Chapter 5 Solutions

5.2.6 Exercises

Exercise 5.1

Notice how we use properties of the expectation operator learned in probability.

(a)

$$E(Y_t) = E(8 + 3t + X_t) = 8 + E(3t) + E(X_t) = 8 + 3t$$

(b)

$$\begin{aligned} Cov(Y_t, Y_{t-k}) &= E((8 + 3t + X_t) - (8 + 3t))((8 + 3(t-k) + X_{t-k}) - (8 + 3(t-k))) \\ &= E(X_t X_{t-k}) = Cov(X_t, X_{t-k}) \\ &= \gamma_k \end{aligned}$$

(c) NO, $\{Y_t\}$ is NOT stationary. The expected value is a function of t even though the covariance is not a function of t.

□

Exercise 5.2

(a) No, X_t is not stationary since the Expected value depends on t.

(b)

$$E(Y_t) = E(7 - 3t + X_t) = 7 - 3t + 3t = 7$$

Notice that $Y_t - E(Y_t) = -3t + X_t$ and $Y_{t-k} - E(Y_{t-k}) = -3(t-k) + X_{t-k}$.

Substituting these last results into the definition of covariance, we realize that

$$Cov(Y_t, Y_{t-k}) = E[(X_t - 3t)(X_{t-k} - 3(t-k))] = Cov(X_t, X_{t-k}) = \gamma_k$$

because $E[(X_t - 3t)(X_{t-k} - 3(t-k))] = E[(X_t - E(X_t))(X_{t-k} - E(X_{t-k})]$ and they tell us in the problem that this covariance is free of t.

Neither the expected value of Y_t, nor the covariance depend on t. Thus Y_t is stationary.

□

Exercise 5.3

(a) $E(Y_t) = (-1)^t E(X_t) = 0$

(b) $\gamma_{k,Y} = Cov(Y_t, Y_{t-k}) = Cov((-1)^t X_t (-1)^{t-k} X_{t-k}) = (-1)^{2t-k} Cov(X_t, X_{t-k}) = (-1)^k \gamma_{k,X}$

For $k = 0$, we get $\gamma_{0,Y} = Var(Y_t) = \gamma_{0,X} = \sigma_X^2$

For $k = 1$, we get $\gamma_1 = -\gamma_{1,X}$

For $k = 2$ and every even k we get $\gamma_k = \gamma_{k,X}$

(c) The mean is constant,i.e., does not depend on time t, the variance is σ_X^2, a constant, not depending on t, and the covariance depends only on k, not on t, assuming that the covariance function for X does not depend on t.

Note:$(-1)^{2t} = 1; (-1)^{-k} = (-1)^k$

□

Exercise 5.4

Using the expectation operator, and the property of $E(W) = 0, Var(W) = 1$, we can find that $Cov(Y_t, Y_{t-1}) = 0.46$ of model (a) and -0.4683749 for model (b).

$Var(Y_t) = 1 + 0.4683749^2$ for both (a) and (b).

By definition, $\rho_1 = \frac{Cov(Y_t, Y_{t-1})}{Var(Y_t)}$.

For ρ_1 to be equal to -0.3796633, the $Cov(Y_t, Y_{t-1}) = -0.46$. Thus model (b) corresponds to the given first order correlation coefficient.

□

5.3.3 Exercises

Exercise 5.5

Use R Program ch5simcode.R, in particular, the section dedicated to Figure 5.7. In that section, we provide two random noise series. Add three more. Also add for each of the series,

```
acf(y,main=" put title indicating whether any autocorrelations significant", lwd=2)
pacf(y, main="add title saying whether any partial autocorr significant", lwd=2)
```

The images that you obtain for the Sample ACF should be very similar to those of Figure 5.7. Occasionally, do not be surprised to find the 5% expected significant autocorrelations or borderline significant, scattered along the whole Sample ACF image. When using data, the Sample ACF does not look like the mathematical ACF.

□

5.3.8 Exercises

□

Exercise 5.7

Example 5.6 is referring to an MA(1) model with coefficient 0.5, but the result is more general for any MA(1) with any coefficient, namely:

$$\begin{aligned}
\gamma_2 = Cov(Y_t, Y_{t-2}) &= E[(Y_t - 0)(Y_{t-2} - 0)] \\
&= E(w_t + \beta_1 w_{t-1})(w_{t-2} + \beta_1 w_{t-3}) \\
&= E(w_t w_{t-2}) + \beta_1 E(w_t wt - 3) + \beta_1 E(w_{t-1} w_{t-2}) + \beta_1^2 E(w_{t-1} w_{t-3}) \\
&= 0
\end{aligned}$$

Do similar operations for γ_3.

