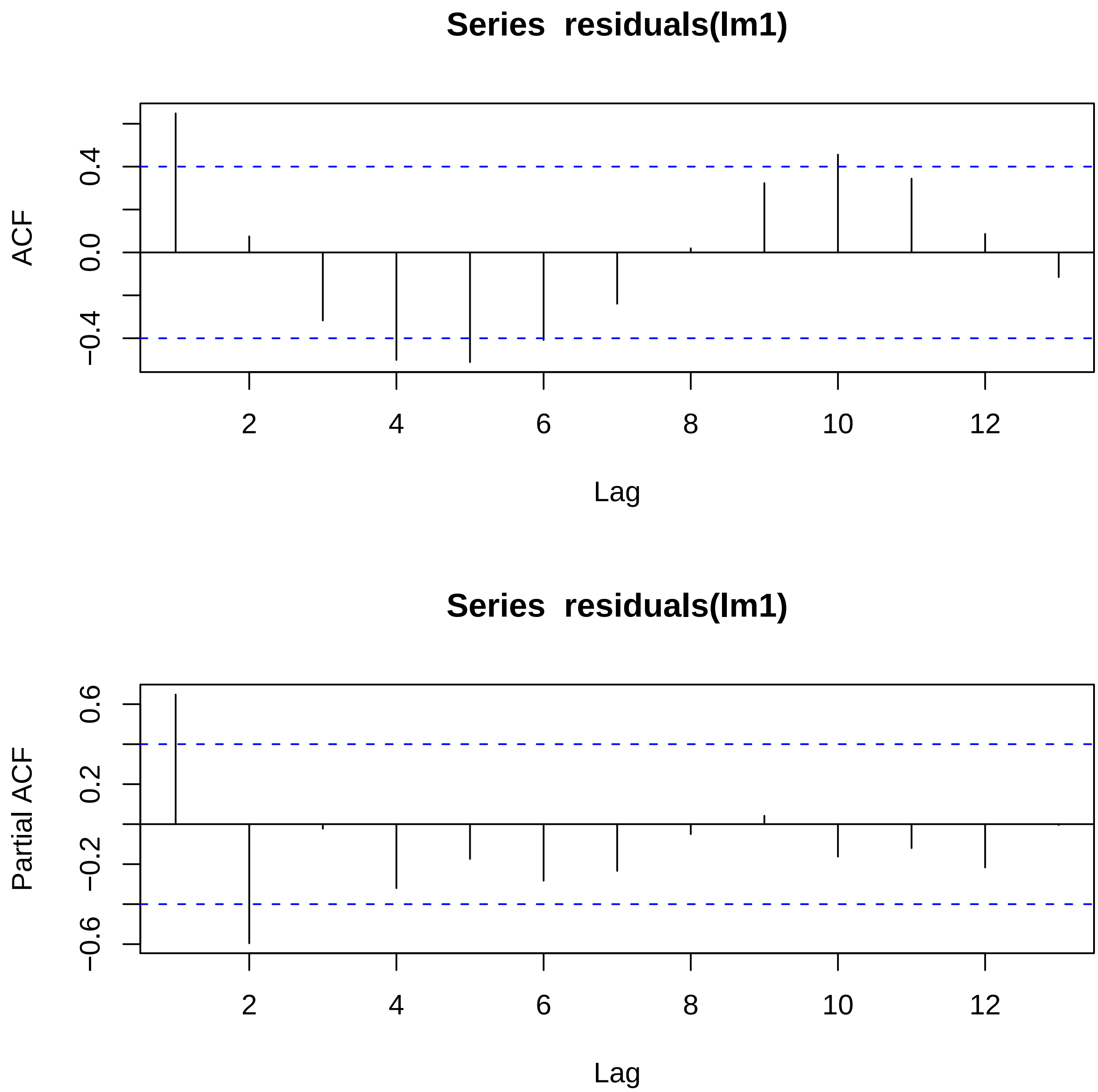


Figure 22: ACF and PACF plots of the residuals from the linear model reported in (b).

It appears that an AR(2) model is appropriate.

(d)

```
>  1=‘  ‘(   ,   =(2,0,0),  = ‘ ‘  ‘ ( ‘ =  (   )))
```

```
>  1

 :
  ((  =   ,   = (2, 0, 0),   =      ( (   =
  (   )))

   :
     1      2
     1.3310  -0.8289  -6672.500  3.4161
 .   0.1156   0.1093   1910.761  0.9602

  2      94.08:       = -90.12,    = 188.25
```

All coefficient estimates are significant. The sale increased by 3.41 units per year. The intercept term refers to the mean sale at year 0, which is not interesting. If, for example, we are interested in using 1982 as the base year, the model fit should be modified with the covariates being time(units)-1982 where now year 0 refers to 1982 and the intercept is the mean sale in 1982. Below are the modified R codes.

```
>  1=  (   ,   =(2,0,0),
  =     ( (  =  (   )-1982))
>  1

 :
  ((  =   ,   = (2, 0, 0),   =      ( (   =
  (   ) -
   1982))

   :
     1      2
     1.3310  -0.8289    98.2637  3.4161
 .   0.1154   0.1077     8.5687  0.6613

  2      94.08:       = -90.12,    = 188.25
```

Note that now all other parameter estimates are unchanged.

For the fitted AR(2) model,

```
>#
> o 1=1.331;o 2=-0.8289
> o 1 2+4*o 2
 1 -1.544039
>#
> 2*o /  (o 1/(2*    (-o 2)))
 1 8.365753
```

Thus, besides an increasing trend, the data is quasi-periodic with period around 8.4 years, which is somewhat lower than that suggested from the time plot of the data; the latter seems to suggest a period around 10.

(e)

```
>    ( 1)
            ( 1)  :

      :
  (   ) :
>#
>    ( 1,    =.2)
```

Note that in carrying out the model diagnostics, we have to specify the option tol=0.2; this option relaxes the criterion for finding the range of maximum lags in the Ljung-Box tests. The default option is tol=0.1 which uses a stringent criterion. Here, the small sample size necessitates the relaxation, lest the Ljung-Box tests cannot be carried out.

According to the above diagnostics, the fitted model provides a good fit to the data.

(f) Now, we repeat the analysis with the logarithmically transformed units.

```
>  2=  ( (   ),    =(2,0,0),
  =    ( =  (   )))
>  2

   :
   ( =  (   ),    = (2, 0, 0),    =    (  =
  (   )))

       :
```

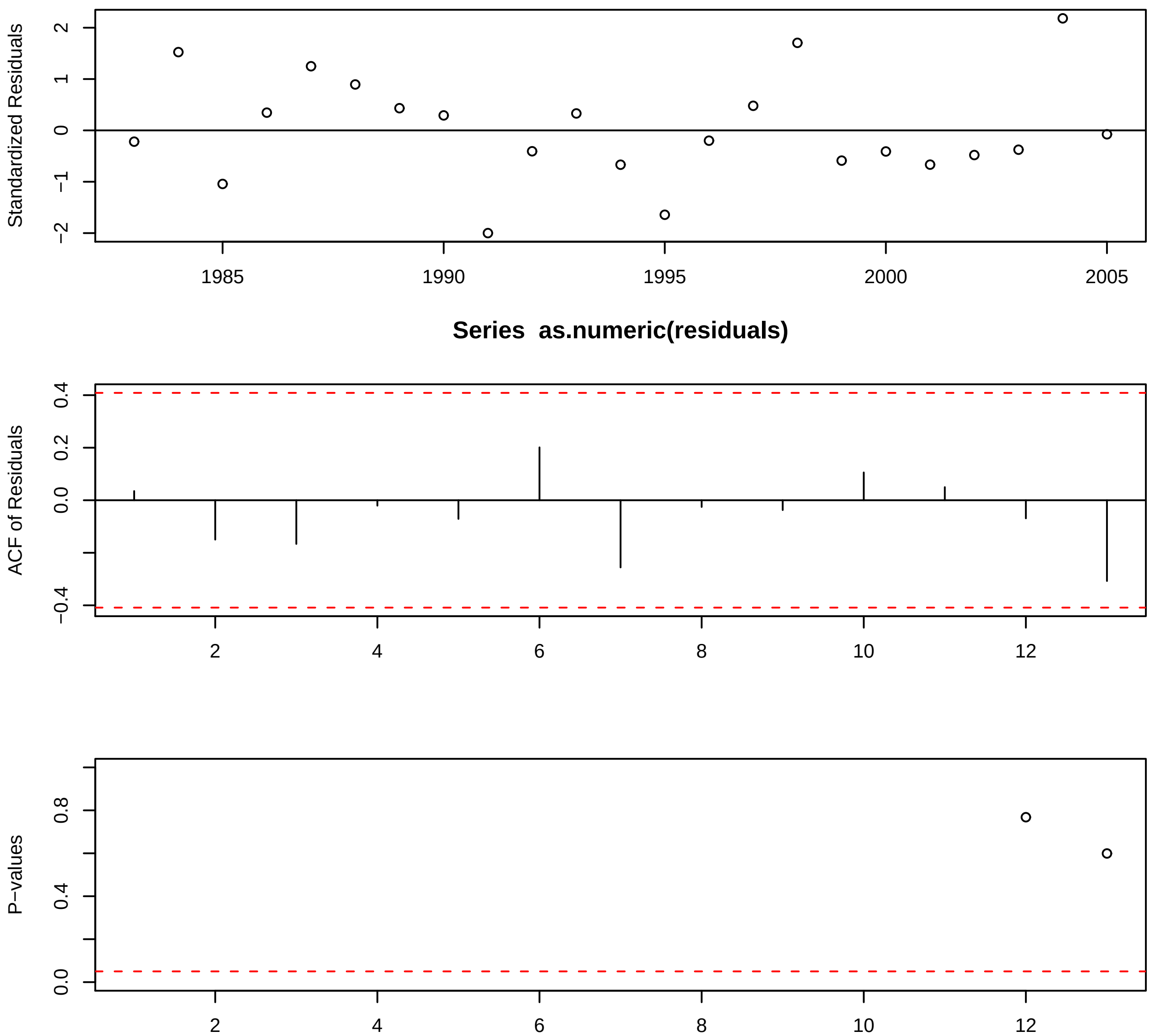


Figure 23: Model diagnostics for the fitted model reported in (d).

```
         1          2
    1.2869    -0.7750     -46.5788   0.0258
.   0.1269     0.1217

      2            0.007025:              = 24.15,      = -40.31
```

```
:
(     ($   .  ))  :        
```

Note that the standard errors of the regression parameters can't be computed, due to collinearity problem. However, if we shift the base year to 1982, the problem goes away.

```
>  2 =   (  (   ),    =(2,0,0),
 =    (   =  (   )-1982))
>  2

  :
  (  =   (   ),    = (2, 0, 0),    =     (    =
 (   ) -
  1982))

     :
       1       2
      1.2869  -0.7750       4.5918  0.0258
 . .  0.1269   0.1265       0.0760  0.0058

   2         0.007025:           = 24.15,    = -40.31
>  1=1.2869;  2=-0.7749
>  1 2+4*  2
 1 -1.443488

>#      (2)      -
> 2* /   (  1/(2*    (-  2)))
 1 8.365615
>#            8.4
>    ( 2 )
          .   ( 2 ) :

>#
>    ( 2 ,    =.2)
```

Now, the interpretation is that the annual sales increased by 2.6% annually with a quasi-period closed to 8.4 years. The model diagnostics reported below suggest that the model also provides good fit to the data on the logarithmic scale.

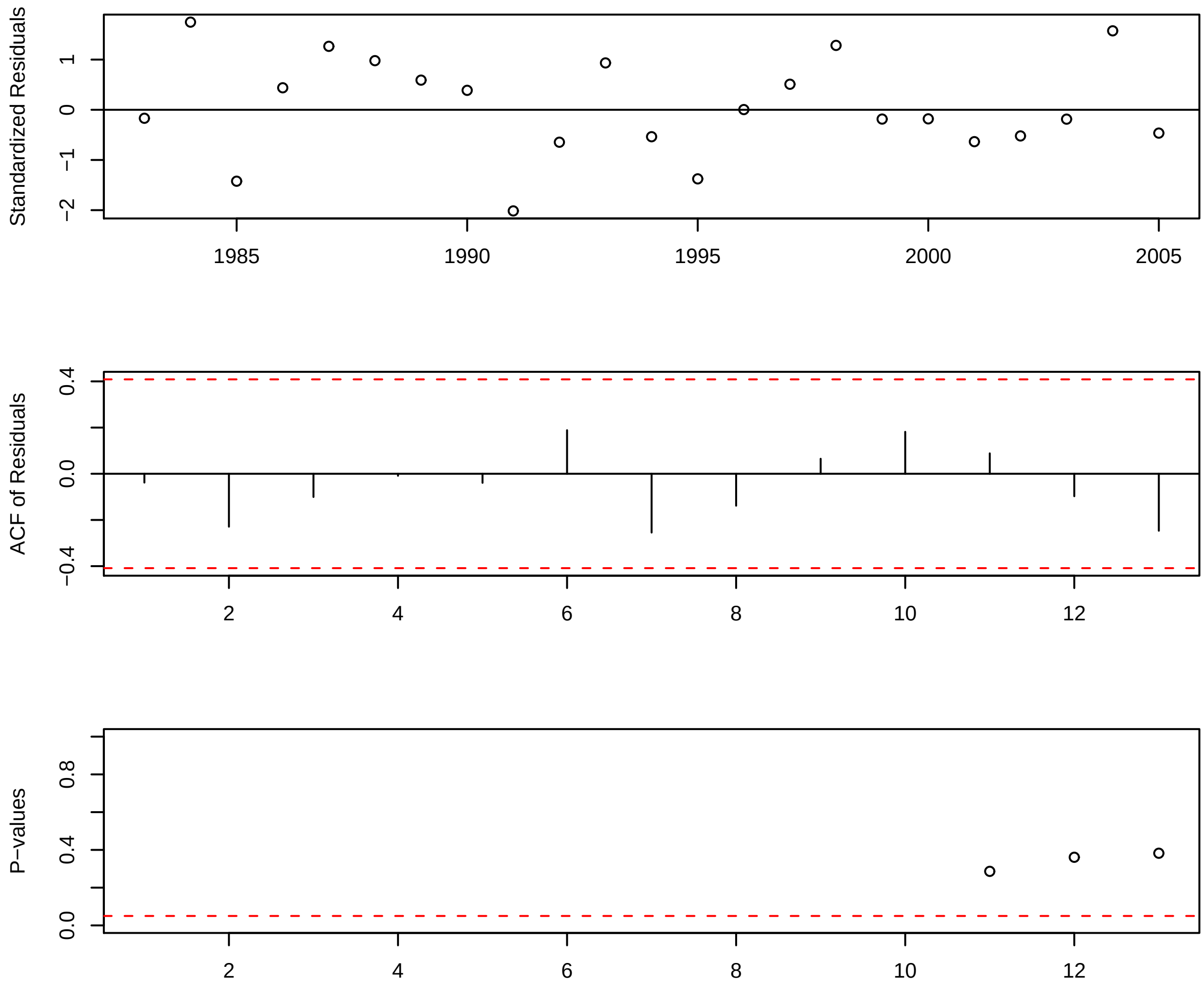


Figure 24: Model diagnostics for the fitted model reported in (d).

Exercise 11.21

In Chapter 8, an IMA$(1, 1)$ model is fitted for the logarithms of monthly oil prices. The results obtained are:

```
> ( .o )
> 1. = ( ( .o ), =(0,1,1))
> 1.
 : ( = ( .o ), = (0, 1, 1))

 :
 1
 0.2956
 . 0.0693

 2 0.006689: = 260.29, =-516.58
> o ( ( 1. ), =' ' ', =' ')
> ( =0)
```

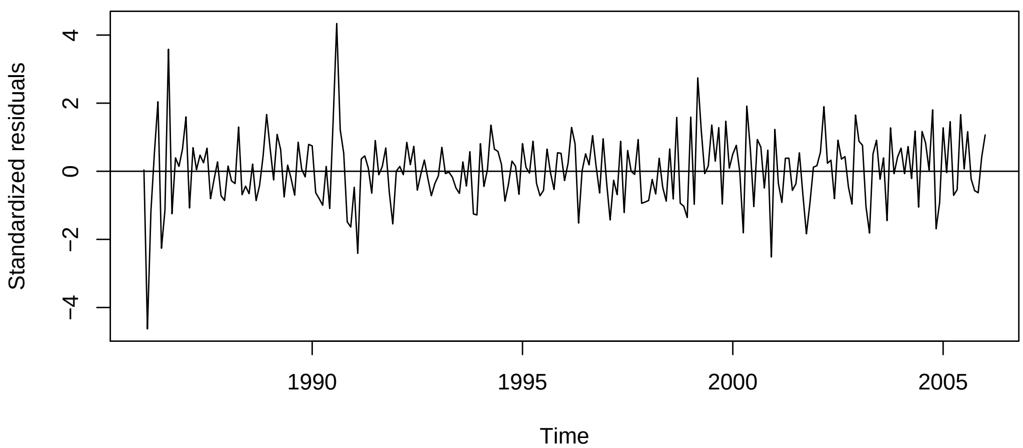


Figure 25: Residuals of model 1

The residuals graph suggests that there may be several outliers in this series. Let's check whether there are innovative or additive outliers in this time series.

```
> (1. )
       ,1]      ,2]       ,3]
      2.000000 8.000000 56.00000
  '2 -4.326085 4.007243 4.07535
> (1. )
       ,1]      ,2]       ,3]
      2.000000 8.000000 56.000000
  '1 -4.875561 3.773707 4.570056
```

The results above show that there may be outliers at Febrary 1986, August 1986, and August 1990. Since among the tests for AO and IO the largest magnitude appears as λ_1 at time $t = 1$, let's try to fix the innovative outlier at time 1 in the model.

```
> 2. =' '( ( . ), =(0,1,1), =(2))
> (2. )
      ,1]       ,2]
     8.000000 56.000000
  '2 4.070551  4.253151
> (2. )
      ,1]       ,2]
     8.000000 56.000000
  '1 3.864698  4.740229
```

In model 2 we still detect possible outliers at $t = 8$ and 56. We continue to add the innovative outlier effect at $t = 56$ into our model.

```
> 3. =' '( ( . ), =(0,1,1), =(2,56))
> (3. )
      ,1]
     8.000000
  '2 4.100764
> (3. )
      ,1]
     8.000000
  '1 3.937741
```

After fitting model 3, the outlier at $t = 8$ is still significant. Since $|\lambda_{2,8}| > |\lambda_{1,8}|$, we incorporate the additive outlier at $t = 8$ into the model 3 to get model 4.

```
> 4.  = (  ( .o  ),  =(0,1,1), =(2,56),
+  = (  =  (  ( .o  ))==8))
>   ( 4. )
1 "  "
>   ( 4. )
1 "  "
> 4.

 :  ( =  ( .o  ),   = (0, 1, 1),  =
 (  =  (  ( .o  )) == 8),  = (2, 56))

 :
      1              -2      -56
    0.2696  0.1534  -0.3966  0.3589
 .  0.0593  0.0439   0.0752  0.0732

  2     0.005276:      = 288.76, =-567.52
```

We can see that after adding 3 outliers into the original IMA(1, 1) model, model 4 has no more significant outliers. In model 4, all outlier estimates are significant. Compare the results of model 4 with model 1, in which the outliers were not taken into account. The estimate of ϕ has not changed very much, but the AIC is much better ($-567.52 < -516.58$). So the model IMA(1, 1) + 2 IO's and 1 AO is better than the original IMA(1, 1) model.

Chapter 12

Exercise 12.1

Let us plot the absolute returns and squared returns for the CREF data.

```
> library(TSA)
> data(CREF)
> r.cref=diff(log(CREF))*100
> win.graph(width=4.875, height=2.5,pointsize=8)
> plot(abs(r.cref))
```

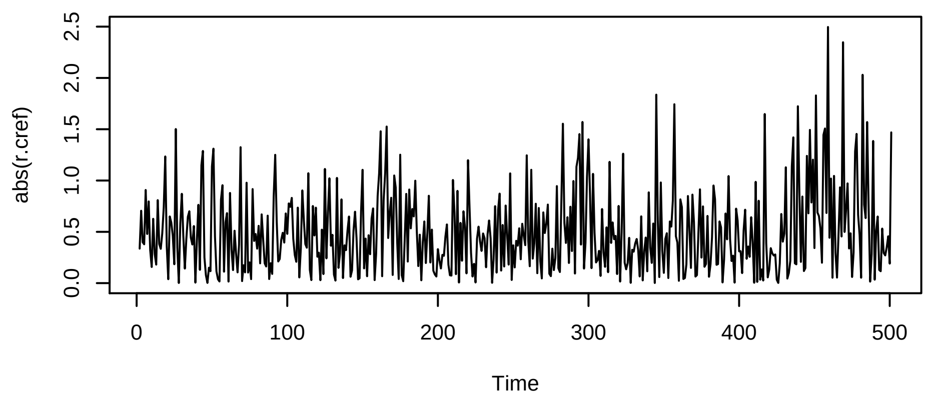


Figure 1: Absolute Values of Daily CREF Stock Returns

```
> win.graph(width=4.875, height=2.5,pointsize=8)
> plot(r.cref^2)
```

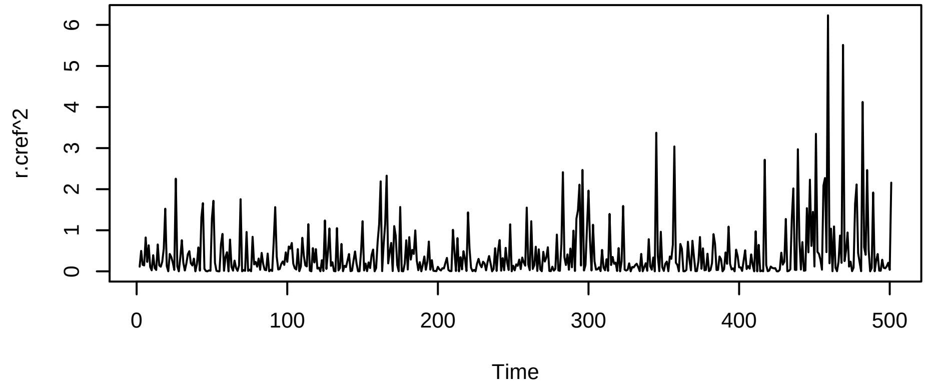


Figure 2: Squared Daily CREF Stock Returns

From the above 2 plots we can clearly see that the returns became more volatile towards the end of the study period.

Exercise 12.2

Let us plot the absolute returns and squared returns for the CREF data.

```
>     ‘  (    )
> ‘ ‘(  .  )
> o   (  (‘ (  .  $   ‘ )   =1),  =’  ’,  ‘=’‘  ’,
+    ‘=’‘ (     )’)
```

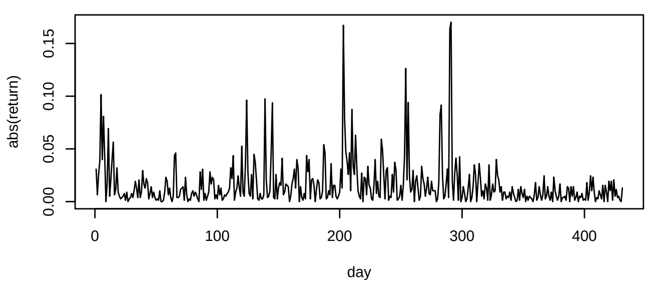


Figure 3: Abs of Daily Returns of USD/HKD Exchange Rate: 1/1/05-3/7/06

```
> 'o ( =4.875, =2.5,o =8)
> o ( ( . $ ' 2, =1), o =' ', '=' ' ',
+ '=' 2')
```

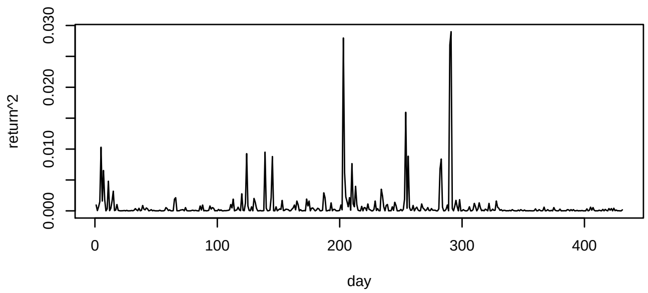


Figure 4: Squared Daily Returns of USD/HKD Exchange Rate: 1/1/05-3/7/06

From the above 2 plots we can see that the volatility clustering is evident particularly during the second half year of 2005.

Exercise 12.3

Since

$$\eta_t = r_t^2 - \sigma_{t|t-1}^2$$

and

$$r_t = \sigma_{t|t-1}\varepsilon_t$$

where $\{\varepsilon_t\}$ is a sequence of i.i.d. random variables with zero mean and unit variance and ε_t is independent of $r_{t-j}, j = 1, 2, \ldots$, we have

$$\begin{aligned}
\mathrm{E}\left(\eta_t\right) &= \mathrm{E}\left(\sigma_{t|t-1}^2\left(\varepsilon_t^2 - 1\right)\right) \\
&= \mathrm{E}\left(\mathrm{E}\left(\sigma_{t|t-1}^2\left(\varepsilon_t^2 - 1\right)\middle| r_{t-1}, r_{t-2}, \ldots\right)\right) \\
&= \mathrm{E}\left(\sigma_{t|t-1}^2\mathrm{E}\left(\left(\varepsilon_t^2 - 1\right)\middle| r_{t-1}, r_{t-2}, \ldots\right)\right) \\
&= \mathrm{E}\left(\sigma_{t|t-1}^2 \times 0\right) \\
&= 0,
\end{aligned}$$

and for every $k > 0$,

$$\begin{aligned}
\mathrm{E}\left(\eta_t\eta_{t-k}\right) &= \mathrm{E}\left(\sigma^2_{t|t-1}\left(\varepsilon_t^2-1\right)\left(r^2_{t-k}-\sigma^2_{t-k|t-k-1}\right)\right)\\
&= \mathrm{E}\left(\mathrm{E}\left(\sigma^2_{t|t-1}\left(\varepsilon_t^2-1\right)\left(r^2_{t-k}-\sigma^2_{t-k|t-k-1}\right)\middle| r_{t-1},r_{t-2},\ldots\right)\right)\\
&= \mathrm{E}\left(\sigma^2_{t|t-1}\left(r^2_{t-k}-\sigma^2_{t-k|t-k-1}\right)\mathrm{E}\left(\left(\varepsilon_t^2-1\right)\middle| r_{t-1},r_{t-2},\ldots\right)\right)\\
&= \mathrm{E}\left(\sigma^2_{t|t-1}\left(r^2_{t-k}-\sigma^2_{t-k|t-k-1}\right)\times 0\right)\\
&= 0.
\end{aligned}$$

Hence,

$$\mathrm{Cov}\left(\eta_t,\eta_{t-k}\right) = \mathrm{E}\left(\eta_t\eta_{t-k}\right) - \mathrm{E}\left(\eta_t\right)\mathrm{E}\left(\eta_{t-k}\right) = 0,$$

which means that $\{\eta_t\}$ is a serially uncorrelated sequence. Similarly, we obtain

$$\begin{aligned}
\mathrm{E}\left(\eta_t r^2_{t-k}\right) &= \mathrm{E}\left(\sigma^2_{t|t-1}\left(\varepsilon_t^2-1\right)r^2_{t-k}\right)\\
&= \mathrm{E}\left(\mathrm{E}\left(\sigma^2_{t|t-1}\left(\varepsilon_t^2-1\right)r^2_{t-k}\middle| r_{t-1},r_{t-2},\ldots\right)\right)\\
&= \mathrm{E}\left(\sigma^2_{t|t-1}r^2_{t-k}\mathrm{E}\left(\left(\varepsilon_t^2-1\right)\middle| r_{t-1},r_{t-2},\ldots\right)\right)\\
&= \mathrm{E}\left(\sigma^2_{t|t-1}r^2_{t-k}\times 0\right)\\
&= 0.
\end{aligned}$$

So

$$\mathrm{Cov}\left(\eta_t,r^2_{t-k}\right) = \mathrm{E}\left(\eta_t r^2_{t-k}\right) - \mathrm{E}\left(\eta_t\right)\mathrm{E}\left(r^2_{t-k}\right) = 0,$$

which means that η_t is uncorrelated with past squared returns.

