Thomas A. Severini

Solutions Manual for Introduction to Statistical Methods for Financial Models

Contents

Changes to the Yahoo Finance database

During the time that *Introduction to Statistical Methods for Financial Models* was in press significant changes were made to the Yahoo Finance database of stock price information.

For the examples in the text and the solutions to the exercises the most important change is in the convention used to specify dates when downloading monthly price data. Previously, the arguments `start` and `end` in the function `get.hist.quote` were specified as the last days of the relevant months. Now, the argument `start` should be taken to be the **first** day of the relevant month; the argument `end` should continue to be taken to be the last day of the relevant month (as was done previously).

For instance, to download the monthly adjusted prices for the stock with symbol "BBY" for December, 2010 to December, 2015 (assigning the data to the variable `x`), the command

```
> x<-get.hist.quote(instrument="BBY", start="2010-12-01", end="2015-12-31",
+      quote="AdjClose", compression="m")
```

should be used. Previously (and according to Chapter 2 of the book) `start` would have been taken to be `"2010-12-31"`.

In addition, for some stocks and dates, changes have been made to the price data in the database. It is not clear if these changes correct previous mistakes or that these new values are incorrect. As of the end of June 2017, changes are still being made to the values in the database; hence, eventually the new database may match closely the previous one. However, for exercises that use the Yahoo data, it is possible that the numerical results resulting from the commands in the solutions will differ from those given here. The same is true for the examples in the text.

Solutions for Chapter 2

2.1. The returns are given by $R_t = (P_t - P_{t-1})/P_{t-1}$, $t = 1, 2, 3, 4$. Hence,

$$R_1 = 0.0400, \quad R_2 = -0.0192, \quad R_3 = 0.0784, \quad R_4 = 0.0091.$$

The log-returns are given by $r_t = \log(1 + R_t)$, $t = 1, 2, 3, 4$. Hence,

$$r_1 = 0.0392, \quad r_2 = -0.0194, \quad r_3 = 0.0076, \quad r_4 = 0.0091.$$

2.2. Returns are given by $R_t = (P_t - P_{t-1})/P_{t-1}$ so that

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} - 1 = \frac{a\exp(bt) - a\exp\left(b(t-1)\right)}{a\exp\left(b(t-1)\right)} = \exp(b) - 1.$$

The log-returns are given by $r_t = \log(1 + R_t)$ so that

$$r_t = \log(\exp(b)) = b.$$

2.3. Using a Taylor's series approximation, $\log(1+x) \doteq x$ for small $|x|$. Hence,

$$r_t = \log(1 + R_t) \doteq R_t \quad \text{for small } \ |R_t|.$$

Including an additional term in the Taylor's series,

$$\log(1+x) \doteq x - \frac{1}{2}x^2$$

for small $|x|$. Hence,

$$r_t = \log(1 + R_t) \doteq R_t - \frac{1}{2}R_t^2 \quad \text{for small } \ |R_t|.$$

2.4. Returns are given by $R_t = (P_t + D_t)/P_{t-1} - 1$. Hence,

$$R_1 = 0.200, \quad R_2 = 0.125, \quad R_3 = 0.080.$$

2.5. Adjusted prices are given by $\bar{P}_3 = P_3 = \$5.40$,

$$\bar{P}_2 = \left(1 - \frac{D_3}{P_2}\right) P_2 = P_2 = \$4.80$$

and

$$\bar{P}_1 = \left(1 - \frac{D_2}{P_1}\right)\left(1 - \frac{D_3}{P_2}\right) P_1 = \left(1 - \frac{0.40}{4.00}\right) 4.00 = \$3.60.$$

2.6. (a) The single-period return at time t is given by

$$R_t = \frac{P_t + D_t}{P_{t-1}} - 1 = \frac{P_t + \alpha P_{t-1}}{P_{t-1}} - 1 = \frac{P_t}{P_{t-1}} - 1 + \alpha.$$

(b) Let $\bar{P}_t$, $t = 0, 1, 2, \ldots, T$ denote the sequence of adjusted prices. Then $\bar{P}_T = P_T$,

$$\bar{P}_{T-1} = \left(1 - \frac{D_T}{P_{T-1}}\right) P_{T-1} = (1 - \alpha)\, P_{T-1},$$

$$\bar{P}_{T-2} = \left(1 - \frac{D_T}{P_{T-1}}\right)\left(1 - \frac{D_{T-1}}{P_{T-2}}\right) P_{T-2} = (1-\alpha)^2 P_{T-2}$$

and so on. The general relationship is

$$\bar{P}_{T-k} = (1-\alpha)^k P_{T-k}.$$

2.7. Consider

$$\mathrm{Cov}(Y_t + X_t, Y_s + X_s) = \mathrm{Cov}(Y_t, Y_s) + \mathrm{Cov}(X_t, X_s) + \mathrm{Cov}(Y_t, X_s) + \mathrm{Cov}(Y_s, X_t).$$