□

Exercise 5.8

For an MA(1) model ($Y_t = \beta_1 W_{t-1} + W_t = (1 + \beta_1 B)W_t)$), $\rho_1 = \frac{\beta_1}{1+\beta_1}$ as we saw in Section 5.3.5.

If $\beta_1 = 1/2$, then $Y_t = \frac{1}{2} W_{t-1} + W_t$, $\rho_1 = 2/5$ and the modulus of the root of the backshift polynomial, found with R as follows

```
Mod(polyroot(c(1,1/2)))
```

is 2, larger than 1. Hence the process is invertible.

If $\beta_1 = 2$, then $Y_t = 2W_{t-1} + W_t$, $\rho_1 = 2/5$, which is the same as in the previous case, but the modulus of the root of the backshift polynomial, found with R as follows

```
Mod(polyroot(c(1,2)))
```

is 0.5, less than 1. Hence the process is not invertible.

All of this implies that were we to have autocorrelation at lag 1 of 2/5, it is not clear whether we should assume that the stochastic model has coefficient $\beta_1 = 1/2$ or $\beta_1 = 2$.

□

Exercise 5.9

$$\rho_1 = \frac{\beta_1}{1 + \beta_1^2} = 0.4.$$

$$\beta_1 = \frac{1 \pm \sqrt{1 - 4(0.4)^2}}{2(0.4)} = \texttt{2or1/2}$$

See Exercise 5.8. We would use the model that is invertible, i.e., the one that has $\beta_1 = 1/2$

$Y_t = \frac{1}{2} W_{t-1} + W_t$

and the justification is in Exercise 5.8.

□

Exercise 5.10

We will generate eight realizations of size 100 from the model

$Y_t = -0.8W_{t-1} + W_t$,

but we will write the program below to do a plot for four time series. The reader can modify slightly the program to obtain the plot for the remaining four.

```
ata=matrix(0,ncol=8,nrow=100) # space to put the data
head(data)
# simulate from the model requested
for(j in 1:8){
data[,j]=arima.sim(n=100, list(ma=-0.8 ), innov= rnorm(100)) # generating MA(1) data
}
head(data)
dim(data)
## Make the data a multiple time series object and make up some time index
data.ts=ts(data,start=c(1998,1),frequency = 12)

#Plots

par(mfrow=c(4,3))
for(j in 1:4){
plot.ts(data[,j], main=paste("Time plot series", j))
acf(data[,j], main=paste("ACF of series", j), ylab)
pacf(data[,j], main=paste("PACF of series", j))
}
```

We observe that all four time series fluctuate randomly around the mean. The ACF of all the time series indicate that there is a significant negative autocorrelation $r_1 \approx 0.5$, which makes sense since according to Section 5.3.5, the theoretical autocorrelation at lag 1 is −0.4878049 and 0 for any other lag, and the sample should indicate that. As indicated in the Chapter, the other borderline and nonsignificant autocorrelations should be ignored (again, see Section 5.3.5). The PACF for all the time series shows more than one significant autocorrelation at low lags of 1, 2, 3... indicating. What we observe is typical of an MA process of order 1. The ACF will show just one significant autocorrelation at lag 1, but the PACF will show more lags significant at low lags 1, 2, 3,.. sometimes more.

□

Exercise 5.13

This will generate a time series sequence of data from

$$X_t = W_t + 0.9W_{t-1}$$

where W_t $N(0, 1)$. The readers, series will be slightly different, since it is random and we are not using set.seed() to fix the data for replication.

```
x=arima.sim(n=100,list(ma=0.9),innov=rnorm(100))
x.ts=ts(x,start=c(1960,1),freq=12)
acf(x.ts, main="ACF of MA(1) with parameter 0.9")
acf(x.ts)$acf  # view all the ACF plotted
acf(x.ts)$acf[2]
```

The sample time series generated from the process given gives $r_1 = 0.499727188$. The reader's r_1 will be slightly different. Since the size of the sample time series, n=100, is not very large, we do not expect the r_1 to be exactly like the theoretical one, $\rho_1 = \frac{0.9}{1+0.9^2} = 0.4972376$, but the reader will appreciate that even with n=100, the estimate is really close to the ρ_1.