Exercise 12.4

Substituting $\sigma^2_{t|t-1} = r_t^2 - \eta_t$ into Equation (12.2.2), we have

$$\begin{aligned}
& r_t^2 - \eta_t = \omega + \alpha r^2_{t-1}\\
\Longrightarrow\ & r_t^2 = \omega + \alpha r^2_{t-1} + \eta_t,
\end{aligned}$$

which is Equation (12.2.5).

Exercise 12.5

Equation (12.2.2) tells us

$$\begin{aligned}
& \sigma^2_{t|t-1} = \omega + \alpha r^2_{t-1}\\
\Longrightarrow\ & \sigma^4_{t|t-1} = \omega^2 + 2\omega\alpha r^2_{t-1} + \alpha^2 r^4_{t-1}.
\end{aligned}$$

Let us take expectation on both sides of the last equation. Denoting $\tau = \mathrm{E}\left(\sigma_{t|t-1}^4\right)$ and since $\mathrm{E}\left(r_t^4\right) = 3\tau$ (Equation (12.2.7)), we get

$$\tau = \omega^2 + 2\omega\alpha\sigma^2 + \alpha^2 3\tau,$$

which is Equation 12.2.8.

Exercise 12.6

The order is like: the uniform distribution on $[-1, 1]$, the normal distribution with mean 0 and variance 4, the t-distribution with 30 d.f., and the t-distribution with 10 d.f..

Reason: Kurtosis is a measure of the "peakedness". A high kurtosis distribution has a sharper "peak" and fatter "tails", while a low kurtosis distribution has a more rounded peak with wider "shoulders". So the uniform distribution which has no peak must have the lowest kurtosis value. Among the other three distributions, the t-distributions have heavier tails than the normal distribution. Furthermore, the t-distribution with 10 d.f. has heavier tails compared to the t-distribution with 30 d.f..

Exercise 12.7

Simulate a time series, of size 500, from a GARCH(1,1) model with standard normal innovations and parameter values $\omega = 0.01, \alpha = 0.1$, and $\beta = 0.8$.

```
> (  )
> . (1234567)
> ' o 1. = ' o . (' o '=o(0.01,0.1), '=0.8, =500)
> o  ( ' o 1. , o=' ', ' = o  ( ), ' = ' ')
```

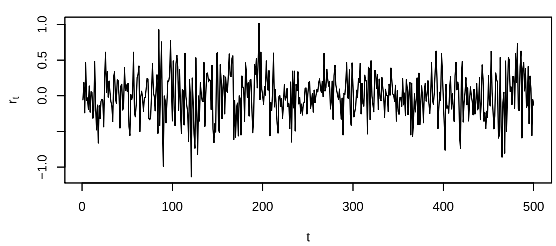


Figure 5: Simulated GARCH(1,1) Process

Let us see the sample ACF, PACF, and EACF of the simulated time series.

```
> acf(garch1.sim)
```

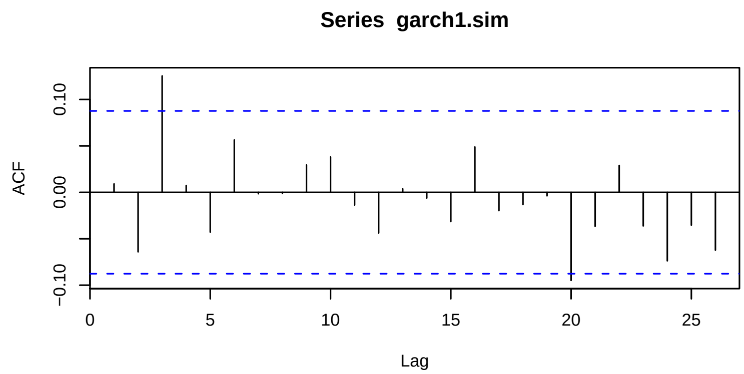


Figure 6: Sample ACF of Simulated GARCH(1,1) Process

```
> pacf(garch1.sim)
```

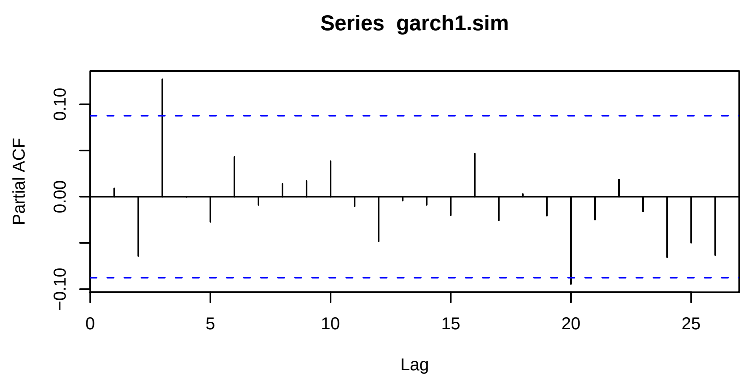


Figure 7: Sample PACF of Simulated GARCH(1,1) Process

```
> eacf(garch1.sim)
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0     x
1 x   x
2 x x x
3   x x
4     x
5 x x x
6 x x x
7 x x   x     x
```

Except for lag 3 and 20 which are mildly significant, the sample ACF and PACF of the simulated data do not show significant correlations. Also, the pattern in the EACF table seems suggest an AR(3) model. Hence, the simulated process seems consistent with the assumption of white noise.

(a) Let us see the sample ACF, PACF, and EACF of the squared simulated GARCH(1,1) time series.

```
> acf(garch1.sim^2)
```

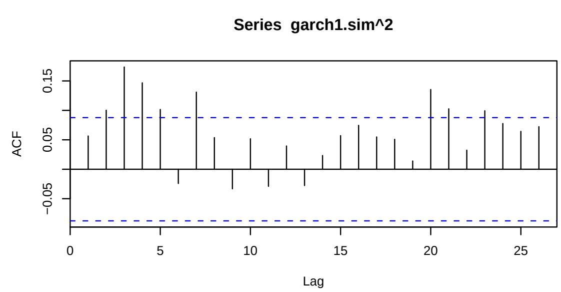


Figure 8: Sample ACF of Squared Simulated GARCH(1,1) Process

```
> pacf(garch1.sim^2)
```

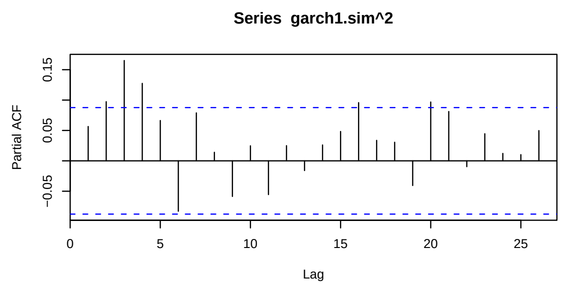


Figure 9: Sample PACF of Squared Simulated GARCH(1,1) Process

```
> eacf(garch1.sim^2)
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0   x x x x   x
1 x       x   x x
2 x         x x   x
3 x x x     x
4 x x x x x
5 x   x
6 x x x x   x
7 x x   x   x
```

The sample ACF and PACF of the squared simulated GARCH(1,1) process show significant autocorrelation pattern in the squared data. Hence the simulated process is serially dependent as it is. But the pattern in the EACF table is not very clear. An ARMA(3,3) model is kind of suggested. As mentioned in the textbook, the fuzziness of the signal in the EACF table is likely caused by the larger sampling variability when we deal with higher moments.

(b) Let us see the sample ACF, PACF, and EACF of the absolute values of the simulated GARCH(1,1) time series.

```
> acf(abs(garch1.sim))
```

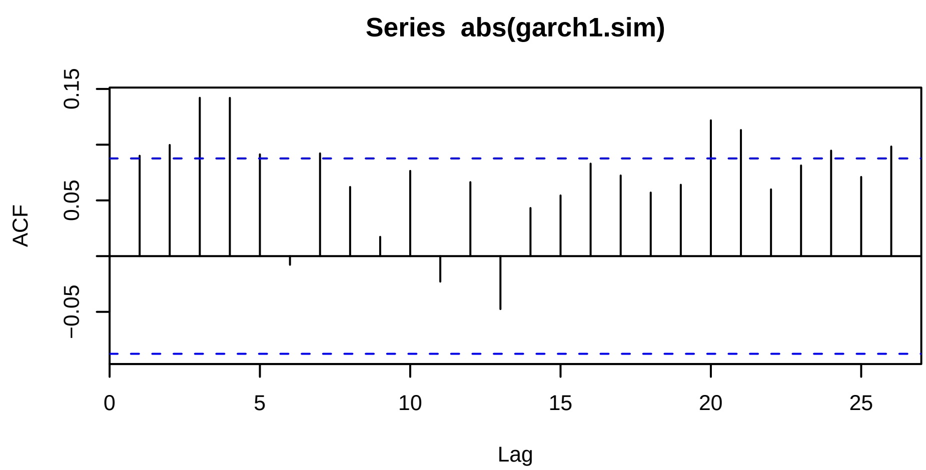


Figure 10: Sample ACF of Abs(Simulated GARCH(1,1) Process)

```
> pacf(abs(garch1.sim))
```

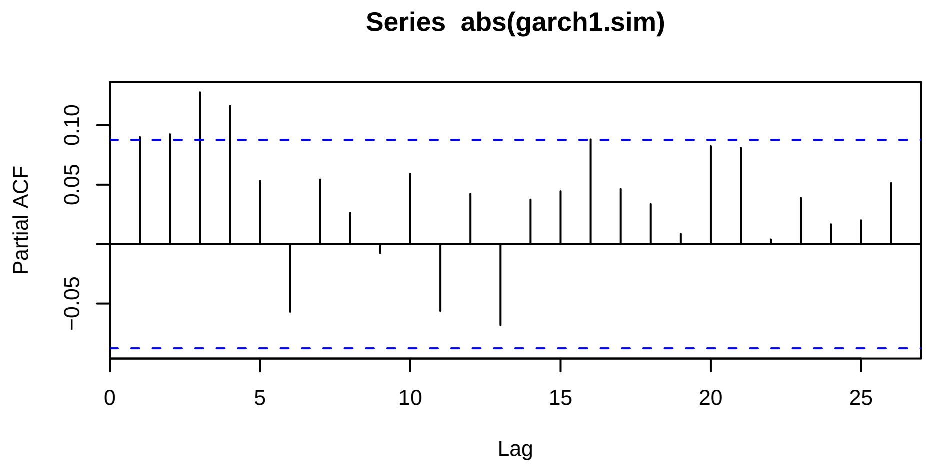


Figure 11: Sample PACF of Abs(Simulated GARCH(1,1) Process)

```
> eacf(abs(garch1.sim))
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 x x x x x   x
1 x       x
2 x x
3 x x       x
4 x x   x   x
5 x   x x x
6 x   x   x x
7 x x x x x   x
```

The sample EACF table for the absolute simulated process suggests convincingly an ARMA(1,1) model, and therefore a GARCH(1,1) model for the original data.

(c)

The McLeod-Li test shows the presence of strong ARCH effects in the data, as we know there are.

(d) Let us see the sample ACF, PACF, and EACF of the squared GARCH(1,1) time series using only the first 200 simulated data.

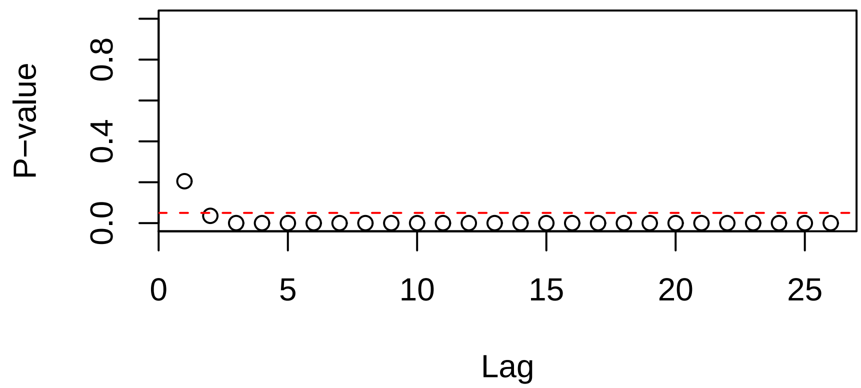


Figure 12: McLeod-Li Test for the Simulated GARCH(1,1) Process

```
> ‘ ɒ 2= ‘ ɒ 1.    1:200
> ‘ɒ ( ‘ ɒ 2 2)
```

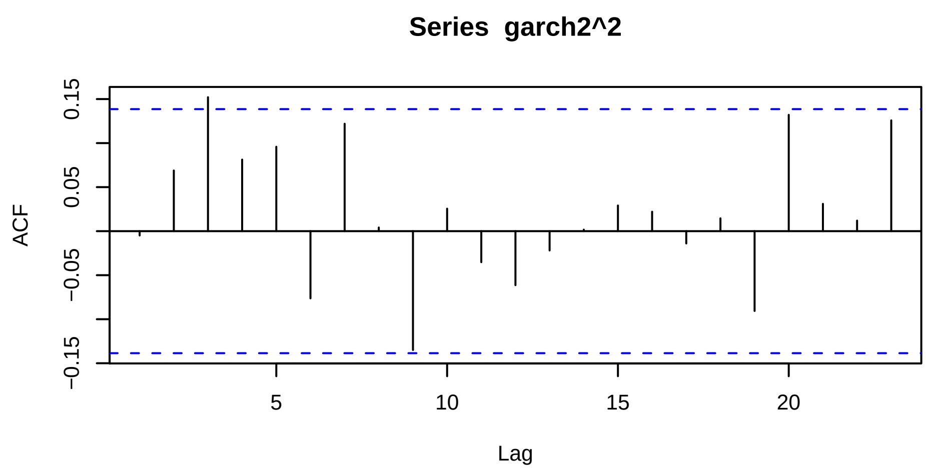


Figure 13: Sample ACF of 200 Squared Simulated GARCH(1,1) Process

```
> pacf(garch2^2)
```

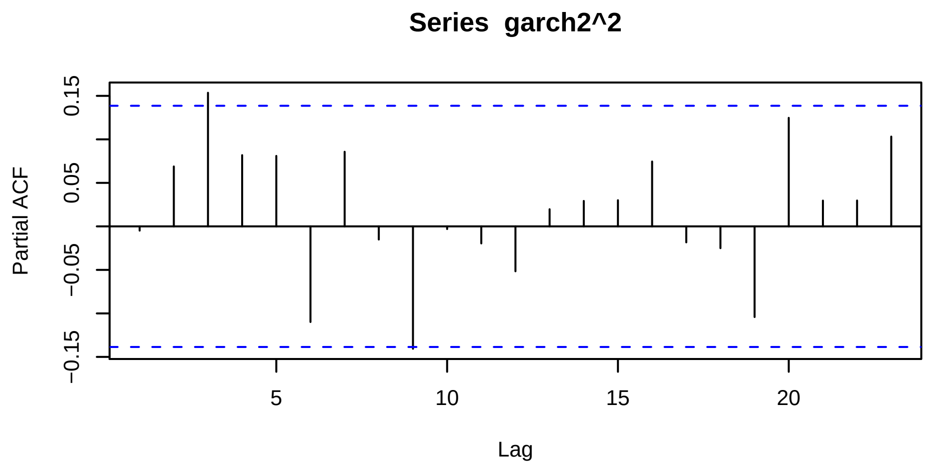


Figure 14: Sample PACF of 200 Squared Simulated GARCH(1,1) Process

```
> eacf(garch2^2)
AR/MA
```

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0			x											
1														
2	x	x							x					
3	x	x	x											
4	x		x											
5	x	x			x									
6	x	x			x	x								
7	x			x		x								

The plots of the ACF and PACF for the first 200 squared simulated data show no significant autocorrelations except for lags 3 and 9. Also, the EACF table seems to suggest the 200 squared data are white noise.

Then let us see the sample ACF, PACF, and EACF of the absolute values of the GARCH(1,1) time series also using the first 200 simulated data.

```
> acf(abs(garch2))

> pacf(abs(garch2))
```

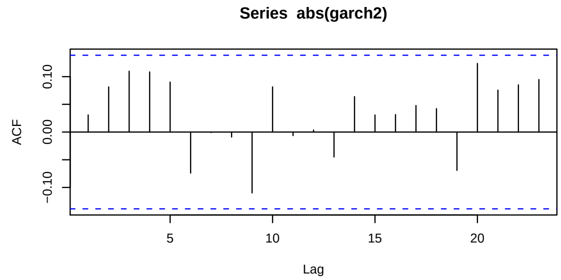


Figure 15: Sample ACF of 200 Abs(Simulated GARCH(1,1) Process)

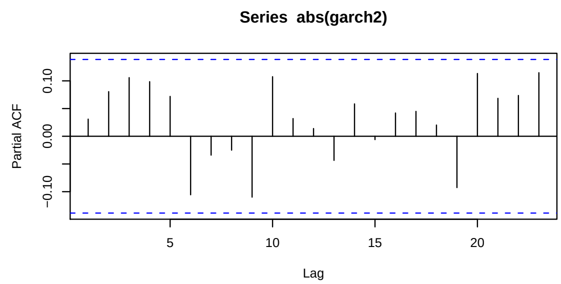


Figure 16: Sample PACF of 200 Abs(Simulated GARCH(1,1) Process)

```
> [illegible] 2))
[illegible]
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0
1 [illegible]
2 [illegible] [illegible]
3 [illegible]
4 [illegible]
5 [illegible] [illegible] [illegible] [illegible] [illegible]
6 [illegible]   [illegible] [illegible] [illegible]
7 [illegible]   [illegible] [illegible]
```

The plots of the ACF and PACF for the first 200 absolute simulated data show no significant autocorrelations. In addition, the EACF table convincingly suggests that the 200 squared data are white noise.

So a GARCH(p, q) time series with size 200 is not enough for us to identify the orders p and q by inspecting its ACF, PACF and EACF.

Exercise 12.8

(a) The plot of the daily CREF bond prices is given below. Generally the bond price has an increasing trend. But deep jumps appear several times during the whole period.

```
> (        )
> o  (        )
```

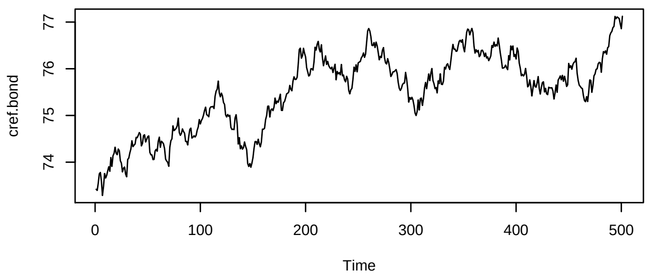


Figure 17: Daily CREF Bond Prices: August 26, 2004 to August 15, 2006

(b) Let us plot the daily bond returns below. It shows no evidence of volatility clustering.

```
>        = (  (        ))*100
> o  (     )
>      ( =0)
```

(c)

The McLeod-Li test suggests that there is no ARCH effects in the return series.

(d) The ACF and PACF plots of the bond returns suggest that the returns have little serial correlation.

```
>  (     )
> o  (     )
```

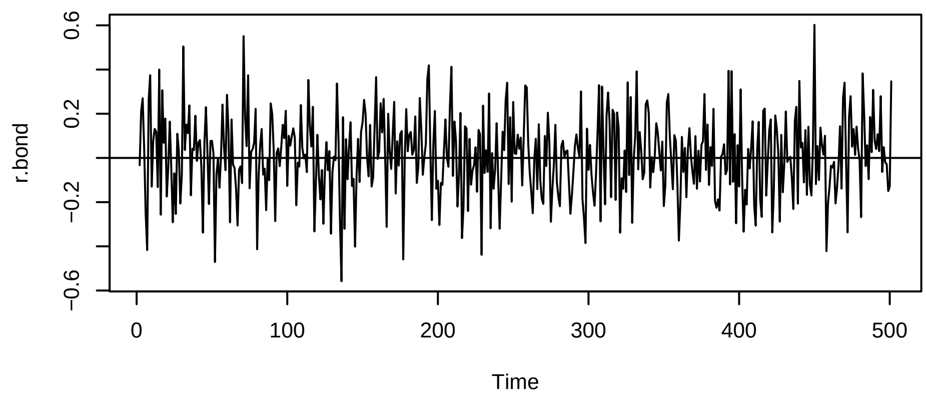


Figure 18: Daily CREF Bond Returns: August 26, 2004 to August 15, 2006

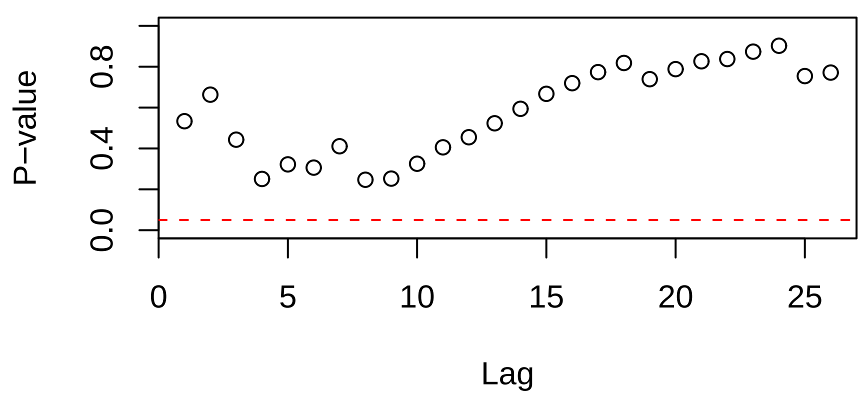


Figure 19: McLeod-Li Test for the returns.

Also, the ACF and PACF plots for the absolute and squared returns are given below. From these plots, no significant autocorrelation is observed, which further supports the claim that the returns of the CREF bond price series appear to be independently and identically distributed.

```
> [illegible]([illegible]([illegible]))
> [illegible]([illegible]([illegible]))
```

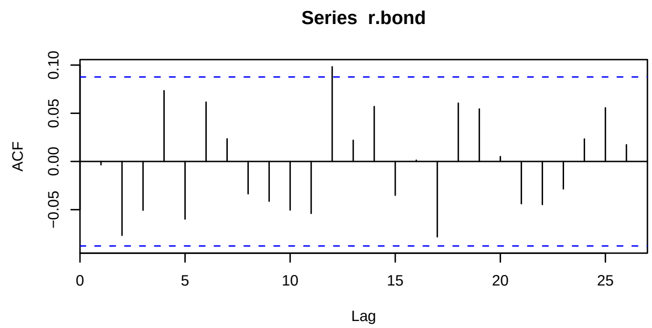


Figure 20: Sample ACF of Daily CREF Bond Returns

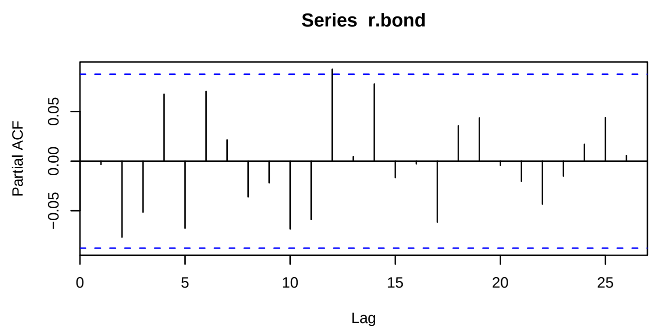


Figure 21: Sample PACF of Daily CREF Bond Returns

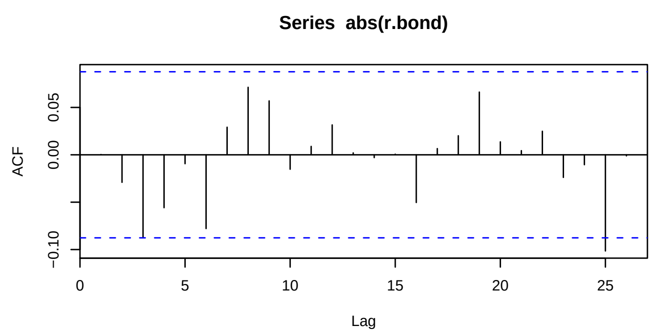


Figure 22: Sample ACF of the Absolute Daily CREF Bond Returns

```
> [illegible]( [illegible] 2)
> [illegible]( [illegible] 2)
```

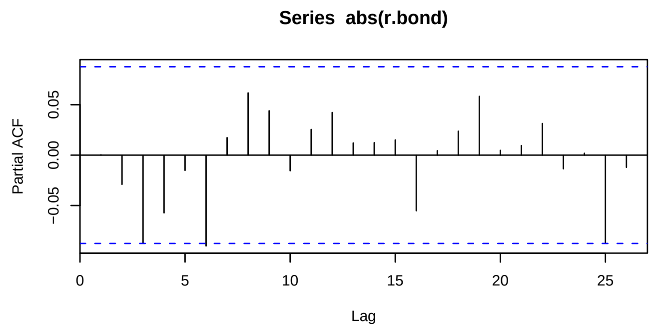


Figure 23: Sample PACF of the Absolute Daily CREF Bond Returns

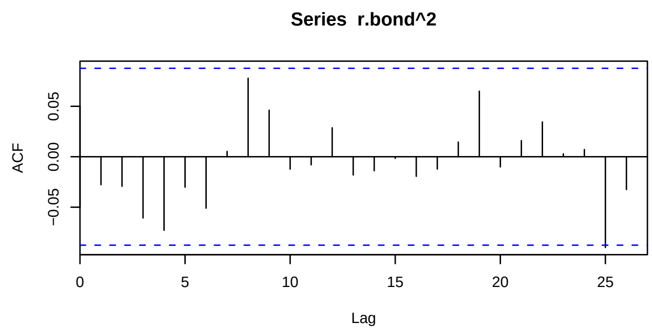


Figure 24: Sample ACF of the Squared Daily CREF Bond Returns

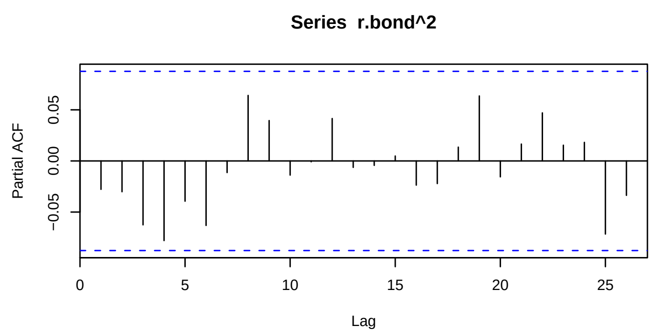


Figure 25: Sample PACF of the Squared Daily CREF Bond Returns

Exercise 12.9

(a) The daily returns of the google stock are plotted below.

```
> c[illegible]‘ ‘([illegible])
> o[illegible]([illegible])
```

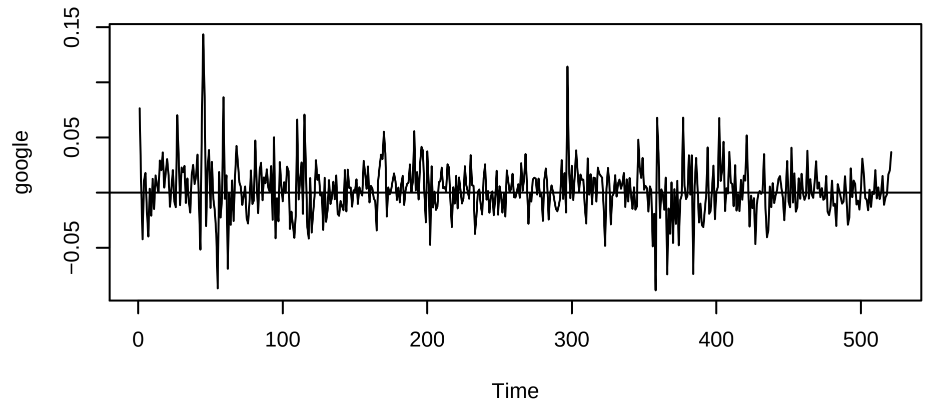


Figure 26: Daily Google Stock Returns: August 14, 2004 to September 13, 2006

From the ACF and PACF of the daily returns, we see that the data are essentially uncorrelated over time.

```
> acf(google)
```

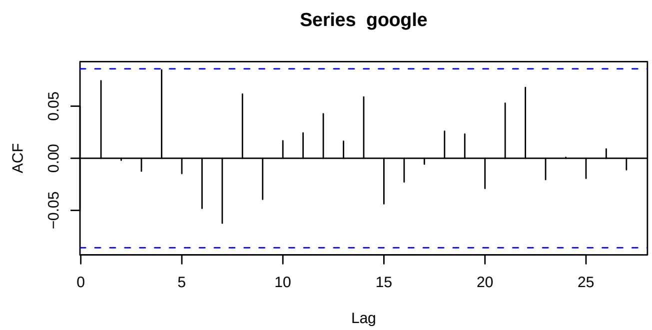


Figure 27: Sample ACF of the Daily Google Stock Returns

```
> o‘ (  )
```

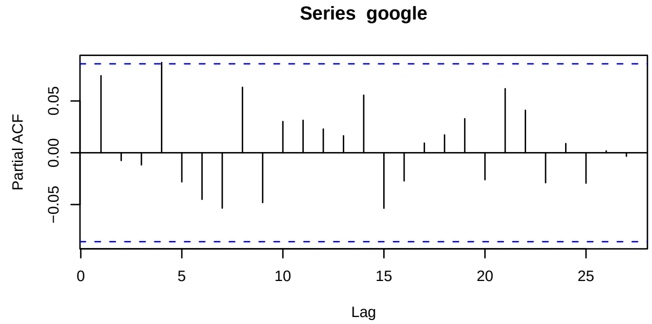


Figure 28: Sample PACF of the Daily Google Stock Returns

(b) The mean of the google daily returns is 0.00269, which is significantly greater than 0 by a one-sample, one-sided t-test. Hence, we should consider a mean+GARCH model for the datai, i.e. $r_t = \mu + \sigma_{t|t-1}\epsilon_t$. Since, the GARCH model fit is invariant to mean shift, the GARCH model fit reported below is the same whether or not we specifically include the mean shift. For convenience, the GARCH models will be fitted to the original returns.

```
>  . (  , ‘  ‘ = ‘ ’)

  ‘ o –

‘ ‘:     = 2.5689,   = 520, o– ‘  = 0.00524 ‘  ‘
 o  :   ‘  ‘  ‘ 0 95
 ‘ :
 0.000962967
 ‘ o   ‘ :
 ‘
0.002685589
```

(c)

The McLeod-Li test suggests significant ARCH effects in the Google return series.

(d) The sample EACF's of the absolute and squared daily google stock returns are given below. Both of them convincingly suggest an ARMA(1,1) model, and therefore a GARCH(1,1) model for the original data.

```
> ‘ ((  ))
```

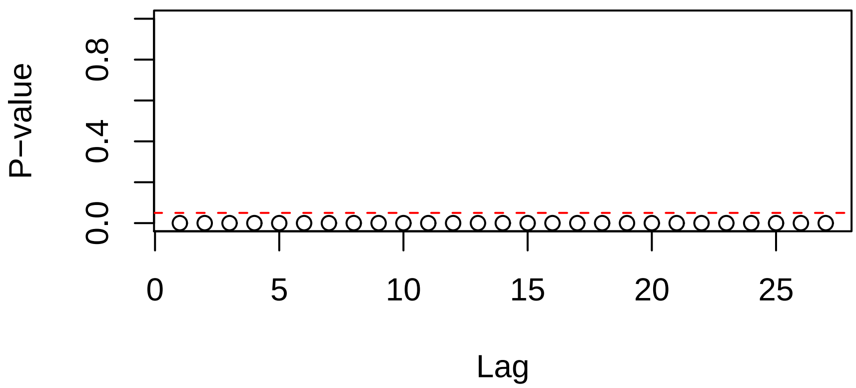


Figure 29: McLeod-Li Test for the Google returns.

[illegible]

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	[illegible]	[illegible]	[illegible]				[illegible]			[illegible]			[illegible]	[illegible]
1	[illegible]													[illegible]
2	[illegible]	[illegible]												[illegible]
3	[illegible]	[illegible]	[illegible]											[illegible]
4	[illegible]		[illegible]											
5	[illegible]		[illegible]		[illegible]									
6		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]								
7	[illegible]		[illegible]	[illegible]	[illegible]		[illegible]							

> [illegible] 2)

[illegible]

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	[illegible]	[illegible]								[illegible]				[illegible]
1	[illegible]									[illegible]				[illegible]
2	[illegible]									[illegible]				[illegible]
3	[illegible]	[illegible]	[illegible]							[illegible]				[illegible]
4	[illegible]	[illegible]	[illegible]											
5	[illegible]	[illegible]	[illegible]											
6	[illegible]	[illegible]	[illegible]	[illegible]										
7		[illegible]	[illegible]			[illegible]								

We fit a GARCH(1,1) model to the (mean-deleted) google stock daily returns data and get the MLE's of the fitted model.

```
>    1=     (      -   (        ),      =(1,1) ,          =0.000001)
>        ( 1)

    :
       (   =      ,       = (1, 1),           = 1 -06)

    :
    (1,1)

      :
                 1                 3
-3.64587 -0.46484  0.08232  0.65376  5.73913

         ( ):
                                    (>  )
 0 5.058 -05   1.232 -05    4.105 4.04 -05 ***
 1 1.264 -01   2.136 -02    5.920 3.22 -09 ***
 1 7.865 -01   3.579 -02   21.978  < 2 -16 ***
---
           :  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

The graph below shows the standardized residuals from the fitted GARCH model. It suggests no particular tendency in the standardized residuals.

```
>      (          ( 1),    =' ',    ='                  ')
```

We can also look at the sample ACF and generalized portmanteau tests of squared and absolute standardized residuals in the following plots. It seems that the standardized residuals $\{\widehat{\varepsilon}_t\}$ is close to independently and identically distributed. So the GARCH(1,1) model provides good fit to the daily google stock returns data.

```
>     (           ( 1) 2,   .      =  .    )
>         ( 1,        ='        ')

>     (   (          ( 1)),   .      =  .    )
>         ( 1,        ='          ')
```

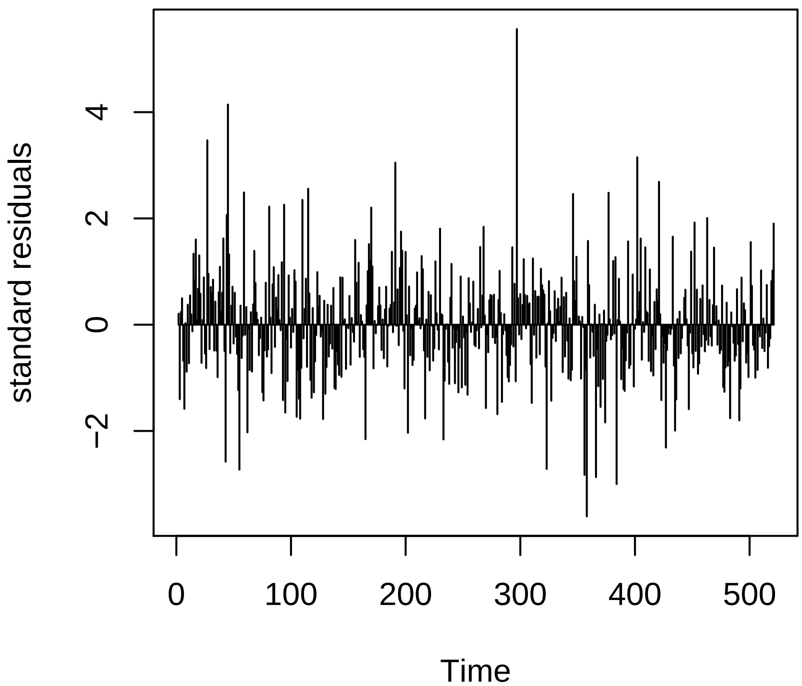


Figure 30: Standardized Residuals from the Fitted GARCH Model of Daily Google Stock Returns

(e) The next graph shows the within-sample estimates of the conditional variances. At the final time point, the squared return equals 0.00135 and the conditional variance estimated to be 0.0003417. These values combined with Equations (12.3.8) and (12.3.9) in the book can be used to compute the forecasts of future conditional variances. For example, the one-step ahead forecast of the conditional variance equals $0.00005 + 0.1264 \times 0.00135 + 0.7865 \times 0.0003417 = 0.00049$. The longer forecasts eventually approach 0.00058, the long run variance of the model.

```
> o   (t      ( 1) ,1 ) 2,  o =' ', '='o      '   '  ' o ', '=' ')
```

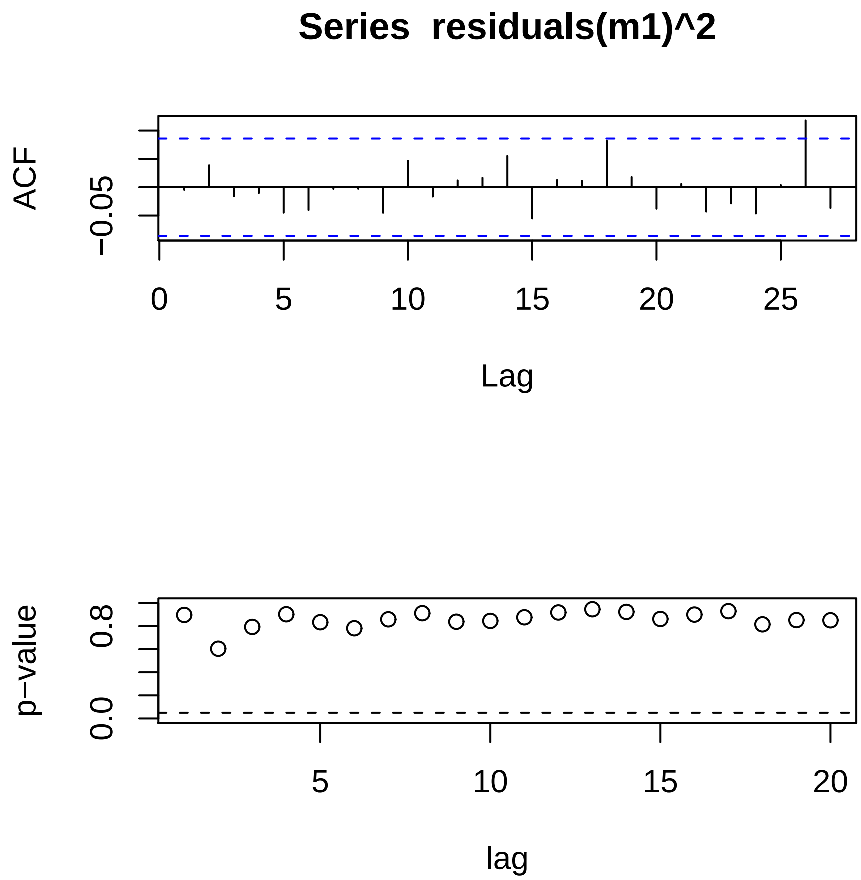


Figure 31: Sample ACF and Generalized Portmanteau Test P-Values of Squared Standardized Residuals from the Fitted Model

(f) The QQ normal plot of the standardized residuals from the fitted GARCH(1,1) model is given below. It shows that the standardized residuals are distributed with heavier tails than that of standard normal distribution on both sides.

```
>      ( ‘  ( 1));       ( ‘  ( 1))
```

(g) The 95% confidence interval for b_1 is $(0.7865 - 1.96 \times 0.03578, 0.7865 + 1.96 \times 0.03578) = (0.7164, 0.8566)$.

(h) According to the GARCH(1,1) model, the stationary variance is

$$\frac{\widehat{\omega}}{1-\widehat{\alpha}-\widehat{\beta}} = \frac{0.00005}{1-0.1264-0.7865} = 0.00058,$$

which is very close to the variance of the raw data, 0.00057. The stationary mean of the mean plus GARCH(1,1) model is simply 0.002686, the mean of the raw returns. (Remember that the stationary mean of a GARCH model is always zero!)

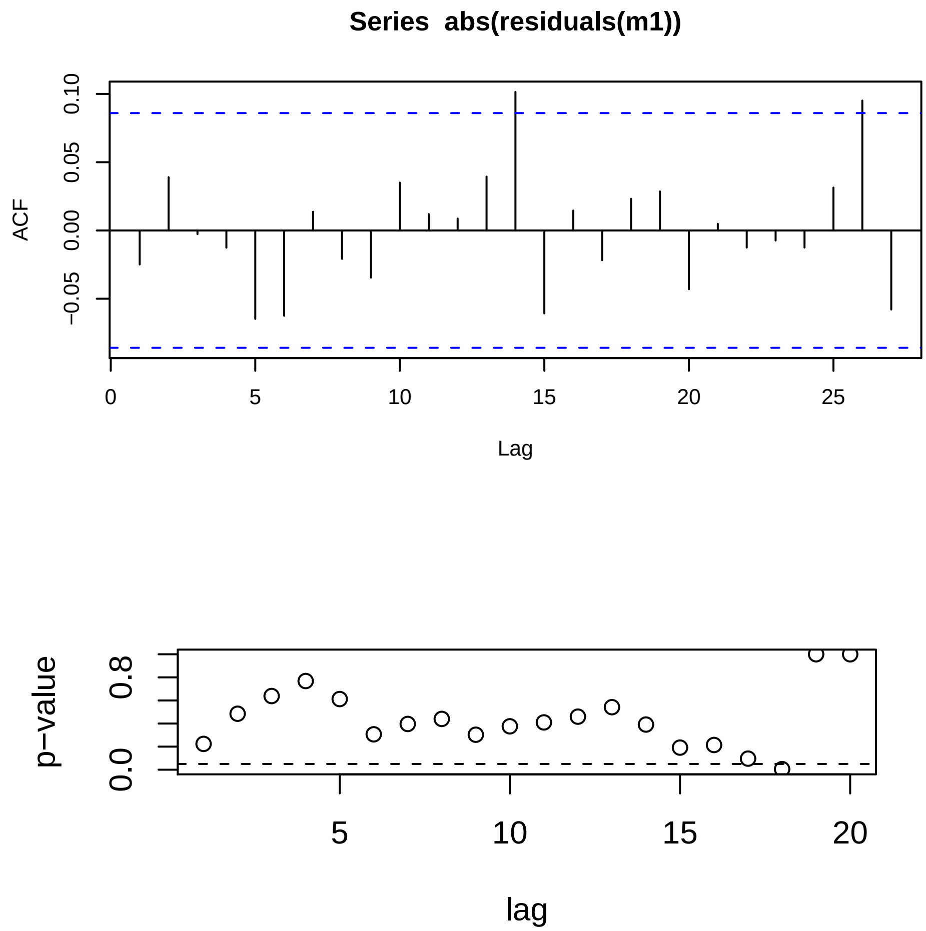


Figure 32: Sample ACF and Generalized Portmanteau Test P-Values of Absolute Standardized Residuals from the Fitted Model

(i)

There is no closed-form solution to this problem. We have to resort to Monte Carlo methods for obtaining the predictive intervals. The first to fifth steps ahead predictive distribution can be simulated by recursively computing (12.3.12) and drawing realizations from (12.2.1), with the initial condition set by the fact that at the final time point, the squared return equals 0.00135 and the conditional variance estimated to be 0.0003417. Below is the required R-code

```
 ‘  . = ‘ (     )
  o =1000
       =
  o=5
```

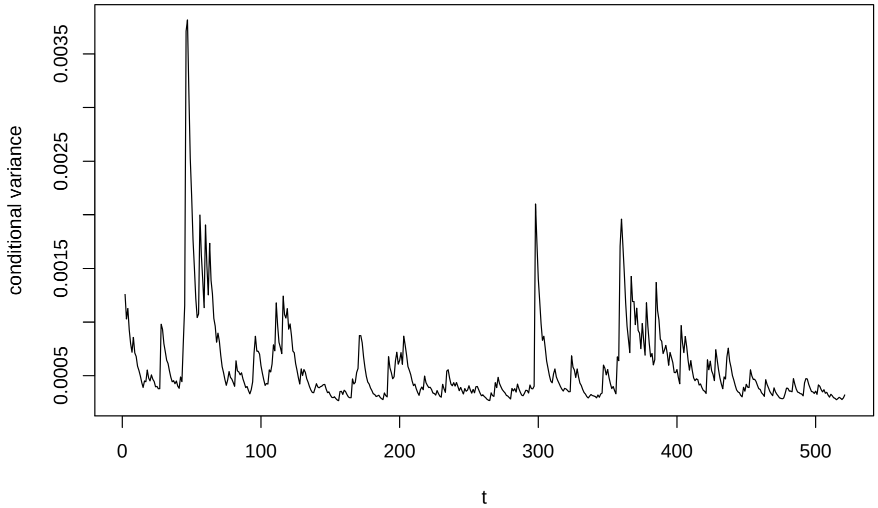


Figure 33: Estimated Conditional Variances of the Daily Google Stock Returns

```
 . (12579)
#     o      o  o       o
 (   1:  o )
   =
#     o        1   5  o     o  o
  .
 (   1:  o)
# o o      2    -1   o
 '2=0.00005246+0.7698*   '2.  1+0.1397*  .     .  1
#                 -     o
 .     =    (1,   =0,  =    '2 .5)
      =o(     ,  .     )
 '2.  1=    '2
 .     .  1=  .     2

      =o-  (       ,      )


#
      =     +   .
```

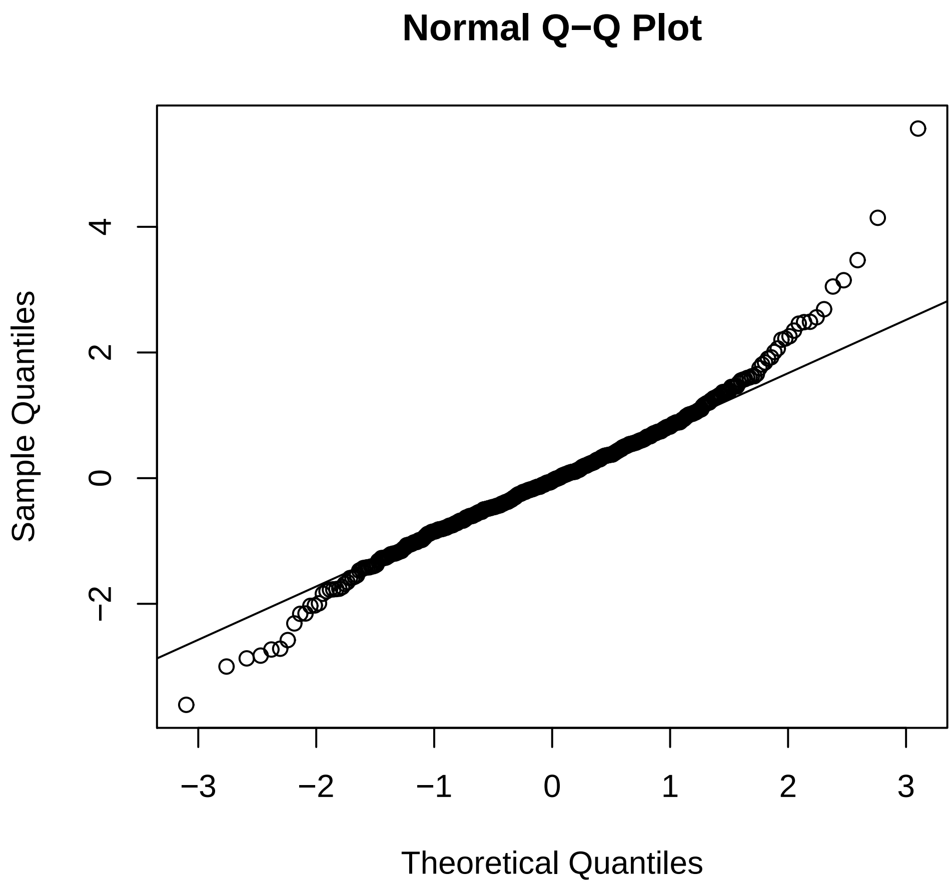


Figure 34: QQ Norm Plot of Standardized Residuals from the Fitted Model of the Daily Google Stock Returns

```
# 95% 1 5

> ,1 (.025,.975)) )
  ,1 ,2 ,3 ,4 ,5
2.5%  -0.04173014 -0.04238438 -0.0409627 -0.04163405 -0.04481970
97.5%  0.04867032  0.04674474  0.0516759  0.04919631  0.05039805
```

Exercise 12.10

(a) In Chapter 8, an IMA(1, 1) model is fitted for the logarithms of monthly oil prices. The plots of sample ACF, PACF, and EACF of the absolute and squared residuals from the fitted

IMA(1, 1) model are given below.

```
> ( .o )
> 1. = ( ( .o ), =(0,1,1))
> ( ( ( 1. )))
```

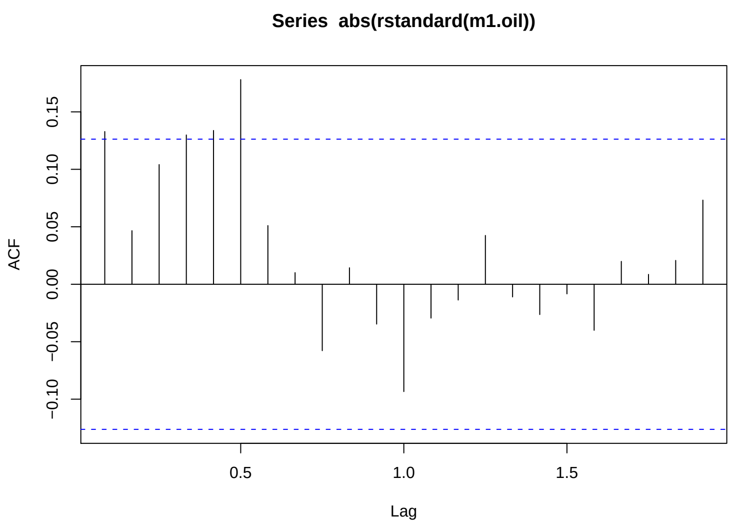


Figure 35: ACF of absolute residuals of m1.oil

```
> o‘ (‘ (    ‘ ‘ ( 1.  )))
```

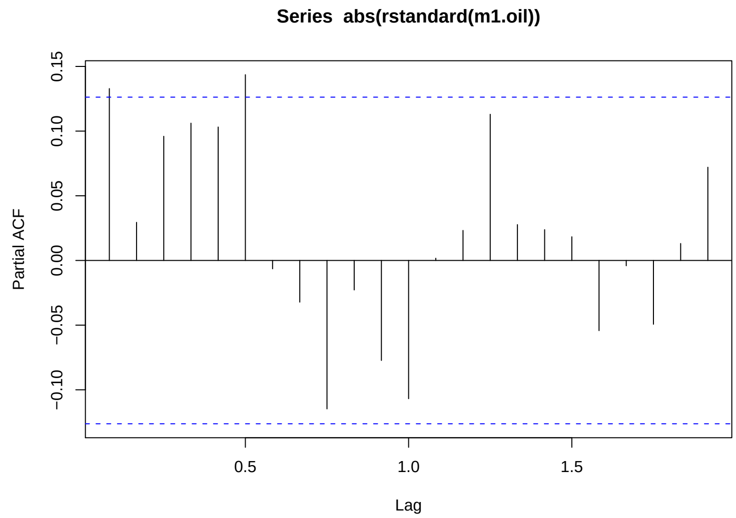


Figure 36: PACF of absolute residuals of m1.oil

```
> [illegible]([illegible]( [illegible]( 1.[illegible] )))
[illegible]
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 [illegible]       [illegible] [illegible]
1 [illegible]           [illegible]
2 [illegible] [illegible]
3 [illegible] [illegible]
4 [illegible] [illegible]   [illegible]
5 [illegible] [illegible] [illegible] [illegible]
6
7       [illegible]
> [illegible]( [illegible]( 1.[illegible] ) 2)
```

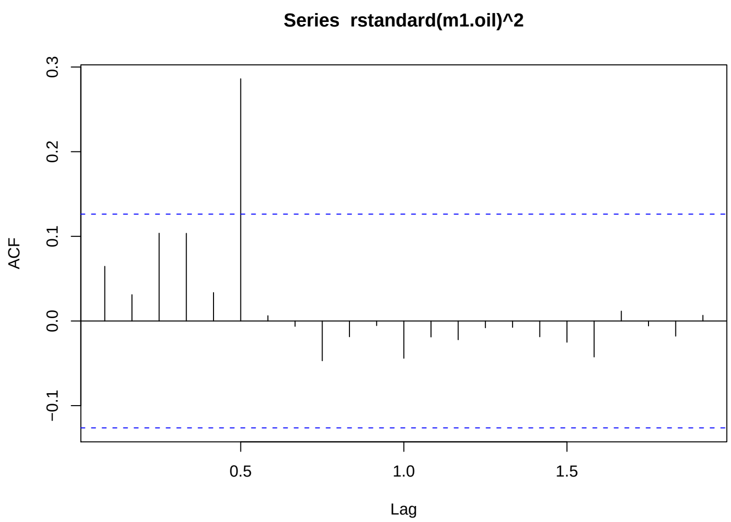


Figure 37: ACF of squared residuals of m1.oil

```
> o'o( ' c ' c ( 1. ) 2)
```

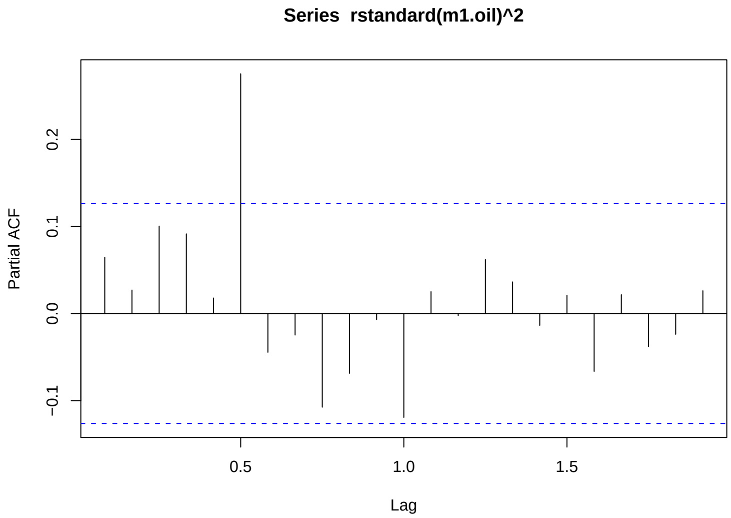


Figure 38: PACF of squared residuals of m1.oil

```
> e'o( ' c ' c ( 1. ) 2)

  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0           x
1 x         x x
2 x x
3 x x
4 x
5   x x   x
6 x   x x   x
7 x   x     x
```

The sample ACF, PACF, and EACF of the absolute residuals and those of the squared residuals in the above graphs display some significant autocorrelations and suggest that a GARCH$(1,1)$ model may be appropriate for the residuals.

(b) Let's fit an IMA$(1,1)$+GARCH$(1,1)$ model to the logarithms of monthly oil prices.