□

5.3.11 Exercises

Exercise 5.14

Use ch1passengers.csv data file in the Base R programs site in timeseriestime.org for Chapter 5. The R code for this problem is in Program ch5passEx5-14.R

After running the code, confirming that we have the same data as in Chapter 1, cleaning the data, and looking at the Sample ACF and Sample PACF of the random term of the decomposition of the data for the years 2012-2017, we conclude that there is nothing left to explain after we remove the trend and the seasonal component. The sample ACF and the sample PACF are those of a white noise time series.

If we were to model this time series, the only terms we would have to include is the trend and the seasonal features.

□

Exercise 5.15

None of the models could represent this process. Neither of them give a $\rho_1 = -0.6$. On addition to that, none of these models have $\mu = 4$.

□

Exercise 5.16

The process is invertible. Writing the process in backshift polynomial form,

$$x_t = (1 + 0.7B + 0.5B^2 + 0.2B^3)w_t$$

and using Polyroot in R, the Modulus of the roots of the polynomial found are 1.581139, 2.0, and 1.581139, all of them larger than 1. Therefore the process is invertible.

$E(X_t) = 0$

$Var(X_t) = (0.7^2 + 0.5^2 + 0.2^2 + 1)\sigma^2 = 1.78\sigma^2$

$\gamma_1 = cov(X_t, X_{t-1}) = E[(X_t - 0)(X_{t-1} - 0)] = 1.15\sigma^2$ and $\rho_1 = 1.11/1.78 = 0.646$

$\gamma_2 = cov(X_t, X_{t-2}) = E[(X_t - 0)(X_{t-2} - 0)] = 0.64\sigma^2$ and $\rho_2 = 0.64/1.78 = 0.3595$

$\gamma_3 = cov(X_t, X_{t-3}) = E[(X_t - 0)(X_{t-3} - 0)] = 0.2\sigma^2$ and $rho_3 = 0.2/1.78 = 0.11235$

$\gamma_k = cov(X_t, X_{t-k}) = 0, k > 3$ and $rho_k = 0$ for $k > 3$.

We simulate a time series of length n=1000. The larger the time series the closer will the time plot and Sample ACF will be to the true process and its acf. The first values of r_k are $r_0 = 1, r_1 = 0.648154404, r_2 = 0.363287792, r_3 = 0.111934054, r_4 = -0.005096919$

(Notice that answers will vary slightly, since we are simulating) Compared to what we saw in the theory part of part (a), these numbers are very close to the theoretical values we got there.

The following program will do the job for us.

```
rho=function(k,beta){
q=length(beta)-1
if (k>q) ACF=0 else {
s1=0; s2=0
for(i in 1:(q-k+1)) s1=s1+beta[i]*beta[i+k]
for(i in 1:(q+1)) s2= s2+beta[i]^2
ACF=s1/s2}
ACF
}

beta=c(1,0.7,0.5,0.2)
rho.k=rep(1,10)
for(k in 1:10) rho.k[k] = rho(k,beta)
rho.k

par(
mfrow=c(2,1),
font.axis=2,
mar=c(5,5,5,5),
font.main=2,
font.lab=2
)

plot(0:10,c(1,rho.k),pch=4,lwd=1.5,cex=0.5,type="h",
ylab=expression(rho[k]),
xlab="k",
main= TeX("$ACF \\; of \\;  X_t= 0.7W_{t-1} +0.5 W_{t-2} +0.2 W_{t-3} +W_t$"))
abline(h=0)
text(x=5,y=0.8, TeX("$  \\; \\rho_k=0.64,0.36,0.11,   k\\;=1,2,3  \\;\\rho_k=0,
\\;\\;k >2\\;  $"))

y=arima.sim(n=1000, list(ma=c(0.7, 0.5,0.2) ), innov= rnorm(1000))
acf(y, main= TeX("$ ACF \\ of \\ sample  \\ from \\ Y_t= 0.7 W_{t-1} +0.5 W_{t-2}
 +0.2 W_{t-3}+W_t $"))
# dev.off() # execute this only when done viewing the plots.
```

To make you aware of how much clutter we get in the ACF if the time series is shorter, try to simulate just a realization of the process of size n=100. That is the reason why it is recommended to focus on the lower lags, k=1,2,3... if you have a small sample. (Of course, if we have seasonality we also care about the seasonal lag, but that is not the case here.) □