```
> 2= ' o    '= ' '(0,1)+ ' o (1,1),c ' '= ( ( .o o )),
+ o c . e ' = )
```

```
> ( 2)

:

:
= (0, 1) + (1, 1), = ( ( . )),
. = )

:
(0, 1) + (1, 1)

:


( ):
1 1 1
0.271071054 0.000938987 0.192378356 0.661455778


:
(> )
1 0.2710711 0.0761841 3.558 0.000374 ***
0.0009390 0.0005427 1.730 0.083604 . 1 0.1923784
0.0749343 2.567 0.010250 * 1 0.6614558 0.1148448
5.760 8.43 -09 ***
---
. : 0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

:
-277.2456 : -1.155190

:
-
- 2 6.637005 0.03620702
- 0.991279 0.1632340
- (10) 10.36138 0.4093809
```

```
                         (15)   23.84615   0.06775269
                         (20)   36.94588   0.01187811
                      2  (10)   3.985653   0.9479919
                      2  (15)   7.576812   0.9396271
                      2  (20)   10.03166   0.9675942
                              2   6.80259    0.8703787
```

```
:
2.343713 2.401724 2.343170 2.367087
```

From the summary of the model, we see that all of the estimates for the parameters are significantly different from 0. But the residuals tests are somewhat contradictory. The p-value for Jarque-Bera test is 0.0362 which means the normality assumption is rejected at a usual confidence level, while the p-value of Shapiro-Wilk test, 0.163, doesn't suggest rejection of the normality assumption.

(c) The time-sequence plot for the standardized residuals from the fitted IMA$(1,1)$+GARCH$(1,1)$ model is given below.

```
> o    o o (o   (( (  2)-  ( (  2)))/ ( (  2)),
+    = '   2',  = '))
```

From the above graph, we can see that at 3 time points the standard residuals are particular large with one residual close to 3 and the other two residuals around 4. This phenomena tells us that there may exist some outliers for this model.

(d) Next we are going to fit an IMA$(1,1)$ model with two IO's at $t = 2$ and $t = 56$ and an AO at $t = 8$.

```
>  3=( ( .o ),  =(0,1,1), =(2,56),
+  = ( = ( ( .o ))==8))
> o  ( ( 3),  =  3',  = ')
```

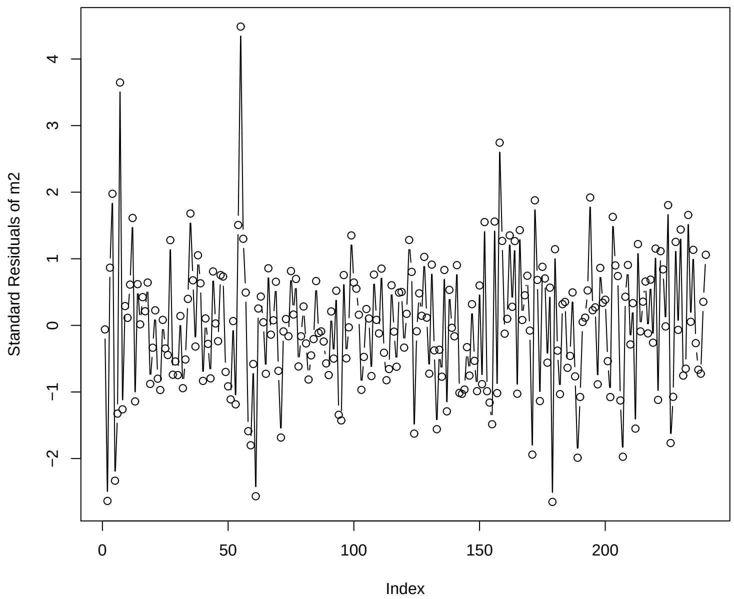


Figure 39: Standard Residuals of m2

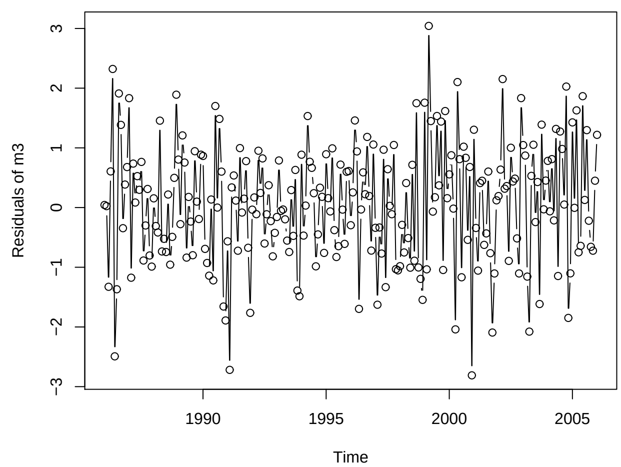


Figure 40: Standard Residuals of m3

Under this model, the standard residuals seem independently and identically distributed. Let's check the normality.

```
>       (    ' ' ( 3))
>       (    ' ' ( 3))
```

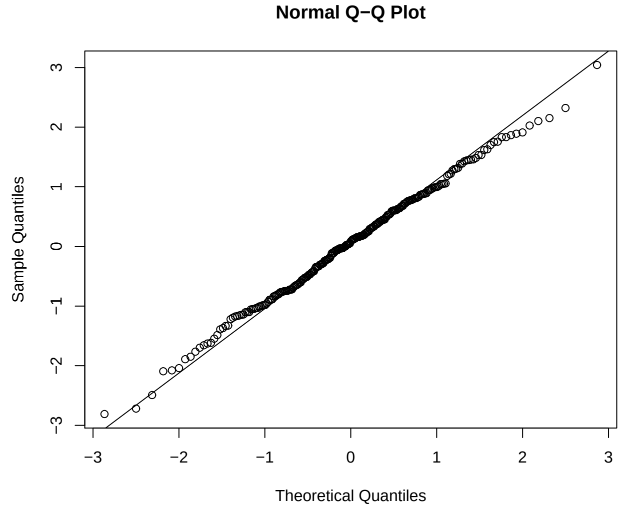


Figure 41: QQ-plot of standard residuals of m3

```
>  'o    .    (   '  ' ( 3))

        'o   -         '

 ' ':      '  ' ( 3)   = 0.9972, o- '   = 0.9539
```

Both of the qqnorm plot and Shapiro-Wilk test accept the normality of the residuals. So there is no discernible volatility clustering.

(e) Comparing the analysis of standard residuals for the two models in part (c) and (d), the IMA$(1, 1)$ model is more appropriate for the oil price data. The volatility clustering is not so discernible that the IMA+GARCH model doesn't fit well.

Chapter 13

Exercise 13.1

Since

$$3\cos(2\pi f t + 0.4) = 3\left[\cos(2\pi f t)\cos(0.4) - \sin(2\pi f t)\sin(0.4)\right]$$

it is clear that

$$A = 3\cos(0.4)\,, B = -3\sin(0.4)\,.$$

Exercise 13.2

In this problem, $A = 1$ and $B = 3$. So

$$R = \sqrt{A^2 + B^2} = \sqrt{10},$$
$$\Phi = \arctan(-B/A) = \arctan(-3)\,.$$

Exercise 13.3

(a) Regress the time series y on $\cos\left(2\pi t\frac{4}{96}\right)$ and $\sin\left(2\pi t\frac{4}{96}\right)$.

```
>  =1:96 #  =96
> ɒ  1=ɒ  (2*o * *4/96)
> ɒ  2=ɒ  (2*o *( *14/96+.3))
>    1=   (2*o * *4/96)
>  =2*ɒ  1+3*ɒ  2
>  1=  ( ɒ  1+   1-1)
>      (  1)

   :
        =  ɒ  1 +   1 - 1)

     :
```

```
                   1                 3
-2.996 +00 -2.063 +00  8.077 -15  2.063 +00  2.996 +00

          :
                                   (>  )
  1 2.000 +00  3.094 -01    6.464 4.51 -09 ***
  1 4.111 -16  3.094 -01 1.33 -15          1
---
          :  0  ***' 0.001  **' 0.01  *' 0.05  .' 0.1   ' 1

                    : 2.144    94
                  : 0.3077,                    : 0.293
 -            : 20.89    2     94   ,  o-     : 3.119 -08
```

The estimates are $A = 2$ and $B = 0$.

(b) According to Equations (13.1.5), for the cosine component at frequency $f = 14/96$, $A = 3\cos(0.6\pi) = -0.927$ and $B = -3\sin(0.6\pi) = -2.853$.

(c) Now regress the time series y on $\cos\left(2\pi t\frac{14}{96}\right)$ and $\sin\left(2\pi t\frac{14}{96}\right)$.

```
>    3=   (2*   * *14/96)
>    3=   (2*   * *14/96)
>   2=  (    3+    3-1)
>        ( 2)

    :
          =        3 +     3 - 1)

         :
                   1                 3
-2.000 +00 -1.414 +00  5.877 -15  1.414 +00  2.000 +00

          :
                                  (>  )
  3  -0.9271      0.2063  -4.494 1.99 -05 ***
  3  -2.8532      0.2063 -13.831  < 2 -16 ***
---
```

```
  . :   0  ***' 0.001  **' 0.01  *' 0.05  .' 0.1   ' 1

          : 1.429    94
     :  0.6923,             : 0.6858
 -     : 105.7     2     94    ,   -       : < 2.2  -16
```

The estimates of A and B are the same as what we got in part (b).

(d) Let us regress the series y on $\cos(2\pi t f)$ and $\sin(2\pi t f)$ for both $f = 4/96$ and $f = 14/96$.

```
>  3=  (     1+   1+   3+   3-1)
>       ( 3)

  :
     =       1 +    1 +    3 +    3 - 1)

   :
                  1                3
-5.128  -14  -5.856  -15  -9.716  -16   8.103  -15   2.862  -14

          :
                                (>   )
  1   2.000  +00   1.884  -15   1.061  +15    <2  -16  ***
  1  -1.395  -15   1.884  -15  -7.400  -01     0.461
  3  -9.271  -01   1.884  -15  -4.920  +14    <2  -16  ***
  3  -2.853  +00   1.884  -15  -1.514  +15    <2  -16  ***
---
  . :   0  ***' 0.001  **' 0.01  *' 0.05  .' 0.1   ' 1

          : 1.306  -14     92
     :      1,             :      1
 -     : 9.152  +29     4     92    ,   -       : < 2.2  -16
```

From the regression, $A_4 = 2, B_4 = 0, A_{14} = -0.927$, and $B_{14} = -2.853$. These are perfect estimates.

(e) Let us try to regress the series y on $\cos(2\pi t f)$ and $\sin(2\pi t f)$ for both $f = 3/96$ and $f = 13/96$.

```
> 4=  (2*  * *3/96)
> 4=  (2*  * *3/96)
> 5=  (2*  * *13/96)
> 5=  (2*  * *13/96)
> 4=  (  4+  4+  5+  5-1)
>       ( 4)

      :
        =      4 +    4 +    5 +    5 - 1)

        :
              1          3
-4.8532 -1.8291  0.1309  1.6302  4.7598

              :
                                    (>   )
  4 -9.012  -16  3.759  -01 -2.40  -15       1
  4  8.967  -17  3.759  -01  2.39  -16       1
  5 -1.599  -16  3.759  -01 -4.25  -16       1
  5 -4.847  -15  3.759  -01 -1.29  -14       1

                   : 2.604    92
             : 1.895  -30,                : -0.04348
 -        : 4.358  -29    4     92    ,   -     : 1
```

We see that all the estimates for A_3, B_3, A_{13}, and B_{13} are 0.

(f) Now we redo the regression in part (b) but add a third pair, $\cos\left(2\pi t\frac{7}{96}\right)$ and $\sin\left(2\pi t\frac{7}{96}\right)$, as predictor variables.

```
> 6=  (2*  * *7/96)
> 6=  (2*  * *7/96)
> 5=  (  1+  1+  3+  3+  6+  6-1)
>       ( 5)

      :
        =      1 +    1 +    3 +    3 +    6 +    6 - 1)
```

```
                     1                    3
-5.067 -14  -6.223 -15  -1.473 -16   8.207 -15   2.780 -14

                                          (>   )
   1  2.000 +00   1.866 -15   1.072 +15    <2 -16 ***
   1 -1.395 -15   1.866 -15  -7.480 -01    0.4566
   3 -9.271 -01   1.866 -15  -4.969 +14    <2 -16 ***
   3 -2.853 +00   1.866 -15  -1.529 +15    <2 -16 ***
   6 -1.278 -15   1.866 -15  -6.850 -01    0.4950
   6 -3.438 -15   1.866 -15  -1.843 +00    0.0687 .
---
            :  0  ***' 0.001  **' 0.01  *' 0.05  .' 0.1   ' 1

                    : 1.293 -14     90
                :     1,                   :     1
 -          : 6.225 +29     6      90   ,   o-      : < 2.2 -16
```

From the new regression, the estimates are $A_4 = 2, B_4 = 0, A_{14} = -0.927, B_{14} = -2.853$, and $A_7 = B_7 = 0$. All of these estimates are perfect.

Exercise 13.4

Let $\{Y_t, t = 1, \ldots, 10\}$ be an arbitrary series of length 10. We denote

$$\widehat{A}_0 = \overline{Y},$$

$$\widehat{A}_j = \frac{1}{5}\sum_{t=1}^{10} Y_t \cos\frac{\pi t j}{5} \quad \text{and} \quad \widehat{B}_j = \frac{1}{5}\sum_{t=1}^{10} Y_t \sin\frac{\pi t j}{5}, \quad \text{for } j = 1, \ldots, 4,$$

$$\widehat{A}_5 = \frac{1}{10}\sum_{t=1}^{10} (-1)^t Y_t \quad \text{and} \quad \widehat{B}_5 = 0.$$

Then for each $t = 1, \ldots, 10$, we have

$$
\begin{aligned}
&\widehat{A}_0 + \sum_{j=1}^{5} \left[\widehat{A}_j \cos \frac{\pi t j}{5} + \widehat{B}_j \sin \frac{\pi t j}{5} \right] \\
&= \overline{Y} + \frac{1}{10} \sum_{s=1}^{10} Y_s \left[2 \sum_{j=1}^{4} \left(\cos \frac{\pi s j}{5} \cos \frac{\pi t j}{5} + \sin \frac{\pi s j}{5} \sin \frac{\pi t j}{5} \right) + (-1)^s \cos (\pi t) \right] \\
&= \frac{1}{10} \sum_{s=1}^{10} Y_s \left[1 + \sum_{j=1}^{10} \cos \frac{2\pi (s-t) j}{10} + (-1)^s \cos (\pi t) - \cos (\pi (s-t)) - \cos (2\pi (s-t)) \right] \\
&= \sum_{s=1}^{10} Y_s 1_{s=t} = Y_t,
\end{aligned}
$$

where 1_X is the indicator function of event X. So $\{Y_t, t = 1, \ldots, 10\}$ is exactly fit by a linear combination of cosine-sine curves at the Fourier frequencies.

Exercise 13.5

(a) Let us use the same parameter values used in Exhibit (13.4) to simulate a signal+noise time series. The plot is given below. It is hard to see periodicities from this plot.

```
> (  )
> (  =4.875,   =2.5,   =8)
>  =1:96
> . (124)
>    = (48,2)
>  1=     1/96;  2=     2/96
>  1=     (1,0,2);  1=     (1,0,2)  #
>  2=     (1,0,3);  2=     (1,0,3)
>  =2* *
>  = 1*  (   1)+ 1*  (   1)+ 2*  (   2)+ 2*  (   2)+
+     (96,0,1)
>  ( , ,  =' ',  =  (  ))
```

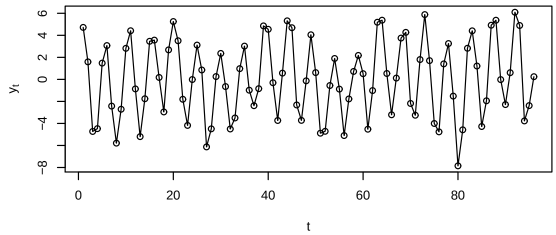


Figure 1: Simulated Signal Plus Noise Time Series

(b) Now let us see the periodogram for the simulated time series.

```
> 'o ( =4.875, =2.5,o =8)
> ( , =' ', ' ='', ' =' ' ' =' o ',
+ =' ', =2, =''))
```

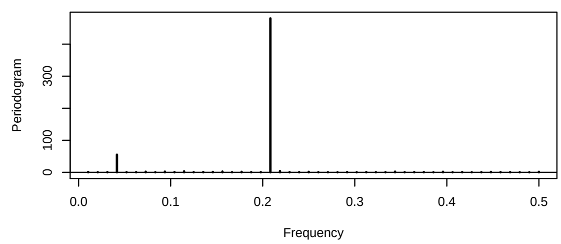


Figure 2: Periodogram for the Simulated Signal Plus Noise Time Series

It is clear that the time series contains two cosine-sine pairs at frequencies of about 0.04 and 0.21 and that the higher frequency component is much stronger. Actually, one frequency is chosen to be $4/96 \approx 0.0417$ and the other is $20/96 \approx 0.2083$.

Exercise 13.6

Equation (13.3.1) tells us that

$$Y_t = \sum_{j=1}^{m} \left[A_j \cos\left(2\pi f_j t\right) + B_j \sin\left(2\pi f_j t\right) \right],$$

where the frequencies $0 < f_1 < f_2 < \ldots < f_m < 1/2$ are fixed and A_j and B_j are independent normal random variables with zero means and $Var\left(A_j\right) = Var\left(B_j\right) = \sigma_j^2$. So the variance function is

$$\begin{aligned}\gamma_k &= Cov\left(Y_t, Y_{t-k}\right)\\ &= \sum_{j=1}^{m} \sigma_j^2 \left[\cos\left(2\pi f_j t\right)\cos\left(2\pi f_j\left(t-k\right)\right) + \sin\left(2\pi f_j t\right)\sin\left(2\pi f_j\left(t-k\right)\right)\right]\\ &= \sum_{j=1}^{m} \sigma_j^2 \cos\left(2\pi f_j k\right),\end{aligned}$$

as given in Equation (13.3.2).

Exercise 13.7

Following the assumption given right before Equation (13.3.10) in the book, we assume that Y_t represents deviations from the sample mean. Then for $-1/2 < f < 1/2$, we have

$$\begin{aligned}\widehat{S}\left(f\right) &= \frac{1}{2} I\left(f\right)\\ &= \frac{n}{4}\left(\widehat{A}_f^2 + \widehat{B}_f^2\right)\\ &= \frac{1}{n}\left[\left(\sum_{t=1}^{n} Y_t \cos\left(2\pi t f\right)\right)^2 + \left(\sum_{t=1}^{n} Y_t \sin\left(2\pi t f\right)\right)^2\right]\\ &= \frac{1}{n}\left[\sum_{t=1}^{n} Y_t^2 + 2\sum_{i<j} Y_i Y_j \cos\left(2\pi\left(j-i\right)f\right)\right]\\ &= \frac{1}{n}\sum_{t=1}^{n} Y_t^2 + \frac{2}{n}\sum_{k=1}^{n-1}\cos\left(2\pi k f\right)\sum_{t=k+1}^{n} Y_t Y_{t-k}\\ &= \widehat{\gamma}_0 + 2\sum_{k=1}^{n-1}\widehat{\gamma}_k \cos\left(2\pi k f\right). \qquad (1)\end{aligned}$$

For $f = 1/2$, we have

$$\begin{aligned}\widehat{S}\left(\frac{1}{2}\right) &= I\left(\frac{1}{2}\right)\\ &= n\widehat{A}_{\frac{1}{2}}^2\\ &= \frac{1}{n}\left[\sum_{t=1}^{n}\left(-1\right)^t Y_t\right]^2\\ &= \frac{1}{n}\left[\sum_{t=1}^{n} Y_t^2 + 2\sum_{i<j} Y_i Y_j \left(-1\right)^{i-j}\right]\\ &= \widehat{\gamma}_0 + 2\sum_{k=1}^{n-1}\widehat{\gamma}_k\left(-1\right)^k. \qquad (2)\end{aligned}$$

Hence, by (1) and (2), Equation (13.3.10) holds.

Exercise 13.8

Let $\{\gamma_k\}, \{\gamma_{X,k}\}$, and $\{\gamma_{Y,k}\}$ be the covariance functions of $\{X_t + Y_t\}, \{X_t\}$, and $\{Y_t\}$ respectively. The the spectral density of $\{X_t + Y_t\}$ is given by

$$S(f) = \gamma_0 + 2\sum_{k=1}^{\infty} \gamma_k \cos(2\pi f k). \tag{3}$$

Since $\{X_t\}$ and $\{Y_t\}$ are independent series, we have

$$\begin{aligned} \gamma_k &= Cov(X_t + Y_t, X_{t-k} + Y_{t-k}) \\ &= Cov(X_t, X_{t-k}) + Cov(Y_t, Y_{t-k}) \\ &= \gamma_{X,k} + \gamma_{Y,k}, \quad k = 0, 1, \ldots. \end{aligned} \tag{4}$$

Plugging (4) into (3), we obtain

$$\begin{aligned} S(f) &= \gamma_0 + 2\sum_{k=1}^{\infty} \gamma_k \cos(2\pi f k) \\ &= (\gamma_{X,0} + \gamma_{Y,0}) + 2\sum_{k=1}^{\infty} (\gamma_{X,k} + \gamma_{Y,k}) \cos(2\pi f k) \\ &= S_X(f) + S_Y(f). \end{aligned}$$

Exercise 13.9

According to Equation (13.5.3), the theoretical spectral density for an MA(1) model is

$$S(f) = \left[1 + \theta^2 - 2\theta\cos(2\pi f)\right]\sigma_e^2.$$

The first derivative of the spectral density with respect to f is $4\pi\theta\sigma_e^2 \sin(2\pi f)$, which is positive when $\theta > 0$ and negative when $\theta < 0$, where $f \in (0, 1/2)$. So when $\theta > 0$ the spectral density for an MA(1) model is an increasing function while for $\theta < 0$ the spectral density decreases.

Exercise 13.10

Figure 3 below shows the theoretical spectral density function for an MA(1) process with $\theta = 0.6$. The density is much stronger for higher frequencies than for low frequencies. High frequency means that the process has a tendency to oscillate back and forth across its mean level quickly.

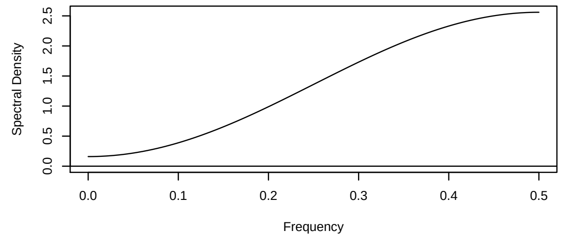


Figure 3: Spectral Density of MA(1) Process with $\theta = 0.6$

Exercise 13.11

Figure 4 below shows the theoretical spectral density function for an MA(1) process with $\theta = -0.8$. The density is much stronger for lower frequencies than for high frequencies. Such a process tends to change slowly from one time instance to the next.

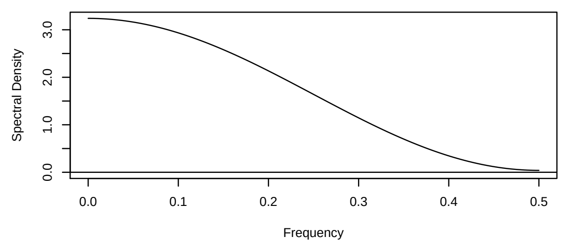


Figure 4: Spectral Density of MA(1) Process with $\theta = -0.8$

Exercise 13.12

According to Equation (13.5.6), the theoretical spectral density for an AR(1) model is

$$S(f) = \frac{\sigma_e^2}{1 + \phi^2 - 2\phi\cos(2\pi f)}.$$

The first derivative of the denominator with respect to f is $4\pi\phi\sin(2\pi f)$, which is positive when $\phi > 0$ and negative when $\phi < 0$, where $f \in (0, 1/2)$. So when $\phi > 0$ the spectral density for an AR(1) model is an decreasing function while for $\phi < 0$ the spectral density increases.

Exercise 13.13

Figure 5 below shows the theoretical spectral density function for an AR(1) process with $\phi = 0.7$. The density function decreases rapidly and is much stronger for lower frequencies than for high frequencies. Such a process tends to change slowly from one time instance to the next.

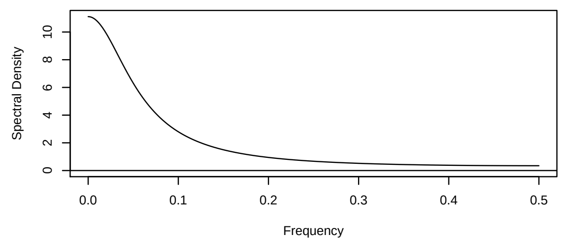


Figure 5: Spectral Density of AR(1) Process with $\phi = 0.7$

Exercise 13.14

Figure 6 below shows the theoretical spectral density function for an AR(1) process with $\phi = -0.4$. The density is much stronger for higher frequencies than for low frequencies. High frequency means that the process has a tendency to oscillate back and forth across its mean level quickly.

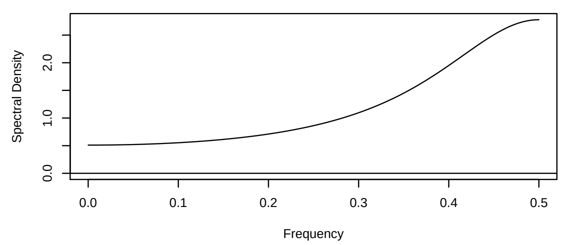


Figure 6: Spectral Density of AR(1) Process with $\phi = -0.4$

Exercise 13.15

Figure 7 below shows the theoretical spectral density function for an MA(2) process with $\theta_1 = -0.5$ and $\theta_2 = 0.9$. The spectral density has a peak at a frequency around 0.25. The process has a main component with frequency arount 0.25.

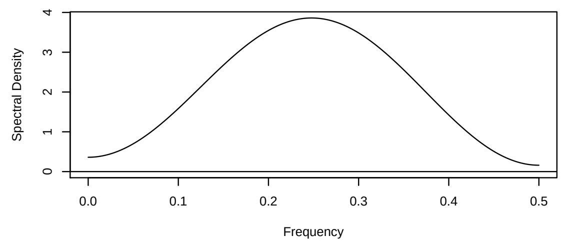


Figure 7: Spectral Density of MA(2) Process with $\theta_1 = -0.5$ and $\theta_2 = 0.9$

Exercise 13.16

Figure 8 below shows the theoretical spectral density function for an MA(2) process with $\theta_1 = 0.5$ and $\theta_2 = -0.9$. The spectral density is strong at $f = 0$ and $f = 0.5$. So the process has two main components. One component tends to change slowly from one time instance to the next; while the other component oscillates back and forth across its mean level quickly. The latter component has a stronger effect on the process.

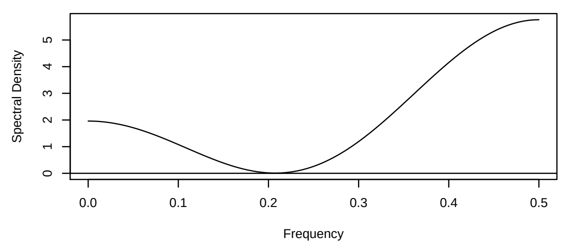


Figure 8: Spectral Density of MA(2) Process with $\theta_1 = 0.5$ and $\theta_2 = -0.9$

Exercise 13.17

Figure 9 below shows the theoretical spectral density function for an AR(2) process with $\phi_1 = -0.1$ and $\phi_2 = -0.9$. The spectral density has a sharp peak at around frequency $f = 0.26$. The process fluctuates mainly with a frequency 0.26.

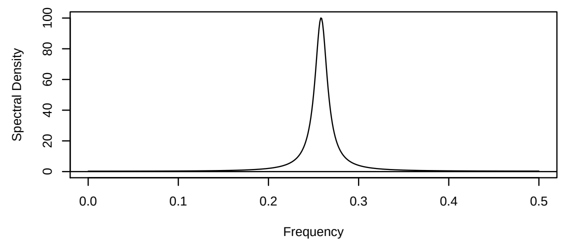


Figure 9: Spectral Density of AR(2) Process with $\phi_1 = -0.1$ and $\phi_2 = -0.9$

Exercise 13.18

Figure 10 below shows the theoretical spectral density function for an AR(2) process with $\phi_1 = 1.8$ and $\phi_2 = -0.9$. The spectral density has a sharp peak at around frequency $f = 0.06$. The process fluctuates mainly with a frequency 0.06.

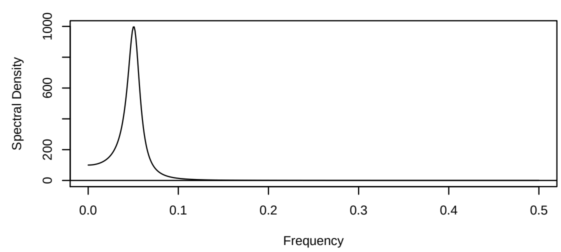


Figure 10: Spectral Density of AR(2) Process with $\phi_1 = 1.8$ and $\phi_2 = -0.9$

Exercise 13.19

Figure 11 below shows the theoretical spectral density function for an AR(2) process with $\phi_1 = -1$ and $\phi_2 = -0.8$. The spectral density has a sharp peak at around frequency $f = 0.35$. The process fluctuates mainly with a frequency 0.35.

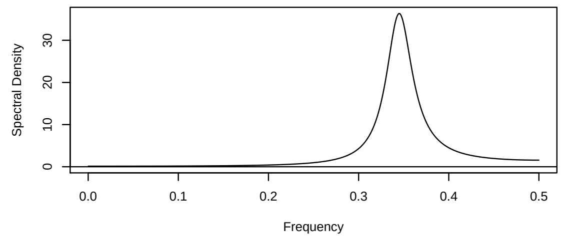


Figure 11: Spectral Density of AR(2) Process with $\phi_1 = -1$ and $\phi_2 = -0.8$

Exercise 13.20

Figure 12 below shows the theoretical spectral density function for an AR(2) process with $\phi_1 = 0.5$ and $\phi_2 = 0.4$. The density is much stronger for lower frequencies than for high frequencies. Such a process tends to change very slowly from one time instance to the next.

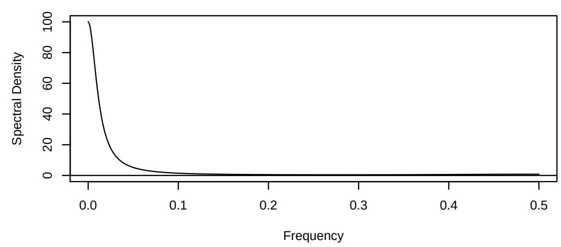


Figure 12: Spectral Density of AR(2) Process with $\phi_1 = 0.5$ and $\phi_2 = 0.4$

Exercise 13.21

Figure 13 below shows the theoretical spectral density function for an AR(2) process with $\phi_1 = 0$ and $\phi_2 = 0.8$. The strongest density appears at $f = 0$ and $f = 0.5$. So the process has two main components. One component tends to change slowly from one time instance to the next; while the other component oscillates back and forth across its mean level quickly.

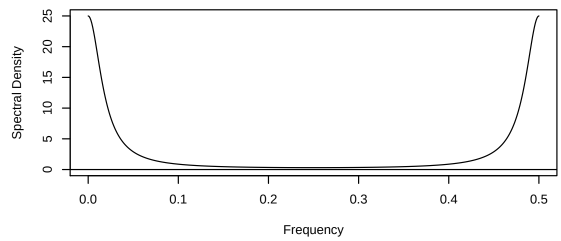


Figure 13: Spectral Density of AR(2) Process with $\phi_1 = 0$ and $\phi_2 = 0.8$

Exercise 13.22

Figure 14 below shows the theoretical spectral density function for an AR(2) process with $\phi_1 = 0.8$ and $\phi_2 = -0.2$. The density is much stronger for lower frequencies than for high frequencies. Such a process tends to change slowly from one time instance to the next.

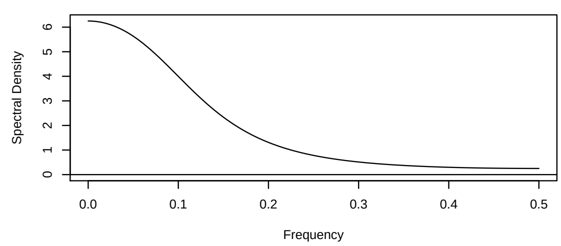


Figure 14: Spectral Density of AR(2) Process with $\phi_1 = 0.8$ and $\phi_2 = -0.2$

Exercise 13.23

Figure 15 below shows the theoretical spectral density function for an AR(2) process with $\phi = 0.5$ and $\theta = 0.8$. The density is much stronger for higher frequencies than for low frequencies. High frequency means that the process has a tendency to oscillate back and forth across its mean level quickly.

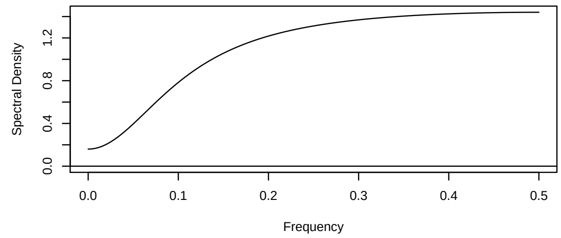


Figure 15: Spectral Density of ARMA(1,1) Process with $\phi = 0.5$ and $\theta = 0.8$

Exercise 13.24

Figure 16 below shows the theoretical spectral density function for an AR(2) process with $\phi = 0.95$ and $\theta = 0.8$. The density is much stronger for frequencies greater than 0.2 than for frequencies less than 0.2. The process has a tendency to oscillate back and forth across its mean level quickly. Frequencies from 0.2 to 0.5 have almost same effect on the process.

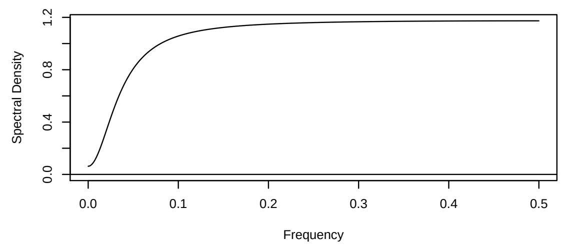


Figure 16: Spectral Density of ARMA(1,1) Process with $\phi = 0.95$ and $\theta = 0.8$

Exercise 13.25

(a) Since $Y_t = (X_t + X_{t-1})/2$, we have $c_0 = 1/2, c_1 = 1/2$ and $Y_t = \sum_{i=0}^{1} c_i X_{t-i}$. The power transfer function for this linear filter is given by

$$\begin{aligned} &\left|C\left(\mathrm{e}^{-2\pi i f}\right)\right|^2 \\ = &\left|\frac{1}{2}\mathrm{e}^{-2\pi i f} + \frac{1}{2}\right|^2 = \frac{1}{2}\left[1 + \cos\left(2\pi f\right)\right]. \end{aligned}$$

(b) Since $c_k = 0$, for $k < 0$, this filter is causal.

(c) The plot of the power transfer function is given below. It shows that this linear filter retains lower frequencies and de-emphasizes higher frequencies.