□

5.4.4 Exercises

Exercise 5.19

This is a good exercise to illustrate how real world time series are much harder to identify a model for than simulated series. Run the following code to see the images requested.

```
random=decompose(AirPassengers,type="mult")$random
par(mfrow=c(3,1))
plot.ts(random,
main="Time plot of random component\n of mult decomp of AirPAssenger")
acf(random, lag=50, main="ACF of random component \n
of mult decomp of Air Passengers", na.action=na.omit)

pacf(random, lag=50, main="PACF of random component \n
of mult decomp of Air Passengers", na.action=na.omit)
dev.off()
```

We realize upon looking at the time plot of the random component that doing multiplicative decomposition did not account enough for the changing variance over time in the series. The time plot looks like the variance is larger before 1954 and after 1958. If we pretend we did not notice that, and interpret the ACF and PACF to identify a stochastic model, then we notice that the ACF looks like the ACF of an AR process: it dies away slowly in a sinusoidal fashion. In those cases, we rely on the PACF for identifying the order of the AR model. Looking at the PACF we notice a significant autocorrelation at lag one of 0.4. Then after a break of one lag, newly significant partial autocorrelations at lags 3 and 7. Those could be artifacts of sampling. So we would be inclined to identify an AR model of order 1 with coefficient $\alpha_1 = 0.4$. This looks a little simplistic, given the additional autocorrelations in the PACF. Perhaps the model is an ARMA..

To investigate further the possibility of the AR(!), we can simulate an AR(1) process with that model coefficient. We do that with the next code. After we run it, several times, we notice that in some of the trials, the realizations give ACF and PACF that could be approximately similar to the AirPassengers one.

```
   # Simulate to see if our guess could be possible.
length(AirPassengers)
x=arima.sim(n=144, list(ar=0.4), innov= rnorm(144))
par(mfrow=(c(3,1)))
```

```
plot.ts(x,
main="Time plot of simulated AR(1) with"~alpha[1]~"=0.4")
acf(x, lag=50, main="Sample ACF of simulated AR(1) with"~alpha[1]~"=0.4",
na.action=na.omit)

pacf(x, lag=50, main="Sample  PACF of simulated AR(1) with"~alpha[1]~"=0.4",
 na.action=na.omit)
```

The ACF and PACF that we saw for the random term of AirPassengers could have been affected by the lack of constant variance. Notice how the simulated data has a time plot with more homogeneous variance across time than the AirPassengers data.

In any case, it is customary to try several possible models, as we will see in later chapters.

□

□

Problems

Problem 5.1

```
install.packages("latex2exp")
library(latex2exp)


y=arima.sim(n=200,list(ma=0.5),innov=rnorm(200))
y.t=ts(y,start=2000)

par(mfrow=c(3,1))

plot.ts(y.t,ylab=TeX("$y_t$"),
main=TeX("$Realization \\; of\\; Y_t=0.5W_{t-1}+ W_t$"))
acf(y.t,main= TeX("$ACF \\; of \\;realization \\; of \\; Y_t=0.5W_{t-1}+ W_t$")
,ylim=c(-1,1))
acf(y.t)$acf[2]  # to get r_1

pacf(y.t,main= TeX("$PACF \\; of \\;realization \\; of \\; Y_t=0.5W_{t-1}+ W_t$")
,ylim=c(-1,1))
pacf(y.t)$acf[1]

dev.off()
# Notice that the ACF of a MA(1) process will show significant autocorrelation
# value at lag=1 and no other significant autocorrelations (Section 5.3.4,
# Example 5.6). The PACF's $r_{11}=r_1$ so you will notice that as well, when
#
```

After running this R program, we observe that the $r_1 = 0.4021641$ is statistically significant and it is equal to r_{11}. So everything is according to what we would expect from theory. According to the theory (See Section 5.3.4). For k¿1, both the ACF and the PACF show no other significant autocorrelation. So this is evidence of an MA(1) process. According to theory in Section 5.3.4,

$$0.4 = \frac{\beta_1}{1 + \beta_1^2}$$

which gives as values for β_1 both 0.5 and -0.5. Since only 0.5 gives positive autocorrelation, we keep 0.5. And confirm that the model is

$$Y_t = W_t + 0.5W_{t-1}$$

□

Problem 5.3

Given a time series data set, there is a very tedious way to find out whether the time series is stationary. You split the series into equally long intervals and calculate the sample mean and the sample variance at each interval. Also, compute the sample correlogram. For example, consider a time series X_t that consists of the annual prices of minis (cars) from 1959. Observed values are: $496.95, 487.21, 468.82, 482.35, 401.48, 405, 386.35, 378.76, 395.3, 412.88, 415.85, 415.15$

For checking whether the mean is constant over time, we could divide the data into four intervals of 4 numbers each. Then you would find out that $\bar{x}_1 = 484.326, \bar{x}_2 = 429.61, \bar{x}_3 = 386.80, \bar{x}_4 = 414.626$.