```
> =(1:50)/100
> 1=1/2*(1+  (2*  ))
> o    ,1,   =" ",     =            ')
```

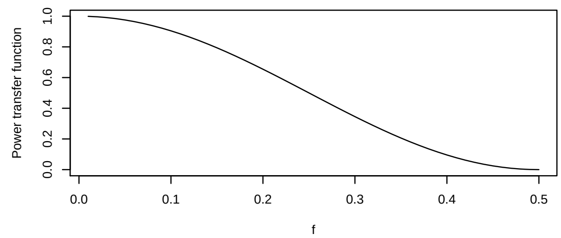


Figure 17: Power Transfer Function

Exercise 13.26

(a) Since $Y_t = X_t - X_{t-1}$, we have $c_0 = 1, c_1 = -1$ and $Y_t = \sum_{i=0}^{1} c_i X_{t-i}$. The power transfer function for this linear filter is given by

$$\begin{aligned} &\left|C\left(\mathrm{e}^{-2\pi i f}\right)\right|^2 \\ =& \left|1-\mathrm{e}^{-2\pi i f}\right|^2 = 2 - 2\cos\left(2\pi f\right). \end{aligned}$$

(b) Since $c_k = 0$, for $k < 0$, this filter is causal.

(c) The plot of the power transfer function is given below. It shows that this linear filter retains higher frequencies and de-emphasizes lower frequencies.

```
> 2=2-2*  (2*  )
> o    ,2,   =" ",     =            ')
```

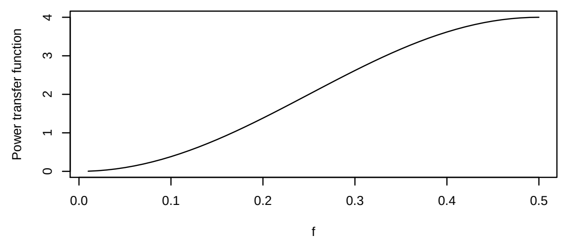


Figure 18: Power Transfer Function

Exercise 13.27

(a) Since $Y_t = (X_{t+1} + X_t + X_{t-1})/3$, we have $c_{-1} = c_0 = c_1 = 1/3$ and $Y_t = \sum_{i=-1}^{1} c_i X_{t-i}$. The power transfer function for this linear filter is given by

$$
\begin{aligned}
& \left|C\left(\mathrm{e}^{-2\pi i f}\right)\right|^2 \\
= & \left|\frac{1}{3}\mathrm{e}^{-2\pi i f} + \frac{1}{3} + \frac{1}{3}\mathrm{e}^{2\pi i f}\right|^2 \\
= & \frac{1}{9}\left[1 + 2\cos(2\pi f)\right]^2 \\
= & \frac{1}{9}\left[1 + 4\cos(2\pi f) + 4\cos^2(2\pi f)\right].
\end{aligned}
$$

(b) Since $c_{-1} = 1/3$, this filter is not causal.

(c) The plot of the power transfer function is given below. It shows that this linear filter retains lower frequencies and de-emphasizes frequencies in $(0.3, 0.4)$.

```
> p3=(1+2*p  (2*p  )) 2/9
> p   ,p3,   =" ",   =       p    ')
```

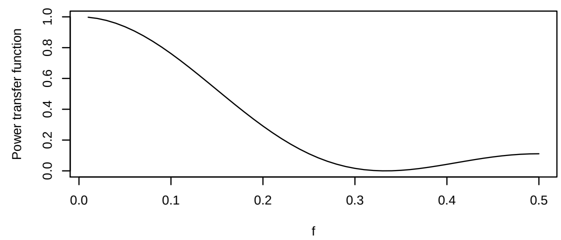


Figure 19: Power Transfer Function

Exercise 13.28

(a) Since $Y_t = (X_t + X_{t-1} + X_{t-2})/3$, we have $c_0 = c_1 = c_2 = 1/3$ and $Y_t = \sum_{i=0}^{2} c_i X_{t-i}$. The power transfer function for this linear filter is given by

$$
\begin{aligned}
&\left|C\left(\mathrm{e}^{-2\pi i f}\right)\right|^2 \\
&= \left|\frac{1}{3} + \frac{1}{3}\mathrm{e}^{-2\pi i f} + \frac{1}{3}\mathrm{e}^{-4\pi i f}\right|^2 \\
&= \frac{1}{9}\left[1 + \cos(2\pi f) + \cos(4\pi f)\right]^2 + \frac{1}{9}\left[\sin(2\pi f) + \sin(4\pi f)\right]^2 \\
&= \frac{1}{9}\left[3 + 4\cos(2\pi f) + 2\cos(4\pi f)\right] \\
&= \frac{1}{9}\left[1 + 4\cos(2\pi f) + 4\cos^2(2\pi f)\right],
\end{aligned}
$$

which is the same as the power transfer function of the filter defined in Exercise (13.27).

(b) Since $c_k = 0$, for $k < 0$, this filter is causal.

Exercise 13.29

Same as Exercise 13.26.

Exercise 13.30

(a) Since $Y_t = (X_{t+1} - 2X_t + X_{t-1})/3$, we have $c_{-1} = 1/3, c_0 = -2/3, c_1 = 1/3$ and $Y_t = \sum_{i=-1}^{1} c_i X_{t-i}$. The power transfer function for this linear filter is given by

$$
\begin{aligned}
&\left|C\left(\mathrm{e}^{-2\pi i f}\right)\right|^2 \\
&= \left|\frac{1}{3}\mathrm{e}^{-2\pi i f} - \frac{2}{3} + \frac{1}{3}\mathrm{e}^{2\pi i f}\right|^2 \\
&= \left|\frac{2}{3}\cos(2\pi f) - \frac{2}{3}\right|^2 \\
&= \frac{4}{9}\left[1 + \cos^2(2\pi f) - 2\cos(2\pi f)\right].
\end{aligned}
$$

(b) The plot of the power transfer function is given below. It shows that this linear filter retains higher frequencies and de-emphasizes lower frequencies.

```
> [illegible] =4.875, [illegible] =2.5, [illegible] =8)
> [illegible] =(1:50)/100
> [illegible]1=4/9*(1+[illegible] (2*[illegible] ) 2-2*[illegible] (2*[illegible] ))
> [illegible] ,[illegible]1, [illegible] =" ", [illegible] ')
```

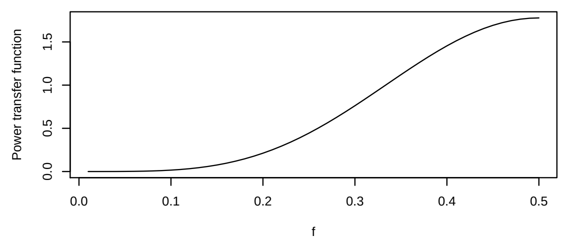


Figure 20: Power Transfer Function

Exercise 13.31

(a) Suppose $\{Y_t\}$ is a white noise process with variance γ_0. The sample spectral density at Fourier frequency $f \in [0, 1/2]$ is

$$
\widehat{S}(f) = \frac{n}{4}\left(\widehat{A}_f^2 + \widehat{B}_f^2\right),
$$

where

$$\widehat{A}_f = \frac{2}{n}\sum_{t=1}^{n} Y_t \cos(2\pi t f) \qquad \text{and} \qquad \widehat{B}_f = \frac{2}{n}\sum_{t=1}^{n} Y_t \sin(2\pi t f).$$

So

$$\begin{aligned}\widehat{S}(f) &= \frac{1}{n}\left(\sum_{t=1}^{n} Y_t \cos(2\pi t f)\right)^2 + \frac{1}{n}\left(\sum_{t=1}^{n} Y_t \sin(2\pi t f)\right)^2 \\ &= \frac{1}{n}\left[\sum_{t=1}^{n} Y_t^2 + 2\sum_{1\le i<j\le n} Y_i Y_j \cos(2\pi i f - 2\pi j f)\right]\end{aligned}$$

which implies

$$\mathrm{E}\widehat{S}(f) = \frac{1}{n}\left[\sum_{t=1}^{n} \mathrm{E}Y_t^2 + 2\sum_{1\le i<j\le n} 0\cdot\cos(2\pi i f - 2\pi j f)\right] = \gamma_0 = S(f).$$

(b) For any two Fourier frequencies f_1 and f_2,

$$\begin{aligned}Cov\left(\widehat{A}_{f_1}, \widehat{B}_{f_2}\right) &= \frac{2}{n^2} Cov\left(\sum_{t=1}^{n} Y_t \cos(2\pi t f_1), \sum_{t=1}^{n} Y_t \sin(2\pi t f_2)\right) \\ &= \frac{2\gamma_0}{n^2}\sum_{t=1}^{n} \cos(2\pi t f_1)\sin(2\pi t f_2) = 0.\end{aligned}$$

(c) For any two unequal Fourier frequencies f_1 and f_2,

$$\begin{aligned}Cov\left(\widehat{A}_{f_1}, \widehat{A}_{f_2}\right) &= \frac{2}{n^2} Cov\left(\sum_{t=1}^{n} Y_t \cos(2\pi t f_1), \sum_{t=1}^{n} Y_t \cos(2\pi t f_2)\right) \\ &= \frac{2\gamma_0}{n^2}\sum_{t=1}^{n} \cos(2\pi t f_1)\cos(2\pi t f_2) = 0.\end{aligned}$$

Exercise 13.32

Note that there is a typo in the first printing of the book as $\phi_2 = -0.8$, not 0.8.

(a) Figure 21 gives the average sample spectral density by Fourier frequencies (circles) and the theoretical spectral density (solid line). The approximation seems to be quite acceptable.

(b) The figure below compares the standard deviation of the sample spectral density with the large-sample analogue.

(c) Figures 23 and 24 below show the QQ plots comparing the observed quantiles with those of a chi-square distribution with two degrees of freedom for Fourier frequencies $f_1 = 7/48$,

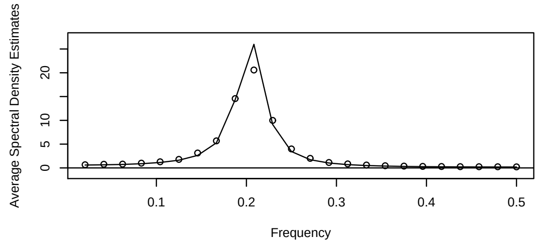


Figure 21: Average Sample Spectral Density: Simulated AR(2), $\phi_1 = 0.5$, $\phi_2 = -0.8$, $n = 48$

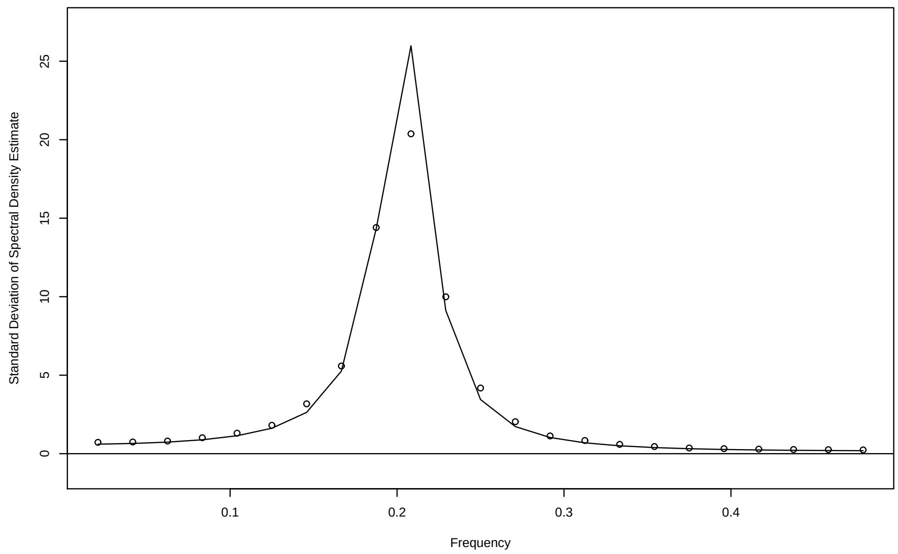


Figure 22: Standard deviation of the Sample Spectral Density: Simulated AR(2), $\phi_1 = 0.5$, $\phi_2 = -0.8$, $n = 48$

and $f_2 = 20/48$. The agreement with chi-square appears to be acceptable in each of these 2 plots. But one of these results tell us that the sample spectral density is not an acceptable estimator of the underlying theoretical spectral density, because the sample spectral density is inconsistent with too much variability to be a useful estimator.

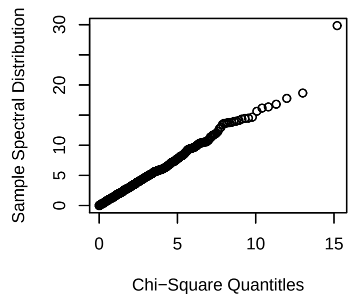


Figure 23: QQ Plot of Spectral Distribution at $f = 7/48$

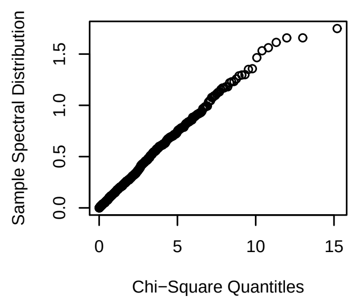


Figure 24: QQ Plot of Spectral Distribution at $f = 20/48$

Exercise 13.33

(a) Figure 25 gives the average sample spectral density by Fourier frequencies (circles) and the theoretical spectral density (solid line). The approximation seems to be quite acceptable.

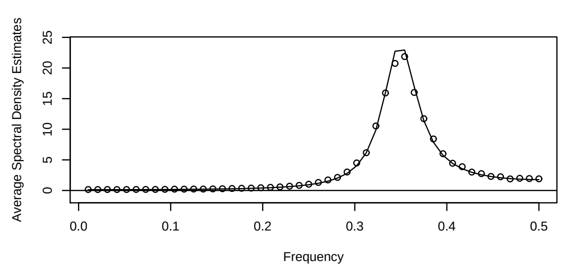


Figure 25: Average Sample Spectral Density: Simulated AR(2), $\phi_1 = -1$, $\phi_2 = -0.75$, $n = 96$

(b) Same question as Part (a).

(c) Figures 26, 27, and 28 below show the QQ plots comparing the observed quantiles with those of a chi-square distribution with two degrees of freedom for Fourier frequencies $f_1 = 7/96$, $f_2 = 20/96$, and $f_3 = 31/96$. The agreement with chi-square appears to be acceptable in each of these 3 plots. But one of these results tell us that the sample spectral density is an acceptable estimator of the underlying theoretical spectral density, because the sample spectral density is inconsistent with too much variability to be a useful estimator.

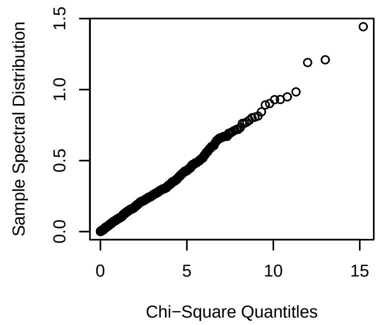


Figure 26: QQ Plot of Spectral Distribution at $f = 7/96$

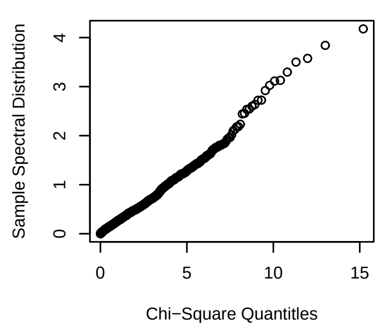


Figure 27: QQ Plot of Spectral Distribution at $f = 20/96$

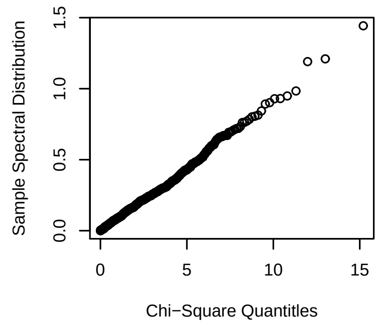


Figure 28: QQ Plot of Spectral Distribution at $f = 31/96$

Exercise 13.34

The periodogram of the simulated series of $n = 1000$ zero mean, unit variance, normal white noises is given below.

```
>  =1000
>    (357864)
>     ‘o (    =4.875,     =2.5,o     =8)
>  o o( ,  =’  ’,  ‘ =   ‘ ’, ‘ =’     o ’,  o =’ ’,  =2,  =’’)
```

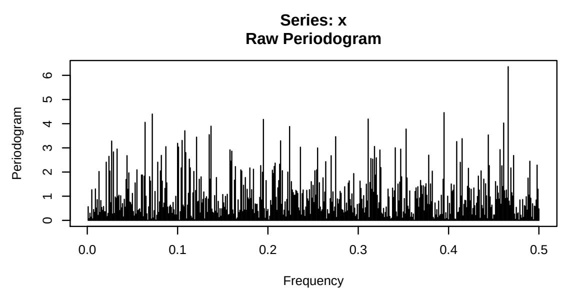


Figure 29: Periodogram of Simulated Normal White Noises

We know that the theoretical spectral density for a white noise process is constant for all frequency in $-1/2 < f \leq 1/2$. Figure 29 fits this feature.

Chapter 14

Exercise 14.1

Let f be a Fourier frequency. The smoothed sample spectral density with Daniell spectral window is

$$\overline{S}\left(f\right)=\frac{1}{2m+1}\sum_{j=-m}^{m}\widehat{S}\left(f+\frac{j}{n}\right). \tag{1}$$

Dividing both sides of Equation (1) by $S\left(f\right)$, we get

$$\frac{\overline{S}\left(f\right)}{S\left(f\right)}=\frac{1}{2\left(2m+1\right)}\sum_{j=-m}^{m}\frac{2\widehat{S}\left(f+\frac{j}{n}\right)}{S\left(f\right)}.$$

Let us assume that the spectral density changes very little over a small interval of frequencies. We already know from the book that the sample spectral density values at the Fourier frequencies are approximately uncorrelated and $2\widehat{S}\left(f\right)/S\left(f\right)$ has approximately a chi-square distribution with two degrees of freedom. So, from the above equation we have

$$\begin{aligned}Var\left(\frac{\overline{S}\left(f\right)}{S\left(f\right)}\right)&\approx\frac{1}{4\left(2m+1\right)^{2}}\sum_{j=-m}^{m}4\left[\frac{S\left(f+\frac{j}{n}\right)}{S\left(f\right)}\right]^{2}\\&\approx\frac{1}{2m+1},\end{aligned}$$

which implies

$$Var\left(\overline{S}\left(f\right)\right)\approx\frac{S^{2}\left(f\right)}{2m+1}.$$

Exercise 14.2

(a) The following panel gives the Daniell spectral window with $m=5$ and its 2nd and 3rd convolutions.

```
>   o‘ =o‘
> o‘ (      =(1,3))
> o   (    ("‘      ", (5,5,5)),  =2, ‘  =’’,   =(0,.1),
+    ‘ = o         (( )))
> ‘    ( =0)
```

```
> o    (        ("  '        ", ρ(5,5)),   =2,  '    ='',      =ρ(0,.1),
+     ' =  o         (`( )))
> '      ( =0)
> o   `(        ("  '        ", ρ(5)),   =2,  '    ='',      =ρ(0,.1),
+     ' =  o         (`( )))
> '      ( =0)
> o' =   o'
```

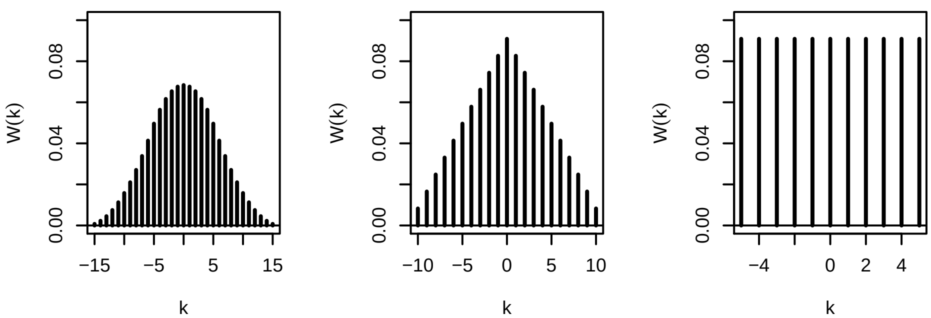


Figure 1: The Daniell Spectral Window and Its Convolutions

(b) The bandwidths and degrees of freedom for each of the spectral windows in part (a) are given below. $n = 100$ is used.

```
>  1=        ("  '       ", ρ(5,5,5))
>  2=        ("  '       ", ρ(5,5))
>  3=        ("  '       ", ρ(5))
>  1=2/(    (( 1$ρ  ) 2)*2-  1$ρ    1  2)
> ρ1=ρ(0:15) 2
>    1=1/100*     (    (ρ1*( 1$ρ  ))*2)
>  2=2/(    (( 2$ρ  ) 2)*2-  2$ρ    1  2)
> ρ2=ρ(0:10) 2
>    2=1/100*     (    (ρ2*( 2$ρ  ))*2)
>  3=2/(    (( 3$ρ  ) 2)*2-  3$ρ    1  2)
> ρ3=ρ(0:5) 2
>    3=1/100*     (    (ρ3*( 3$ρ  ))*2)
> ρ( 1,   1)
 1  39.84931337   0.05477226
> ρ( 2,   2)
 1  32.86419753   0.04472136
> ρ( 3,   3)
 1  22.00000000   0.03162278
```

(c) The following panel gives the modified Daniell spectral window with $m = 5$ for the first graph, convolution of $m = 5$ and $m = 7$ for the second graph, and convolution of three windows with m's of 5, 7, and 11 for the third graph.

```
>   o  =o
> o  (     =(1,3))
> o    (    ("         ", (5)),  =2,    ='',    =(0,.11),
+    =  o          (( )))
>      ( =0)
> o    (    ("         ", (5,7)),  =2,    ='',    =(0,.11),
+    =  o          (( )))
>      ( =0)
> o    (    ("         ", (5,7,11)),  =2,    ='',    =(0,.11),
+    =  o          (( )))
>      ( =0)
> o  =   o
```

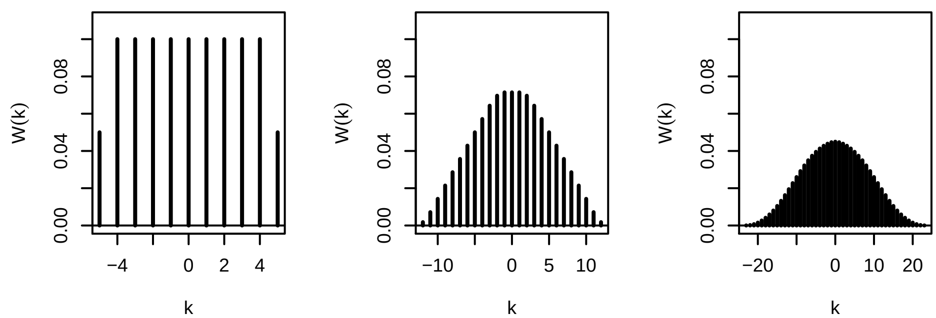


Figure 2: The Modified Daniell Spectral Window and Its Convolutions

(d) The bandwidths and degrees of freedom for each of the spectral windows in part (c) are given below. $n = 100$ is used.

```
>   4=      ("         ", (5))
>   5=      ("         ", (5,7))
>   6=      ("         ", (5,7,11))
>   4=2/(    (( 4$   ) 2)*2- 4$    1  2)
> 4=(0:5) 2
>    4=1/100*     (    (4*( 4$   ))*2)
>   5=2/(    (( 5$   ) 2)*2- 5$    1  2)
> 5=(0:12) 2
>    5=1/100*     (    (5*( 5$   ))*2)
>   6=2/(    (( 6$   ) 2)*2- 6$    1  2)
> 6=(0:23) 2
```

```
>   6=1/100*    (   (ρ6*( 6$ρ  ))*2)
> ρ( 4,  4)
 1  21.05263158  0.02915476
> ρ( 5,  5)
 1  36.97241  0.05000
> ρ( 6,  6)
 1  59.53696211  0.08093207
```

Exercise 14.3

(a) For Daniell rectangular spectral window, $W_m(k) = 1/(2m+1)$, $k = -m, \ldots, m$. So

$$\frac{1}{n^2}\sum_{k=-m}^{m} k^2 W_m(k) = \frac{2}{n^2(2m+1)}\sum_{k=1}^{m} k^2.$$

Since

$$\sum_{k=1}^{m} k^2 = \frac{m(m+1)(2m+1)}{6}, \quad \forall m = 1, 2, \ldots,$$

we have

$$\begin{aligned}
&\frac{1}{n^2}\sum_{k=-m}^{m} k^2 W_m(k) \\
&= \frac{2}{n^2(2m+1)}\frac{m(m+1)(2m+1)}{6} \\
&= \frac{2}{n^2(2m+1)}\left(\frac{m^3}{3} + \frac{m^2}{2} + \frac{m}{6}\right).
\end{aligned}$$

(b) If $m = c\sqrt{n}$ for some constant c, then the right-hand-side of the expression in part (a) becomes

$$\begin{aligned}
&\frac{2}{n^2(2m+1)}\left(\frac{m^3}{3} + \frac{m^2}{2} + \frac{m}{6}\right) \\
&= \frac{2}{\sqrt{n}\left(2c\sqrt{n}+1\right)}\left(\frac{c^3}{3} + \frac{c^2}{2\sqrt{n}} + \frac{c}{6n}\right) \longrightarrow 0, \quad \text{as } n \longrightarrow \infty.
\end{aligned}$$

(c) If $m = c\sqrt{n}$ for some constant c, then the approximate variance of the smoothed spectral density given by the right-hand-side of Equation (14.2.4) becomes

$$S^2(f)\sum_{k=-m}^{m} W_m^2(k) = \frac{S^2(f)}{2c\sqrt{n}+1} \longrightarrow 0, \quad \text{as } n \longrightarrow \infty.$$

Exercise 14.4

Suppose $\overline{S}(f) \approx c\chi_v^2$, where c is a constant and χ_v^2 is a chi-square variable with degrees of freedom v. By the approximate unbiasedness of $\overline{S}(f)$, we have

$$\mathrm{E}\left(\overline{S}(f)\right) \approx cv \approx S(f). \tag{2}$$

From the approximate variance of $\overline{S}(f)$, we get

$$Var\left(\overline{S}(f)\right) \approx 2vc^2 \approx S^2(f) \sum_{k=-m}^{m} W_m^2(k). \tag{3}$$

Then from Equations (2) and (3), we estimate the values of c and v as

$$c = \frac{S(f)\sum_{k=-m}^{m} W_m^2(k)}{2}, v = \frac{2}{\sum_{k=-m}^{m} W_m^2(k)}.$$

Exercise 14.5

The periodogram of the series $\{Y_t\}$ is given by

```
>  =1:48
>  =  (2*o *0.28* )
>       'o (      =4.875,      =2.5,o        =8)
>   o (  ,   ='   ',  ' ='       ' ',  ' ='       ',   ='  ',   =2,   ='')
> '     ( =0)
```

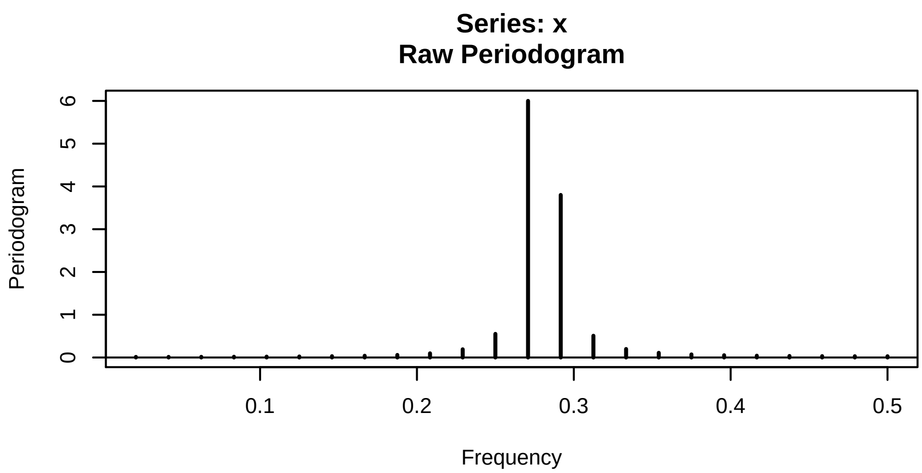


Figure 3: Periodogram of $Y_t = \sin[2\pi(0.28)t]$, $t = 1, \ldots, 48$.

Note that $f = 0.28$ is not a Fourier frequency. So the peak at $f = 0.28$ does not appear. Instead, the power at this frequency is blurred across several nearby frequencies giving the appearance of a much wider peak.

Exercise 14.6

Using the modified Daniell spectral window with the span of 11, we estimate the spectrum of the logarithms of the raw rainfall values.

```
> [illegible]( [illegible] )
> [illegible] =4.875, [illegible] =2.5, [illegible] =8)
> [illegible]( [illegible] ), [illegible] = [illegible](11), [illegible] = [illegible](.08,0.5), [illegible] ='', [illegible] =' [illegible]( [illegible] )',
+ [illegible] =' [illegible] ')
```

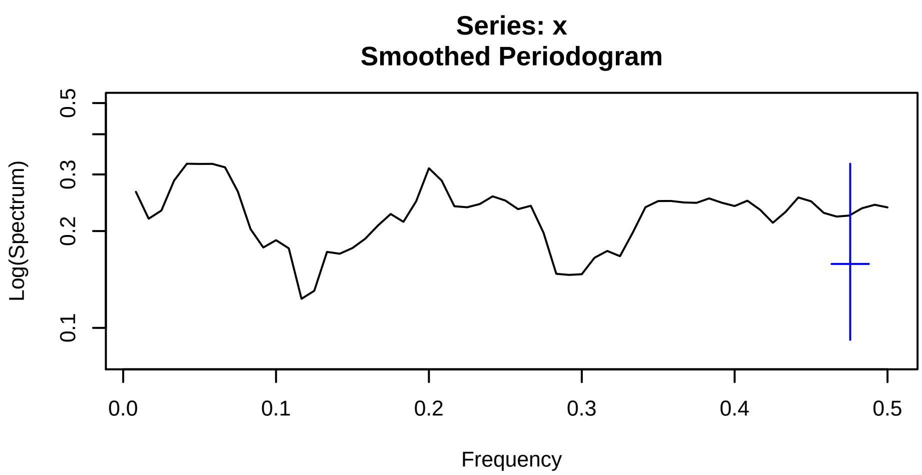


Figure 4: Log(Spectrum) of Log(Larain)

Exercise 14.7

(a) The time series plot of [illegible] 1 is given below. This time series seems stationary.

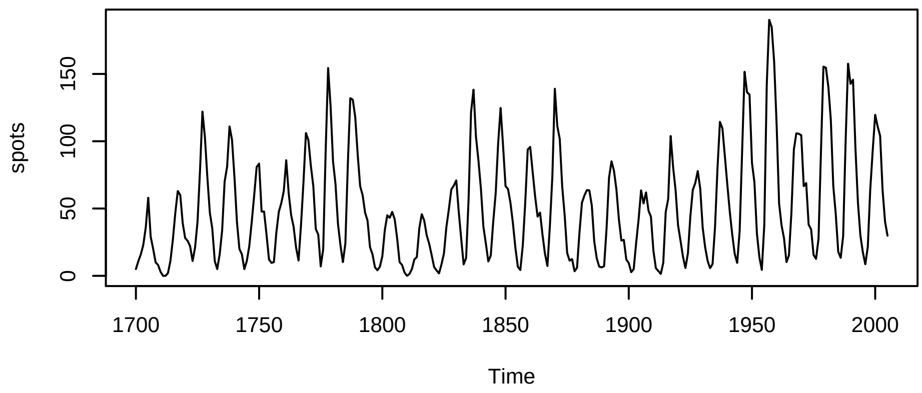


Figure 5: Sunspot Numbers Time Series

(b) The estimated spectrum using a modified Daniell spectral window convoluted with itself and a span of 3 for both is given below. The peak at around $f = 0.09$ is significant suggesting that the number of sunspots fluctuates with a cycle period of about $1/f = 11$ years. This accords with the observation from the data plot in part (a).

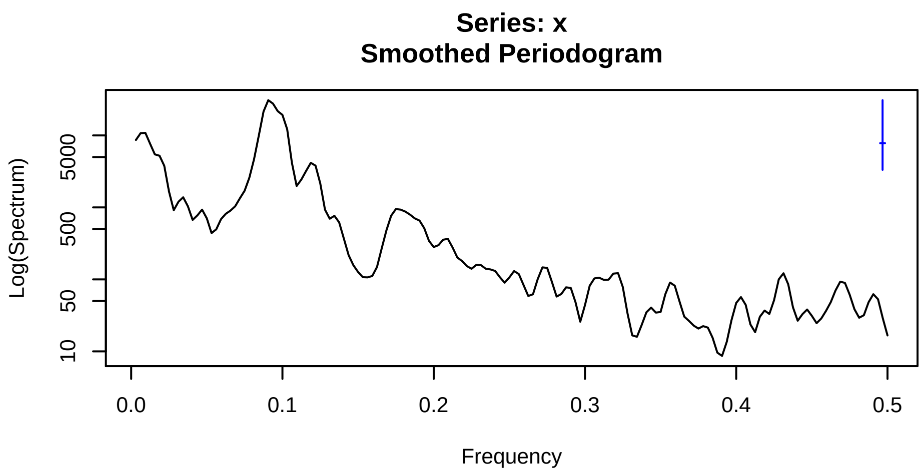


Figure 6: Estimated Spectrum for the Sunspot Number

(c) Then let us estimate the spectrum using an AR model with the order chosen to minimize the AIC. The chosen AR order is 9. The graph below suggests a peak at around $f = 0.095$. This suggests a cycle of about $1/f = 10.53$ years in the fluctuation of the annual sunspot number.

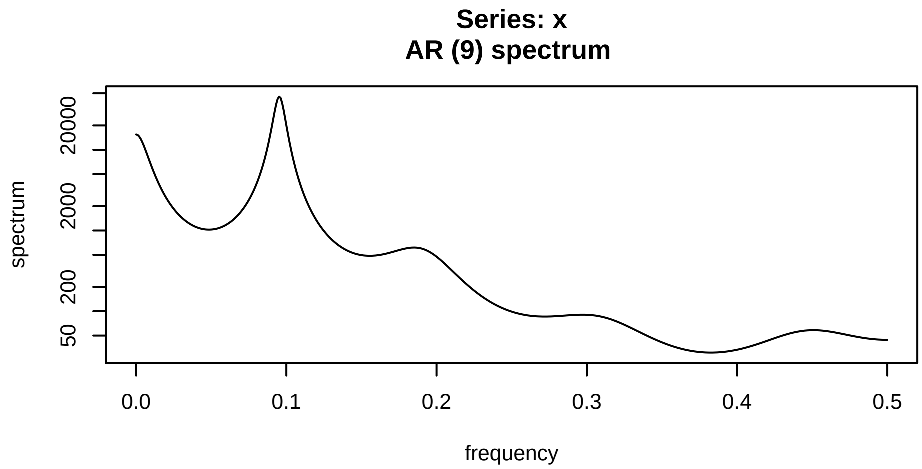


Figure 7: Estimated Spectrum for the Sunspot Number by AR Method

(d) Overlay the estimates obtained in parts (b) and (c) above onto one plot. They seem to to agree to a reasonable degree.

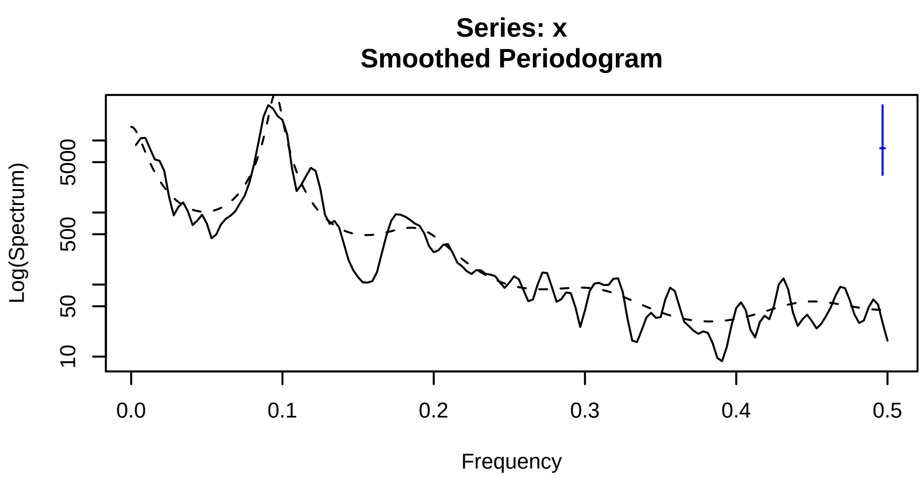


Figure 8: Log(Spectrum) of Sunspot Number

Exercise 14.8

(a) The following 3 figures are the modified Daniell spectral windows of using *span=3,5*, and 7.

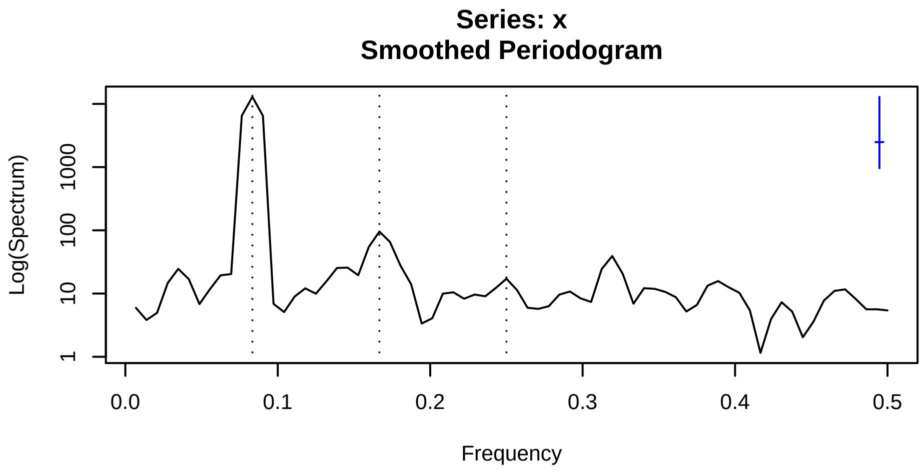


Figure 9: Modified Daniell Spectral Window with $Span = 3$ for

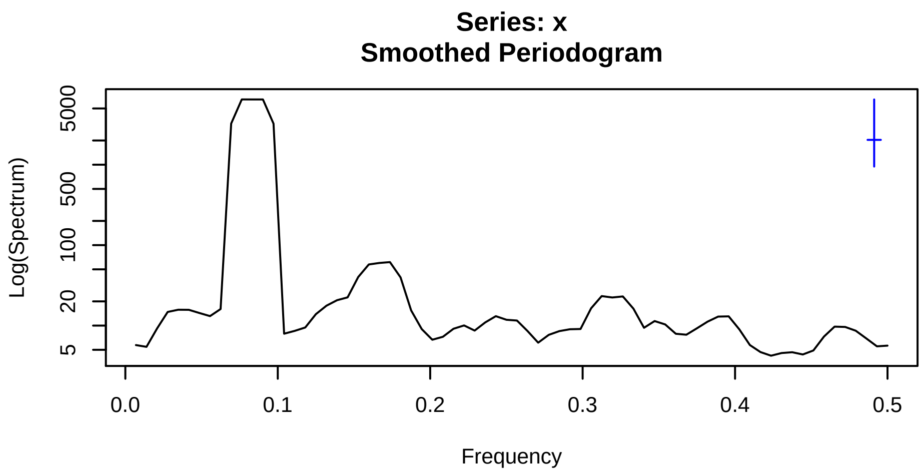


Figure 10: Modified Daniell Spectral Window with $Span = 5$ for

(b) The estimate with $span = 3$ among the three estimates in part (a) best represents the spectrum of the process. With the smallest bandwidth, it shows clearly the prominent peak at $f = 1/12$ representing the strong annual seasonality and some secondary peaks at about $f = 2/12, 3/12$, representing higher harmonics of the annual frequency. On the contrary, with bigger bandwidth, the peaks at those frequencies in the windows with $span = 5$ or 7 are flatted. Given the length of the 95% confidence interval shown in Figure 9, we can conclude that the peaks at around $f = 1/12, 2/12$, and $3/12$ are probably real.

Exercise 14.9

(a) The plot for time series is given below. It seems that this time series is stationary.

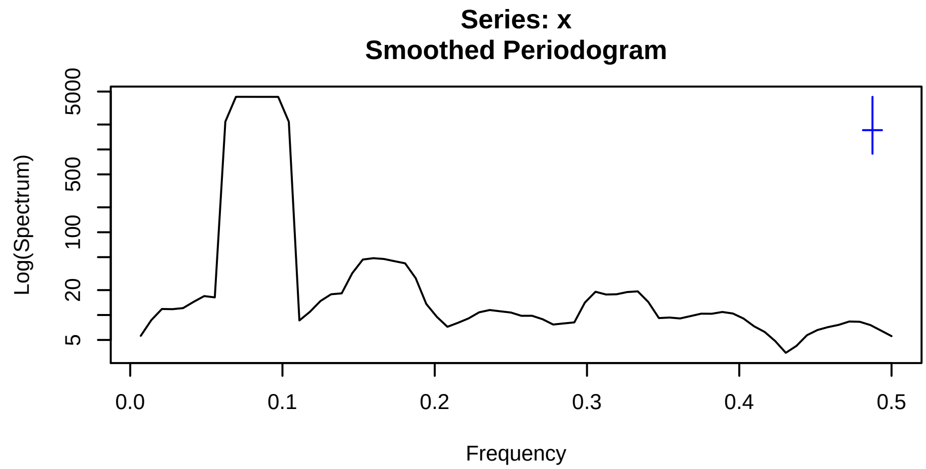


Figure 11: Modified Daniell Spectral Window with $Span = 7$ for

```
>      ‘  (    )
>    ‘(      )
> o  ()
>        ‘o (     =4.875,        =2.5,o        =8)
> o     (    ,  o ="  ")
```

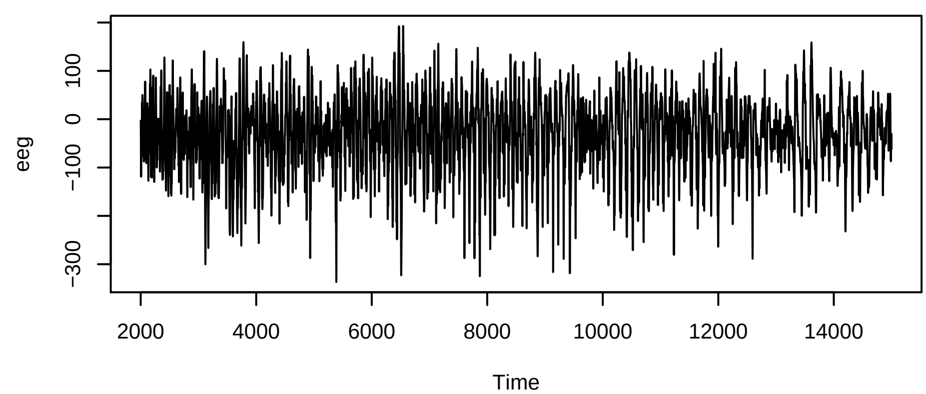


Figure 12: Plot for EEG Time Series

(b) By using a modified Daniell spectral window convolved with itself and an m of 50 for both components of the convolution, we give the sample spectral density plot for . The sample spectral density has a unique and strong peak at a small frequency, which suggests that an AR model may well fit it.

```
>  =        ("                 ",   (50,50))
>           (       =4.875,        =2.5,        =8)
>  1=    (    ,       =  ,    ='',    ='',
+     ='        ',    ='   (                )')
```

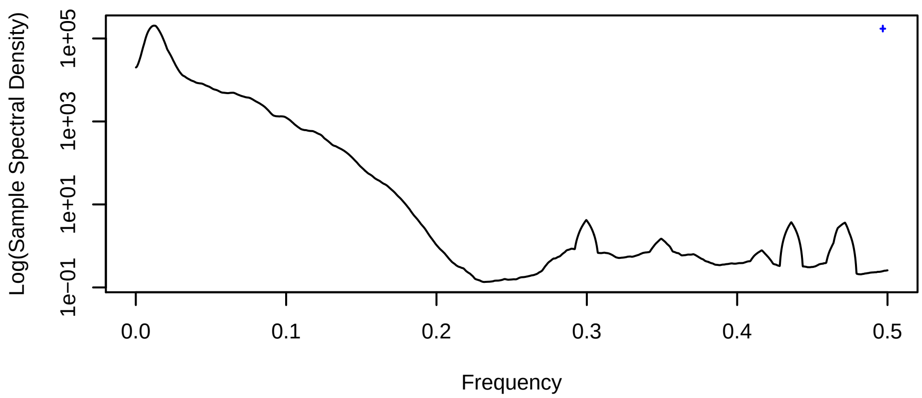


Figure 13: Sample Spectral Density for EEG

(c) Using an AR model with order chosen to minimize the AIC, we obtain the sample spectral estimation given below. The chosen best order for AR model is 41.

```
>  2=    (    ,      ='  ',    ='',    ='   ',    ='        ',
+     ='   (                 )')
>  2$
 1  "  (41)    "
```

(d) Overlay the above two sample spectral estimates onto one plot below. We can see that the non-parametric smoothing estimation agrees with the estimation from an fitted AR(41) model very well.