The sample standard deviations at each of those chunks of data are $s_1 = 14.28, s_2 = 45.7, s_3 = 8.27, s_4 = 1.55$.

What conclusions do you draw from these results? Explain.

□

Problem 5.9

$$\theta_1(\texttt{process in row1}) = \frac{1 \pm \sqrt{1 - 4(0.4)^2}}{2(0.4)} = \ 2 \text{or} 1/2$$

$$y_t = w_t + 1/2w_{t-1} = (1 + 1/2B)w_t$$

This process is invertible, as the roots of $(1 + 2B)$ are less than 1 in absolute value.

If we had used the root=2, the process would not have been invertible. Choose the coefficient that makes the process invertible.

□

Problem 5.10

(a) $H_0 : \rho_{12} = 0$

$H_a : \rho_{12} \neq 0$

$t = \frac{r_{12}-0}{se_{r_{12}}} = \frac{0.85570}{0.201081} = 4.255499$

4.255499 is beyond two standard errors from 0 (i.e., r_{12} is beyond 2*0.201081=0.402162), so we reject the null hypothesis. P-value less than 0.05.ρ_{12} is significantly different from 0. X_t and X_{t-12} are significantly correlated.

$H_0 : \rho_{10} = 0$

$H_a : \rho_{10} \neq 0$

$t = \frac{r_{10}-0}{se_{r_{10}}} = \frac{0.52219}{0.161467} = 3.234035$

3.234035 is beyond two standard errors from 0(i.e., r_{10} is higher than 2*0.161467=0.322934), so we reject the null hypothesis. P-value less than 0.05. ρ_{10} is significantly different from 0. X_t and X_{t-10} are significantly correlated.

$H_0 : \rho_3 = 0$

$H_a : \rho_3 \neq 0$

$t = \frac{r_3-0}{se_{r_3}} = \frac{0.37952}{0.132702} = 2.859942$

2.859942 is beyond two standard errors from 0 (i.e., r_3 is beyond 0.265404 or 2 standard errors), so we reject the null hypothesis. P-value less than 0.05.ρ_3 is significantly different from 0. X_t and X_{t-3} are significantly correlated.

(b) $H_0 : \rho_1 = \rho_2 = \rho_3 = \rho_4 = \rho_5 = \rho_6 = 0$

H_a : At least one $\rho_k, k = 1, ..., 6$ is not 0.

$$Q_6 = n(n+2)\sum_{i=1}^{6}(n-6)^{-1}r_i^2 = 168(170)\left(\frac{0.79356^2}{167} + \frac{0.59118^2}{166} + \frac{0.37952^2}{165} + \frac{0.28254^2}{164} + \frac{0.13148^2}{163} + \frac{0.12412^2}{162}\right) = 212.4242$$

Critical value for a chi-square distribution at $\alpha = 0.05$ and degrees of freedom=6 is 12.59.

$P - value = P(Q_6 > 212.42) < 0.05.$

So we reject the null hypothesis. The time series X_t is NOT white noise.

□

Problem 5.11

(a)

$$y_t = \frac{1}{2}y_{t-1} + w_t$$

B=2, stationary

(b)

$$y_t = y_{t-1} - \frac{1}{4}y_{t-2} + w_t$$

$1/4(B^2 - 4B + 4)y_t = 1/4(B - 2)^2 = 0$ B=2, stationary.

(c)

$$y_t = 1/2 y_{t-1} + 1/2 y_{t-2} + w_t$$

nonstationary because one of the roots is unity. $-1/2(B^2 + B - 2)y_t = -1/2(B-1)(B+2)y_t = w_t$. The roots are B=1, $B = -2$.

(d)

$$y_t = -1/4 y_{t-2} + w_t$$

stationary. The roots of $1 + 1/4B^2 = 0$ are $B = \pm 2i$ which are complex numbers with $i = \sqrt{-1}$, each having an absolute value of 2, exceeding unity.

□

5.11 Quiz

Question 5.1

$-\frac{5}{8}\sigma^2$ □

Question 5.3

0 □

Question 5.5

−0.5 □

Question 5.7

The process is mean non-stationary but covariance stationary □

Question 5.9

It is invertible, it has a root outside the unit circle

□

□