```
>    (    ,    = (51,51),   ='',
+     ='        ',    ='   (                )')
>  2=    (    ,      ='  ',    ='',    = )
>        ( 2$    , 2$    ,    ='    ')
```

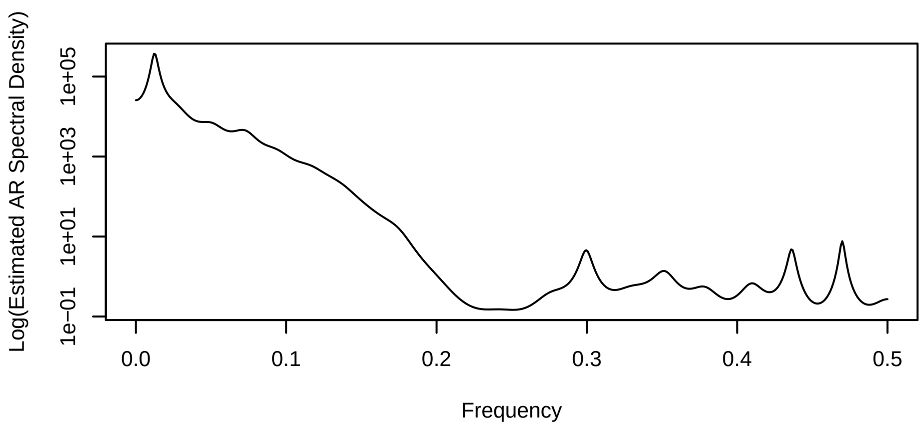


Figure 14: AR Spectral Estimation for EEG

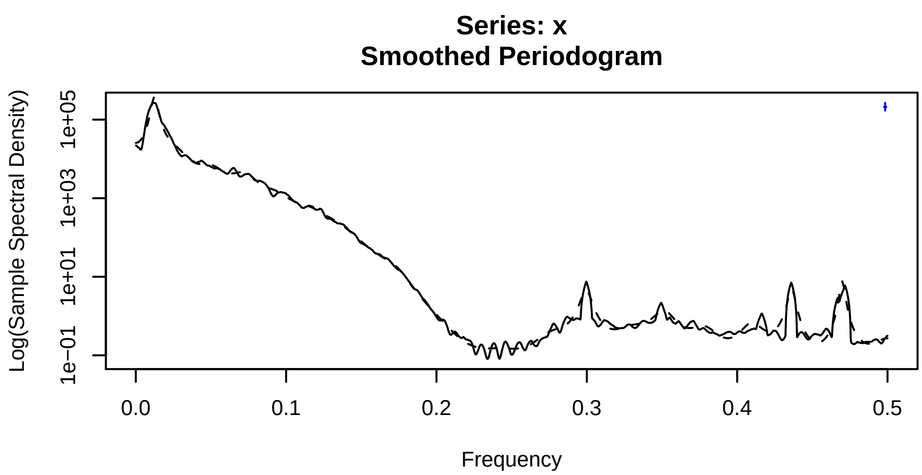


Figure 15: Estimated Spectral Densities

Exercise 14.10

(a) The time series plot of the first difference of the logarithms of the electricity values is given below. This process seems stationary.
(b) The following 3 figures display the smoothed spectrum of the first difference of the logarithms using a modified Daniell spectral window and *span* $= 25, 13$, and 7. As the *span* value becomes smaller, the bandwidth becomes narrower. So in Figure 19, we clearly see 6 peaks at around these frequencies: $0.09, 0.18, 0.25, 0.34, 0.41$ and 0.5. However, possible leakage appears at these peaks.
(c) Now let us use a spectral window that is a convolution of two modified Daniell windows each with *span* $= 3$. Also use a 10% taper. By tapering and using a smaller span value, the 6 peaks are more clear: $f_1 = 0.0825 = 1/12, f_2 = 0.1675 = 2/12, f_3 = 0.25 = 3/12, f_4 =$

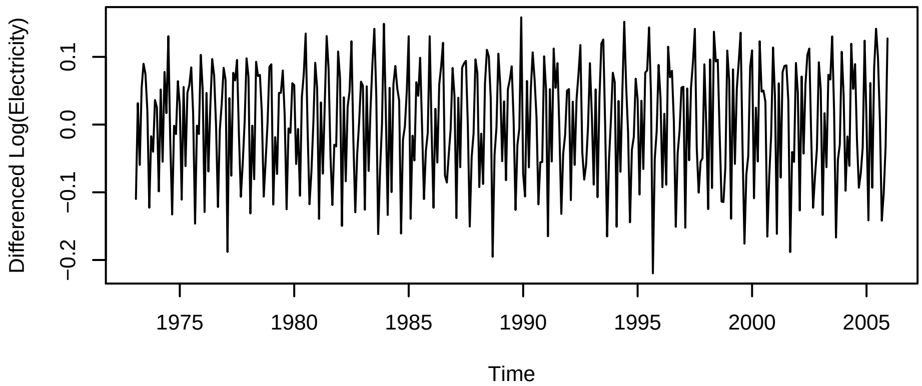


Figure 16: The First Difference of the Log(Electricity)

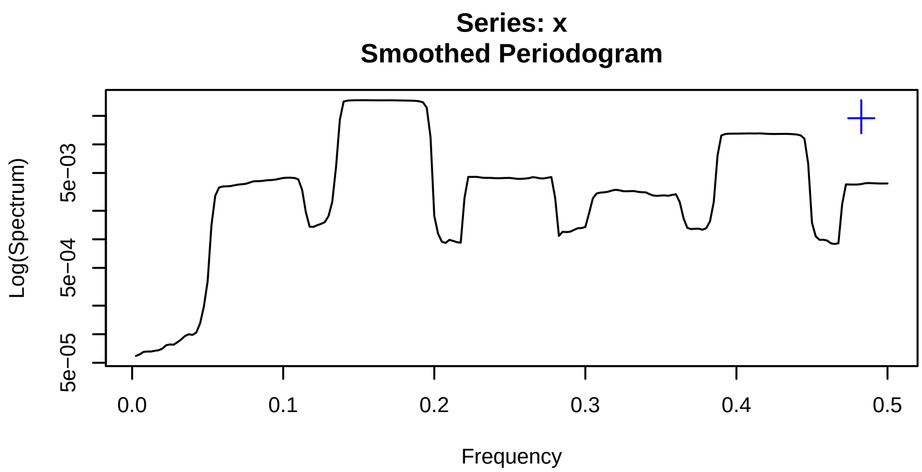


Figure 17: Smoothed Log(Spectrum) in a Modified Daniell Window with $Span = 25$

$0.3325 = 4/12$, $f_5 = 0.4175 = 5/12$, $f_6 = 0.5 = 6/12$.

(d) Using an AR model with order chosen to minimize the AIC, we obtain the sample spectral estimation given below. The chosen best order for AR model is 25. This spectrum shows peaks at the same frequencies as those in part (c), representing multiples of the fundamental frequency of $1/12$.

(e) Now overlay the estimates obtained in parts (c) and (d) above onto one plot. We can see that the non-parametric smoothing estimation agrees the estimation from an fitted AR(25) model well.

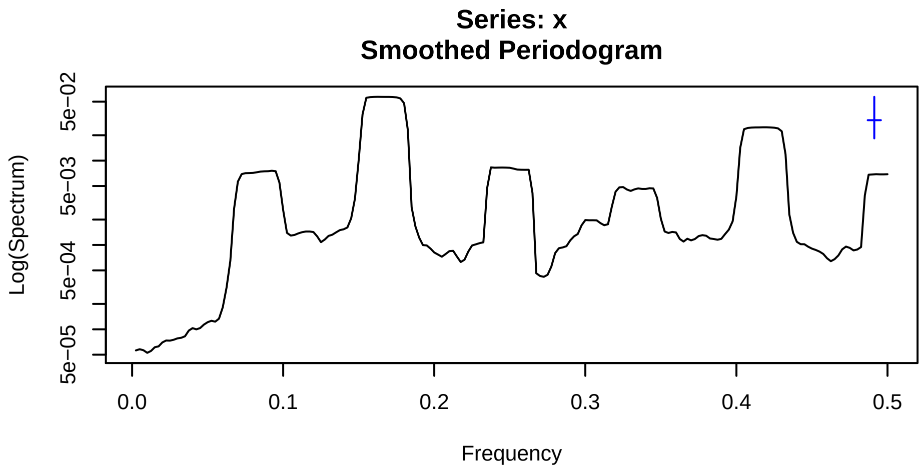


Figure 18: Smoothed Log(Spectrum) in a Modified Daniell Window with $Span = 13$

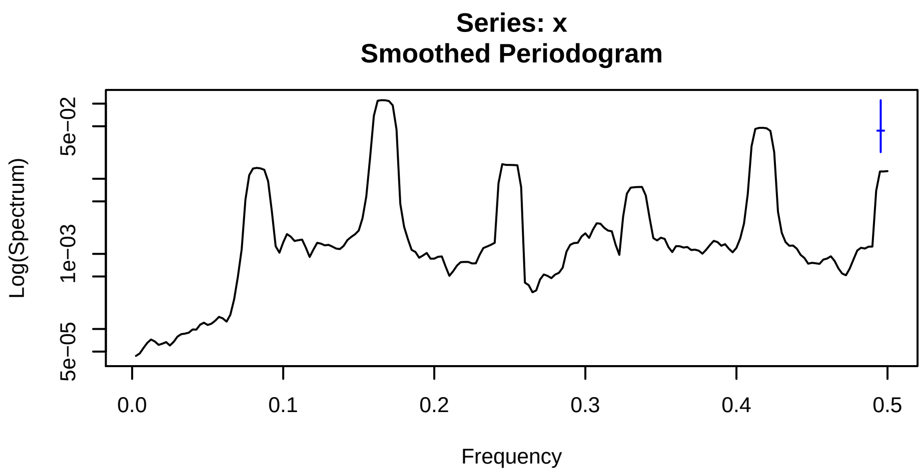


Figure 19: Smoothed Log(Spectrum) in a Modified Daniell Window with $Span = 7$

Exercise 14.11

(a) Let us estimate the spectrum using a spectral window that is a convolution of two modified Daniell windows each with $span = 7$. Compared with Exhibit (14.24) in the book, the bandwidth is bigger in this window and the peaks are wider. But the peak frequencies are still clearly seen as multiples of 1/12.
(b) Now estimate the spectrum using a single modified Daniell spectral window with $span = 7$. Compared with those results in part (a) and Exhibit (14.24), the peaks here are flat. So we can not figure out at what frequencies the peaks exactly are.
(c) Then estimate the spectrum using a single modified Daniell spectral window with $span = 11$. Compared with the results in part (b), the peaks here are even flatter. Again we have no idea at what frequencies have a peak.

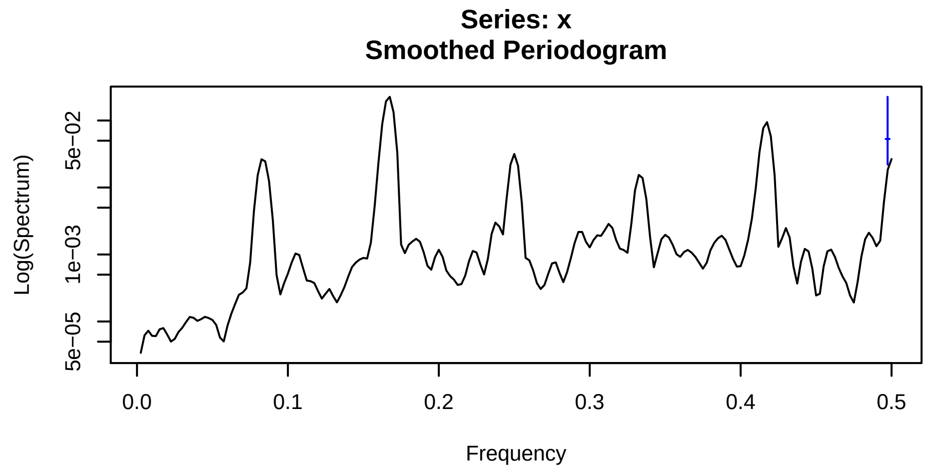


Figure 20: Smoothed Log(Spectrum) in a Convoluted Modified Daniell Window Each with $Span = 3$ and Taper= 0.1

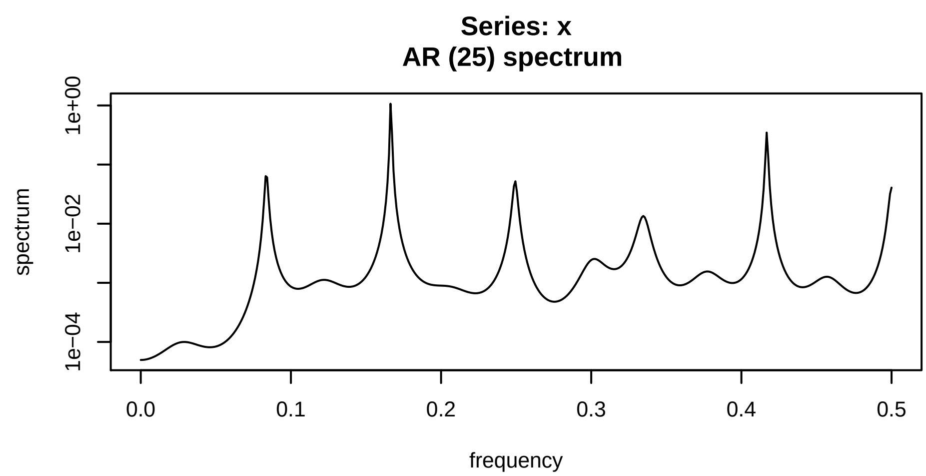


Figure 21: Estimated Spectrum for the Differenced Log(Electricity) by AR Method

(d) Among the 4 different estimates, I prefer the estimate shown in Exhibit (14.24) because the other 3 have more smoothing than the convolution of two modified Daniell windows each with $span = 3$.

Exercise 14.12

(a) Let us estimate the spectrum using $span = 25$ with the modified Daniell spectral window.

```
> ‘ ‘ )
> ( ), ‘ =(25), ‘ =’’, =’’, =(.02,13),
+ ‘ = ( ) ’ , ‘ =’ ’)
```

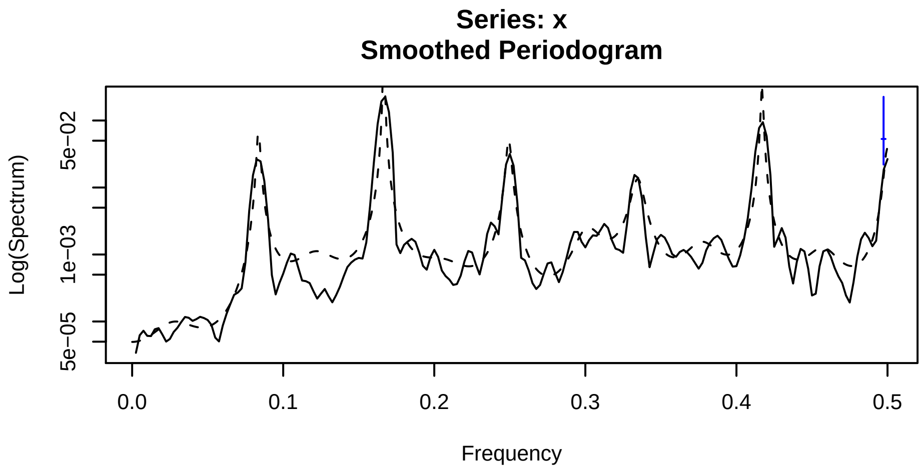


Figure 22: Estimated Spectral Densities for Differenced Log(Electricity)

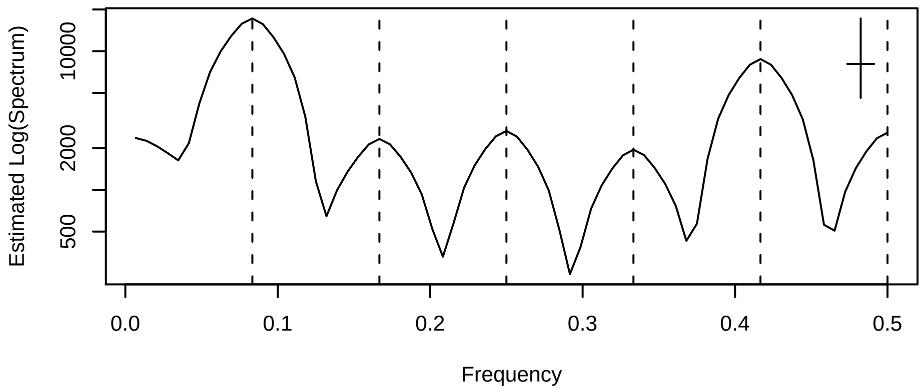


Figure 23: Estimated Spectrum for Milk Production in a Convoluted Modified Daniell Window with $Span = 7$

The bandwidth is 0.012, which is much bigger than the bandwidth (0.0044) in a convolution of two windows each with $m = 7$. The bigger bandwidth implies a window smoothed too much to see any peaks representing the annual seasonality which can be seen from Exhibit 14.23.

(b) Then let us estimate the spectrum using $span = 13$ with the modified Daniell spectral window.

```
>    (      ), o   = (13),  ='',  ='',  = (.02,13),
+      =   (        )'  ='        ')
```

The bandwidth is 0.006, which is bigger than that in Exhibit 14.23 and smaller than that in part (a). So the estimated spetrum here is rougher than that in Exhibit 14.23 but

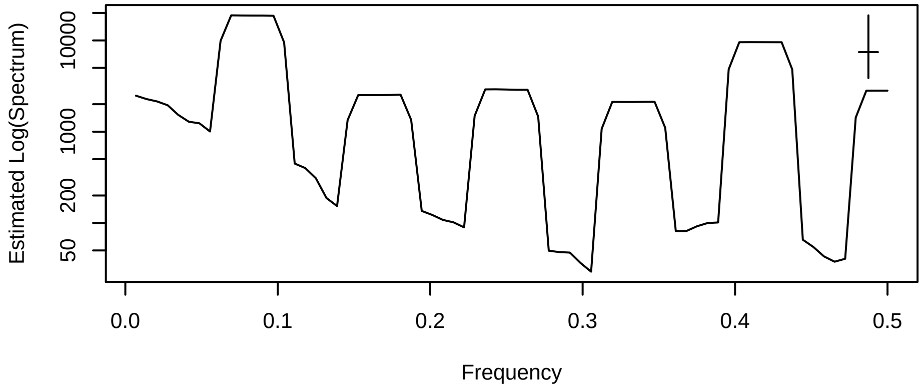


Figure 24: Estimated Spectrum for Milk Production in a Modified Daniell Window with $Span = 7$

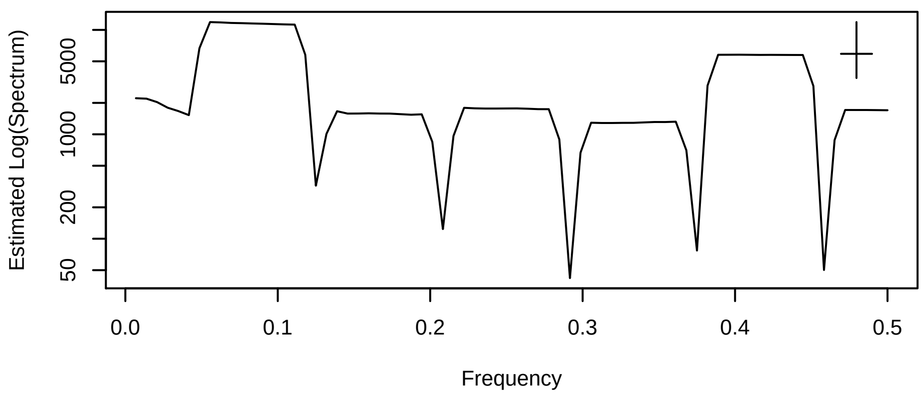


Figure 25: Estimated Spectrum for Milk Production in a Modified Daniell Window with $Span = 11$

smoother than that in part (a). The peaks at frequencies $f = 1/12, 2/12$ and $3/12$ become somewhat clear.

Exercise 14.13

(a) The plot of the first 400 of the [illegible] time series is given below. Compared with the plots of the trombone and euphonium in Exhibit (14.25), the tuba time series has longer cycles.

```
> [illegible]'o ([illegible] =4.875, [illegible] =2.5,o [illegible] =8)
> [illegible]( [illegible]')
```

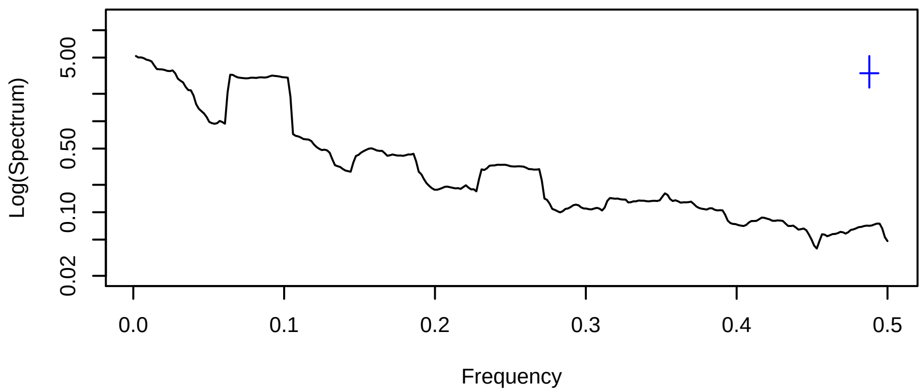


Figure 26: Log(Spectrum) of Log(Flow) with $Span = 25$.

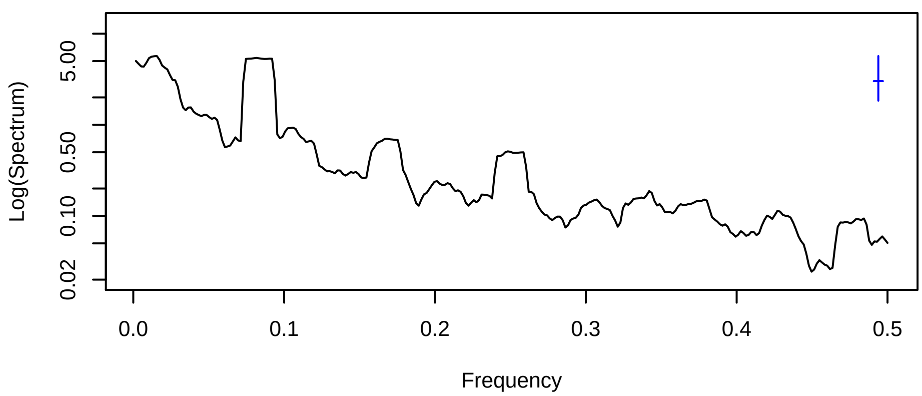


Figure 27: Log(Spectrum) of Log(Flow) with $Span = 13$.

```
> o    (      (   ',    =400),   '  =' '      ', ' o= (-1,+1,2))
```

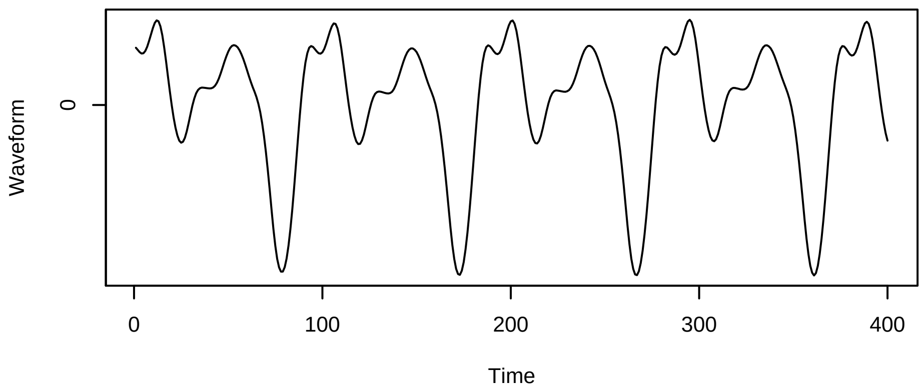


Figure 28: Tuba

(b) Let us estimate the spectrum of the tuba time series using a convolution of two modified Daniell spectral window each with span $m = 11$.

```
> [illegible]([illegible]([illegible]=400), [illegible]=[illegible](11,11), [illegible]='',
+ [illegible]='[illegible]', [illegible]='[illegible]')
```

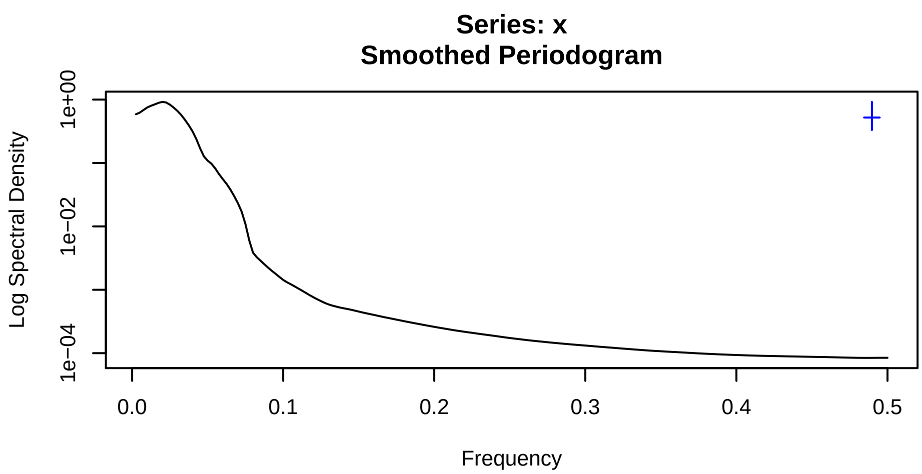


Figure 29: Estimated Spectram

(c) Compared with the estimated spectra of the trombone and euphonium shown in Exhibit (14.24), the estimated spectrum of the tuba time series obtained in part (b) is lower and much smoother.

(d) It's not hard to see that the higher frequency components of the spectrum for the tuba look more like those of the euphonium than the trombone. This reveals one reason why the euphonium is sometimes called a tenor tuba.

Chapter 15

Exercise 15.1

We fit the TAR$(2;1,4)$ model with delay $d=2$ to the logarithms of the predator series in the data file `veilleux`.

```
> library(TSA)
> data(veilleux)
> predator=veilleux[,1]
> predator.eq=window(predator,start=c(7,1))
> predator.tar.3=tar(y=log(predator.eq),p1=4,p2=4,d=2,a=.1,b=.9,print=T)
```

The estimation for this model is given below.

	Estimate	**Std. Error**	**t-statistic**	**p-value**
$\hat{d}$	2			
$\hat{r}$	4.048			
		Lower Regime (14)		
$\hat{\phi}_{1,0}$	0.95	0.79	1.21	0.25
$\hat{\phi}_{1,1}$	0.82	0.20	4.08	0.00
$\tilde{\sigma}_1^2$	0.0598			
		Upper Regime (39)		
$\hat{\phi}_{2,0}$	4.06	0.57	7.10	0.00
$\hat{\phi}_{2,1}$	0.91	0.14	6.37	0.00
$\hat{\phi}_{2,2}$	-0.26	0.21	-1.24	0.22
$\hat{\phi}_{2,3}$	-0.20	0.20	-0.98	0.33
$\hat{\phi}_{2,4}$	-0.32	0.15	-2.14	0.04
$\tilde{\sigma}_2^2$	0.0638			

From the above table we can see that the fitted TAR model with delay $d=2$ has a smaller threshold estimate $\hat{r}=4.048$ than that in the fitted TAR$(2;1,4)$ model with delay $d=3$ ($\hat{r}=4.661$). So only 14 data cases fall in lower regime, while 39 data cases fall in upper regime. It's not as balanced as the situation reported in the text book, where $n_1=30$ and $n_2=23$. In addition, the estimated noise variances in the fitted TAR model, 0.0598 and 0.0638, are larger than those in the TAR model with delay 3, 0.0548 and 0.0560. Also, let us compare the skeletons of the two models. We know that the skeleton of TAR$(2;1,4)$ model with delay 3 converges to a limit cycle. The graph (below) of the skeleton of TAR$(2; 1, 4)$ model with delay 2 tells us it also converges to a limit cycle.

```
> tar.skeleton(predator.tar.3)
```

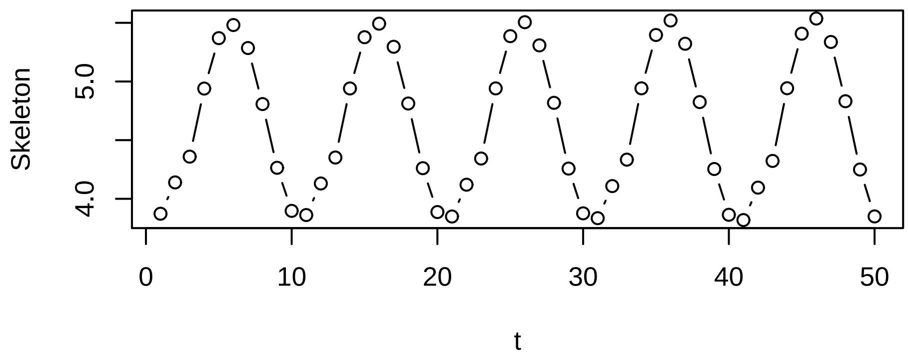


Figure 1: Skeleton of model TAR(2;1,4)

Exercise 15.2

Let us set the maximum order $p = 5$ and use the MAIC estimation method to estimate the parameters.

```
> AICM=NULL
> for(d in 1:5) {spots.tar=tar(y=sqrt(spots),p1=5,p2=5,d=d,a=.1,b=.9)
+ AICM=rbind(AICM, c(d,spots.tar$AIC,signif(spots.tar$thd,5), +
spots.tar$p1,spots.tar$p2)) + }
> colnames(AICM)=c('d','nominal AIC','r','p1','p2')
> rownames(AICM)=NULL
> AICM
```

d	AIC	$\hat{r}$	$\hat{p}_1$	$\hat{p}_2$
1	149.9	5.882	5	5
2	110.5	6.058	3	5
3	124.6	6.596	2	5
4	126.2	8.044	3	5
5	150.5	8.155	4	5

From the above table we find that the TAR$(2; 3, 5)$ with delay $d = 2$ has the smallest AIC. Next we fit this model to the `spots` data.

```
> spots.tar.1=tar(y=sqrt(spots),p1=5,p2=5,d=2,a=.1,b=.9,print=T)
```

	Estimate	Std. Error	t-statistic	p-value
$\hat{d}$	2			
$\hat{r}$	6.058			
		Lower Regime (20)		
$\hat{\phi}_{1,0}$	9.52	1.28	7.44	0.00
$\hat{\phi}_{1,1}$	1.05	0.17	6.17	0.00
$\hat{\phi}_{1,2}$	-1.20	0.29	-4.13	0.00
$\hat{\phi}_{1,3}$	-0.56	0.24	-2.28	0.04
$\tilde{\sigma}_1^2$	0.8211			
		Upper Regime (37)		
$\hat{\phi}_{2,0}$	5.69	0.86	6.60	0.00
$\hat{\phi}_{2,1}$	0.37	0.11	3.47	0.00
$\hat{\phi}_{2,2}$	0.38	0.12	3.25	0.00
$\hat{\phi}_{2,3}$	-0.07	0.11	-0.59	0.56
$\hat{\phi}_{2,4}$	-0.36	0.11	-3.34	0.00
$\hat{\phi}_{2,5}$	-0.11	0.07	-1.50	0.14
$\tilde{\sigma}_2^2$	0.2141			

Notice that the estimate $\hat{\phi}_{2,5}$ with p-value$= 0.14$ is not significant. It guides us to fit the model TAR$(2; 3, 4)$ with delay $d = 2$.

```
> spots.tar.2=tar(y=sqrt(spots),p1=4,p2=4,d=2,a=.1,b=.9,print=T)
```

The fitted TAR$(2; 3, 4)$ model of the lower regime implies that

$$Y_t^{\frac{1}{2}} = 9.52 + 1.05Y_{t-1}^{\frac{1}{2}} - 1.20Y_{t-2}^{\frac{1}{2}} - 0.56Y_{t-3}^{\frac{1}{2}} + 0.91e_t,$$

when $Y_{t-2} \leq 6.058^2 = 36.70$, while the fitted model of the upper regime implies that

$$Y_t^{\frac{1}{2}} = 4.81 + 0.40Y_{t-1}^{\frac{1}{2}} + 0.42Y_{t-2}^{\frac{1}{2}} - 0.01Y_{t-3}^{\frac{1}{2}} - 0.49Y_{t-4}^{\frac{1}{2}} + 0.48e_t,$$

when $Y_{t-2} \geq 36.70$.

Next let us look at the goodness of fit of the TAR$(2; 3, 4)$ model to the square root of `spots` data.

```
> win.graph(width=4.875, height=4.5,pointsize=8)
> tsdiag(spots.tar.2,gof.lag=20)
```

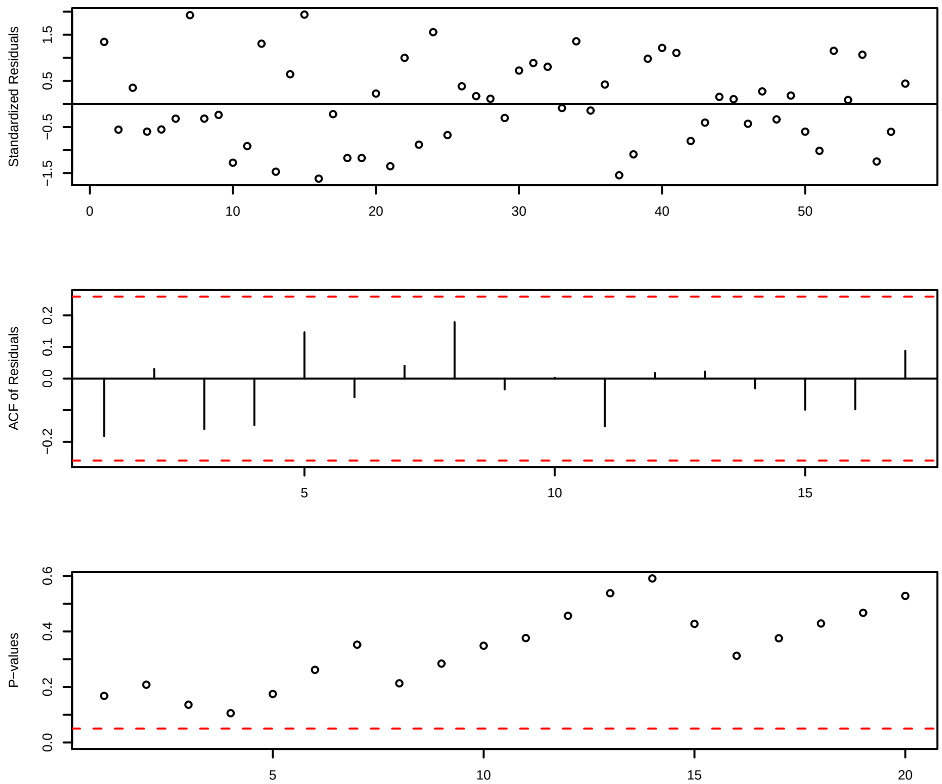


Figure 2: Model Diagnostics of TAR(2;3,4): Sopts Series

```
> win.graph(width=2.5, height=2.5,pointsize=8)
> qqnorm(spots.tar.2$std.res)
> qqline(spots.tar.2$std.res)
```

In Figure 2, the top figure shows that the standardized residuals have no particular pattern. The middle figure shows no significant autocorrelation in the residuals. In the bottom figure, all p-values are above 0.05. Figure 3 displays the QQ normal score plot of the standardized residuals, which is apparently straight and hence the errors appear to be normally distributed. In summary, the fitted TAR$(2;3,4)$ model provides a good fit to the `spots` data.

Exercise 15.3

Let us draw the prediction intervals and the predicted medians for 10 years by using TAR$(2;3,4)$ with delay 2.

```
> pred.sopts=predict(spots.tar.2,n.ahead=10,n.sim=1000)
```

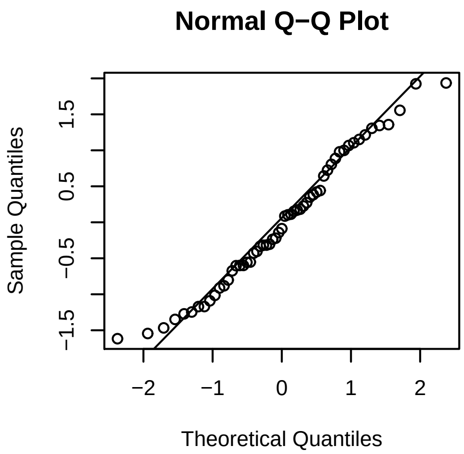


Figure 3: QQ Normal Plot of the Standardized Residuals

```
> yy=ts(c(sqrt(spots),pred.sopts$fit),frequency=1,start=start(spots))
> plot(yy,type='n',ylim=range(c(yy,pred.sopts$pred.interval)),
+ ylab='Sqrt Spots', xlab=expression(t))
> lines(sqrt(spots))
> lines(window(yy, start=end(spots)+c(0,1)),lty=2)
> lines(ts(pred.sopts$pred.interval[2,],start=end(spots)+c(0,1),
+ freq=1),lty=2)
> lines(ts(pred.sopts$pred.interval[1,],start=end(spots)+c(0,1),
+ freq=1),lty=2)
```

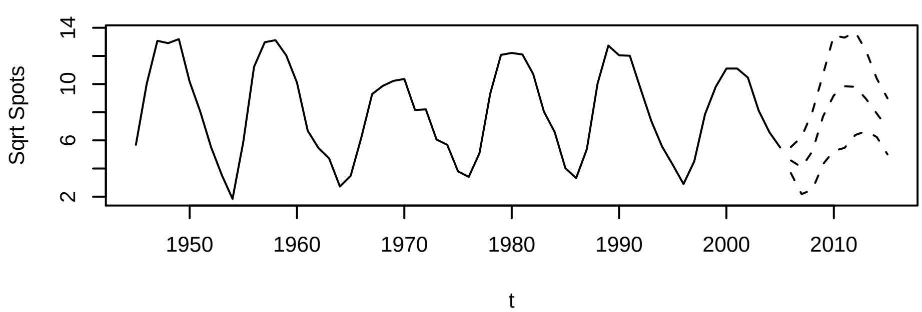


Figure 4: Prediction of the Sqrt of Spots

The middle dashed line in Figure 4 is the median of the predictive distribution and the other dashed lines are the $2.5th$ and $97.5th$ percentiles of the predictive distribution.

Exercise 15.4

The long-run behavior of the skeleton of the fitted TAR$(2;3,4)$ model for the relative sunspot data is given below. The fitted model is stationary and its skeleton converges to a limit point.

```
> tar.skeleton(spots.tar.2)
```

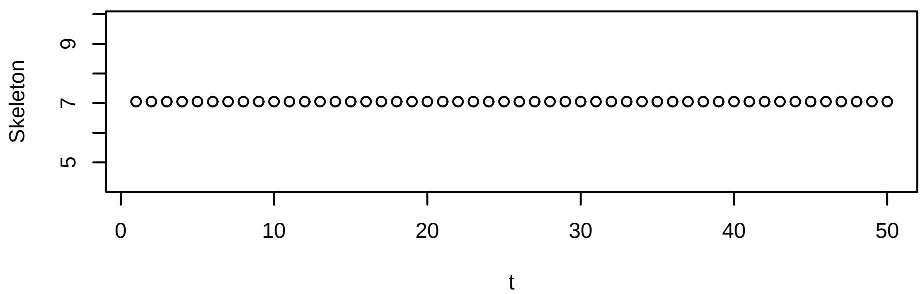


Figure 5: Skeleton of the Fitted Model for the Spots Data

Exercise 15.5

The graph below shows the simulated series with size 1000 from the fitted model TAR$(2;3,4)$ with delay $d=2$.

```
> set.seed(356813)
> plot(y=tar.sim(n=1000,object=spots.tar.1)$y,x=1:1000,ylab=expression(Y[t]),
+ xlab=expression(t),type='o')
```

To compare the spectrum of the simulated realization with that of the data, let us see the graph below. The spectrum of the simulated series fits the sunspots data.

```
> set.seed(2357125)
> yy.1=tar.sim(spots.tar.2,n=1000)$y
> spec.1=spec(yy.1,taper=.1, method='ar',plot=F)
> spec.spots=spec(sqrt(spots),taper=.1, span=c(3,3),plot=F)
> spec.spots=spec(sqrt(spots),taper=.1, span=c(3,3),
+ ylim=range(c(spec.1$spec,spec.spots$spec)),sub='')
> lines(y=spec.1$spec,x=spec.1$freq,lty=2)
```

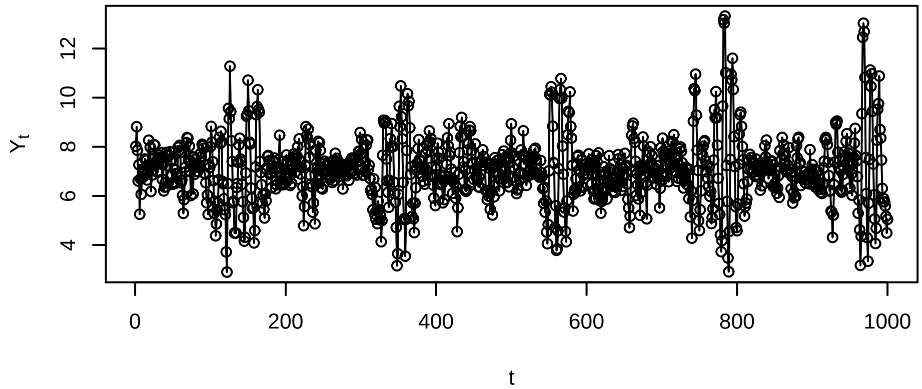


Figure 6: Simulated Series of the Fitted Model ($n = 1000$)

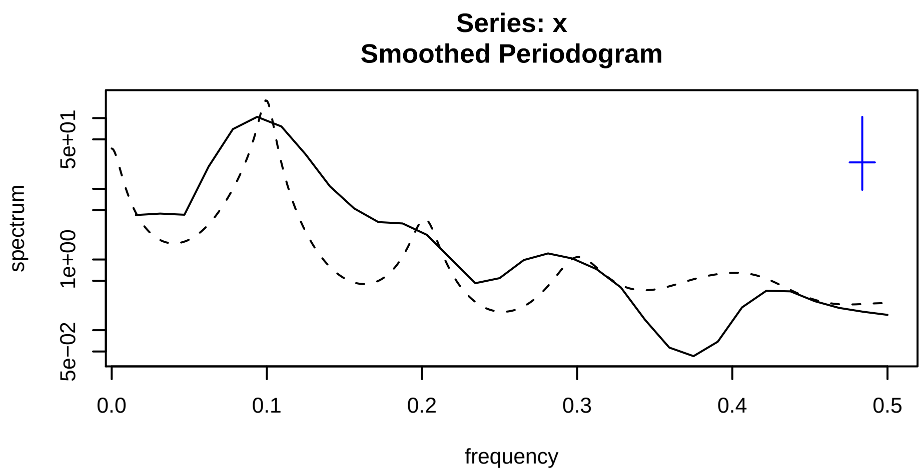


Figure 7: Spectra of Simulated Series and Sqrt Transformed Data

Exercise 15.6

The lagged regression plots for the square-root transformed `hare` series is given below. We can see from Figure 8 that the regression function estimates appear to be strongly nonlinear for lags 2, 3 and 6, suggesting a nonlinear data mechanism.

```
> data(hare)
> win.graph(width=4.875, height=6.5,pointsize=8)
> set.seed(2534567)
> par(mfrow=c(3,2))
> lagplot(sqrt(hare))
```

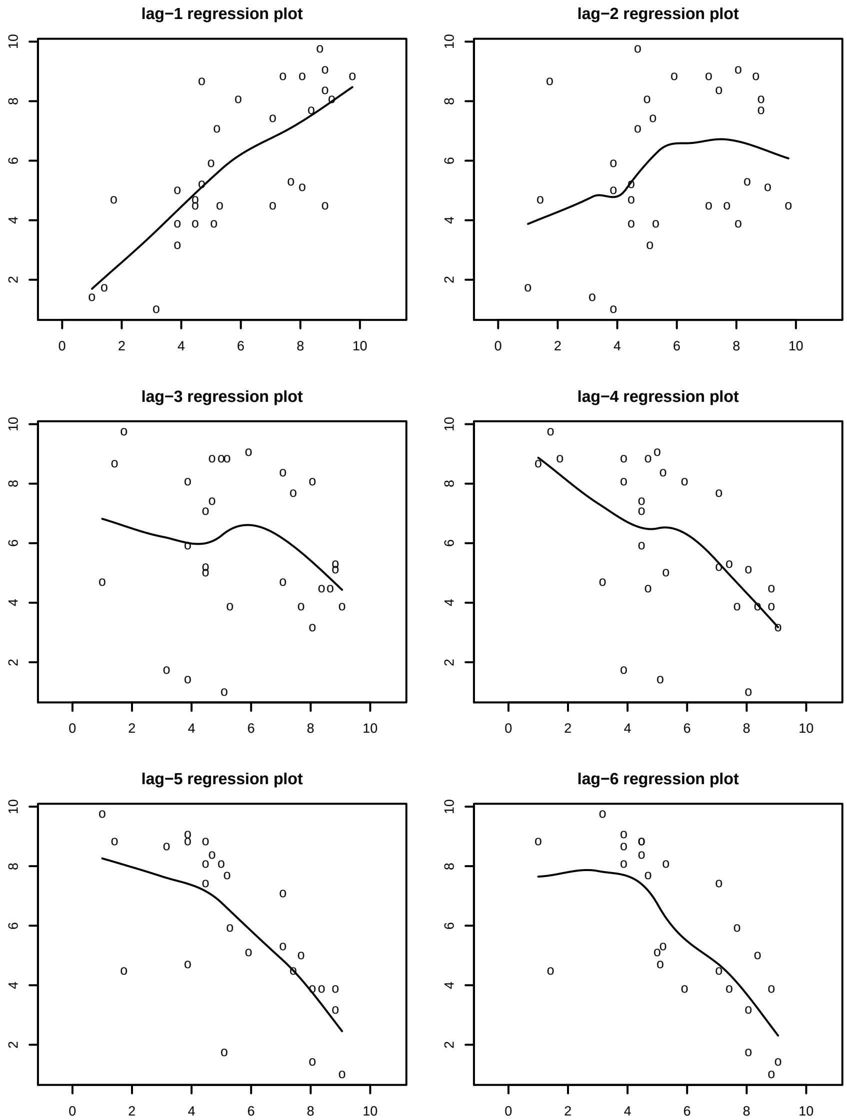


Figure 8: Lagged Regression of Sqrt Transformed Hare Data

Exercise 15.7

The results of formal tests (Keenan's test, Tsay's test and threshold likelihood ratio test) for nonlinearity for the `hare` data are given below.

```
> Keenan.test(sqrt(hare))
$test.stat [1] 8.083568
$p.value [1] 0.009207613
$order [1] 3

> Tsay.test(sqrt(hare))
$test.stat [1] 2.135
```

```
$p.value [1] 0.09923
$order [1] 3

> pvaluem=NULL
> for (d in 1:3){
+ res=tlrt(sqrt(hare),p=5,d=d,a=0.25,b=0.75) + pvaluem= cbind(
pvaluem, c(d,res$test.statistic, + res$p.value))}
> rownames(pvaluem)=c('d','test statistic','p-value')
> round(pvaluem,3)
                  [,1]   [,2]   [,3]
d                1.000  2.000  3.000
test statistic 71.558 54.964 16.807
p-value          0.000  0.000  0.083
```

The p-value of Keenan's test is 0.009, strongly suggesting the nonlinearity of square-root `hare` series. Although the p-value of Tsay's test is $0.09923 > 0.05$, it's still less than 0.10. Moreover, when $d = 1, 2$, the p-value of threshold likehood ratio test is about 0.000 and when $d = 3$ the p-value is 0.083. We conclude it from the three test results that the generating mechanism for `hare` data is nonlinear.

Exercise 15.8

From the MAIC method we set the maximum order to be $p = 3$.

```
> AICM=NULL
> for(d in 1:3) {hare.tar=tar(y=sqrt(hare),p1=3,p2=3,d=d,a=.1,b=.9)
+ AICM=rbind(AICM, c(d,hare.tar$AIC,signif(hare.tar$thd,3), +
hare.tar$p1,hare.tar$p2))}
> colnames(AICM)=c('d','nominal AIC','r','p1','p2')
> rownames(AICM)=NULL
> AICM
     d nominal AIC    r p1 p2
[1,] 1       86.24 7.42  3  2
[2,] 2       83.87 3.87  3  3
[3,] 3       89.37 5.29  3  3
```

Let us fit a TAR(2; 3, 3) model with $d = 2$.

```
> hare.tar.1=tar(y=sqrt(hare),p1=3,p2=3,d=2,a=.1,b=.9,print=T)
```

	Estimate	Std. Error	t-statistic	p-value
$\hat{d}$	1			
$\hat{r}$	3.873			
		Lower Regime (7)		
$\hat{\phi}_{1,0}$	3.89	1.21	3.22	0.049
$\hat{\phi}_{1,1}$	1.25	0.32	3.85	0.03
$\hat{\phi}_{1,2}$	0.38	0.69	0.56	0.61
$\hat{\phi}_{1,3}$	-1.41	0.52	-2.71	0.07
$\tilde{\sigma}_1^2$	1.04			
		Upper Regime (21)		
$\hat{\phi}_{2,0}$	5.05	1.15	4.38	0.00
$\hat{\phi}_{2,1}$	0.84	0.24	3.52	0.00
$\hat{\phi}_{2,2}$	-0.14	0.32	-0.45	0.66
$\hat{\phi}_{2,3}$	-0.52	0.21	-2.45	0.03
$\tilde{\sigma}_2^2$	0.8447			

The estimated threshold is 3.873, which is approximately the 25th percentile of the data. Notice that $\hat{\phi}_{1,2}$ and $\hat{\phi}_{1,3}$ are not significant, which suggests the model TAR$(2;1,3)$ with delay 2. But the nominal AIC of TAR$(2;1,3)$ is 92.53 which is much larger than that of TAR$(2;3,3)$, 83.87.

Let us do model diagnostics.

```
> win.graph(width=4.875, height=4.5,pointsize=8)
> tsdiag(hare.tar.1,gof.lag=20)
```

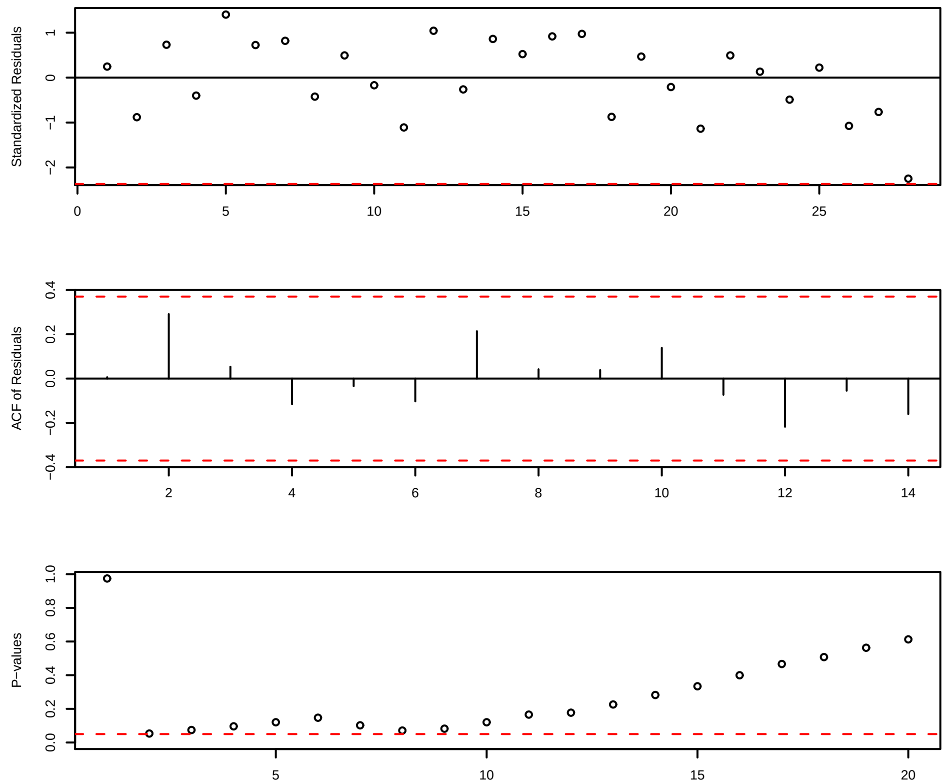


Figure 9: Diag of Lagged Regression of Sqrt Transformed Hare Data

In Figure 9, the ACF and p-values of standard residuals from the fitted model look good. Next let us see the normality of the standard residuals.

```
> win.graph(width=2.5, height=2.5,pointsize=8)
> qqnorm(hare.tar.1$std.res)
> qqline(hare.tar.1$std.res)
```

It looks like the distribution of the residuals has a lighter right tail than standard normal distribution. This may be caused by the small data size. Remember we only have 31 cases in the original data.

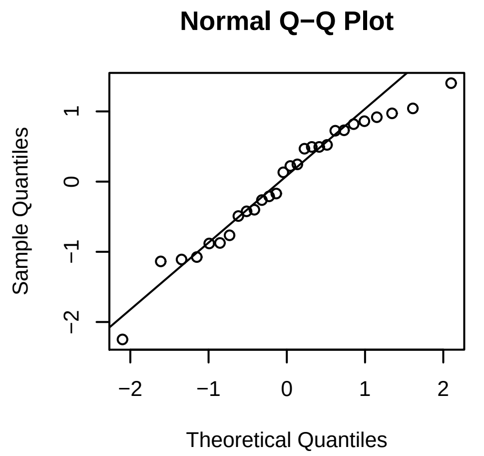


Figure 10: QQ Normal Plot of Standard Residuals

Exercise 15.9

(a) Let $P = (p_{ij})\,, i, j = 1, 2$ be the transition probability matrix of R_t. Then

$$\begin{aligned} p_{11} &= \Pr\left(Y_{t+1} \le r \middle| Y_t \le r\right) \\ &= \Pr\left(\phi_{1,0} + \sigma_1 e_{t+1} \le r\right) \\ &= \Phi\left(\frac{r - \phi_{1,0}}{\sigma_1}\right) \end{aligned}$$

and

$$p_{12} = 1 - p_{11} = \Phi\left(\frac{\phi_{1,0} - r}{\sigma_1}\right),$$

where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal random variable. Similarly,

$$\begin{aligned} p_{21} &= \Pr\left(Y_{t+1} \le r \middle| Y_t > r\right) \\ &= \Pr\left(\phi_{2,0} + \sigma_2 e_{t+1} \le r\right) \\ &= \Phi\left(\frac{r - \phi_{2,0}}{\sigma_2}\right) \end{aligned}$$

and

$$p_{22} = 1 - p_{21} = \Phi\left(\frac{\phi_{2,0} - r}{\sigma_2}\right).$$

Let $\begin{pmatrix} \pi_1 & \pi_2 \end{pmatrix}^T$ be the stationary distribution of $\{R_t\}$. Then $\pi_1 + \pi_2 = 1$ and

$$\begin{pmatrix} \pi_1 & \pi_2 \end{pmatrix} \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} = \begin{pmatrix} \pi_1 & \pi_2 \end{pmatrix}.$$

It can be derived that

$$\begin{pmatrix} \pi_1 \\ \pi_2 \end{pmatrix} = \begin{pmatrix} \frac{p_{21}}{p_{21}+p_{12}} \\ \frac{p_{12}}{p_{21}+p_{12}} \end{pmatrix}.$$

(b) Assume $\{Y_t\}$ is stationary with stationary distribution F. Then for every $y \in \mathrm{R}$, we have

$$F(y) = F(r)\,\Phi\left(\frac{y-\phi_{1,0}}{\sigma_1}\right) + (1-F(r))\,\Phi\left(\frac{y-\phi_{2,0}}{\sigma_2}\right). \tag{0.1}$$

Let $y = r$ in (0.1) to obtain

$$F(r) = \frac{\Phi\left(\frac{r-\phi_{2,0}}{\sigma_2}\right)}{1-\Phi\left(\frac{r-\phi_{1,0}}{\sigma_1}\right)+\Phi\left(\frac{r-\phi_{2,0}}{\sigma_2}\right)}. \tag{0.2}$$

Plugging (0.2) into (0.1) we get the stationary distribution

$$F(y) = \frac{\Phi\left(\frac{r-\phi_{2,0}}{\sigma_2}\right)\Phi\left(\frac{y-\phi_{1,0}}{\sigma_1}\right)+\left(1-\Phi\left(\frac{r-\phi_{1,0}}{\sigma_1}\right)\right)\Phi\left(\frac{y-\phi_{2,0}}{\sigma_2}\right)}{1-\Phi\left(\frac{r-\phi_{1,0}}{\sigma_1}\right)+\Phi\left(\frac{r-\phi_{2,0}}{\sigma_2}\right)}.$$

(c) From part (b), we know

$$\mathrm{E}(Y_t) = \mathrm{E}(Y_{t-1}) = \phi_{1,0}F(r) + \phi_{2,0}(1-F(r))$$

and

$$\begin{aligned} \mathrm{E}(Y_tY_{t-1}) &= \phi_{1,0}\mathrm{E}\left(Y_{t-1}\mathrm{I}_{(Y_{t-1}\le r)}\right) + \phi_{2,0}\mathrm{E}\left(Y_{t-1}\mathrm{I}_{(Y_{t-1}>r)}\right) \\ &= \phi_{1,0}\frac{\int_{-\infty}^{r} y\mathrm{d}F(y)}{F(r)} + \phi_{2,0}\frac{\int_{r}^{\infty} y\mathrm{d}F(y)}{1-F(r)}, \end{aligned}$$

where I_E denotes the indicator function of event E. So we have

$$\begin{aligned} \mathrm{Cov}(Y_t, Y_{t-1}) &= \mathrm{E}(Y_tY_{t-1}) - \mathrm{E}(Y_t)\,\mathrm{E}(Y_{t-1}) \\ &= \phi_{1,0}\frac{\int_{-\infty}^{r} y\mathrm{d}F(y)}{F(r)} + \phi_{2,0}\frac{\int_{r}^{\infty} y\mathrm{d}F(y)}{1-F(r)} - \left(\phi_{1,0}F(r) + \phi_{2,0}(1-F(r))\right)^2. \end{aligned}$$