Let γ_Y denote the covariance function of $\{Y_t : t = 1, 2, \ldots\}$ and let γ_X denote the covariance function of $\{X_t : t = 1, 2, \ldots\}$. Since these processes are both weakly stationary,

$$\mathrm{Cov}(Y_t + X_t, Y_s + X_s) = \gamma_Y(|t-s|) + \gamma_X(|t-s|) + \mathrm{Cov}(Y_t, X_s) + \mathrm{Cov}(Y_s, X_t).$$

However, since we do not know anything about the covariance of Y_t and X_s, it does not follow that the process $Y_1 + X_1, \ldots$ is weakly stationary. For instance, if Y_t and X_s are uncorrelated for all t, s, then it is weakly stationary. However, if the correlation of Y_t and X_s is $1/2$ if $t = s = 1$ and 0 otherwise, then the process is not weakly stationary.

2.8. (a) The mean function is given by

$$\mathrm{E}(Y_t) = \mathrm{E}(X_t - X_{t-1}) = \mathrm{E}(X_t) - \mathrm{E}(X_{t-1}) = 0$$

and the variance function is given by

$$\mathrm{Var}(Y_t) = \mathrm{Var}(X_t - X_{t-1}) = \mathrm{Var}(X_t) + \mathrm{Var}(X_{t-1}) - 2\mathrm{Cov}(X_t, X_{t-1}) = 2\sigma^2 - 2\gamma(1).$$

(b) The covariance function is given by

$$\begin{aligned}
\mathrm{Cov}(Y_t, Y_s) &= \mathrm{Cov}(X_t - X_{t-1}, X_s - X_{s-1}) \\
&= \mathrm{Cov}(X_t, X_s) + \mathrm{Cov}(X_{t-1}, X_{s-1}) - \mathrm{Cov}(X_{t-1}, X_s) - \mathrm{Cov}(X_t, X_{s-1}) \\
&= 2\gamma(|t-s|) - \gamma(|t-s-1|) - \gamma(|t-s+1|).
\end{aligned}$$

(c) The mean and variance functions of the process are constant. Consider the term in the covariance function

$$\gamma(|t-s-1|)+\gamma(|t-s+1|).$$

If $t=s$ this is $2\gamma(|1|)=2\gamma(|t-s|+1)$. If $t\geq s+1$, then

$$\gamma(|t-s-1|)+\gamma(|t-s+1|)=\gamma(t-s-1)+\gamma(t-s+1)=\gamma(|t-s|-1)+\gamma(|t-s|+1).$$

Similarly, if $t\leq s-1$,

$$\gamma(|t-s-1|)+\gamma(|t-s+1|)=\gamma(s+1-t)+\gamma(s-1-t)=\gamma(|t-s|+1)+\gamma(|t-s|-1).$$

It follows that the covariance of Y_t, Y_s is a function of $|t-s|$ and, hence, the process is weakly stationary.

2.9. (a) $\mathrm{E}(X_t)=\mathrm{E}(ZZ_t)=\mathrm{E}(Z)\mathrm{E}(Z_t)=0$; hence, the mean function is 0. Let $\mu=\mathrm{E}(Z_t)$ and $\sigma^2=\mathrm{Var}(Z_t)$. Since $\mathrm{E}(X_t^2)=\mathrm{E}(Z^2Z_t^2)=\mathrm{E}(Z^2)\mathrm{E}(Z_t^2)=(\sigma^2+\mu^2)$, the variance function of the process is $\sigma^2+\mu^2$.

(b) Since $\mathrm{E}(X_t)=0$ for all t,

$$\mathrm{Cov}(X_t,X_s)=\mathrm{E}(X_tX_s)=\mathrm{E}(Z^2Z_tZ_s)=\mathrm{E}(Z^2)\mathrm{E}(Z_tZ_s)=0.$$

(c) Since the mean and variance functions are constant and the $X_1, X_2, \ldots$ are uncorrelated, it follows that $\{X_t : t=1,2,\ldots\}$ is a white noise process. Hence, it is also weakly stationary.

2.10. Since $\mathrm{E}(r_t)$ does not depend on t, clearly $\mathrm{E}(\tilde{r}_t)$ does not depend on t. Let $\sigma^2=\mathrm{Var}(r_t)$ and consider $\mathrm{Var}(\tilde{r}_t)$. Using the fomula for the variance of a sum,

$$\mathrm{Var}(\tilde{r}_t)=21\sigma^2+2\sum_{i<j}\mathrm{Cov}(\tilde{r}_{21(t-1)+i},\tilde{r}_{21(t-1)+j})$$

where the sum in this expression is over all i,j from 1 to 21 such that $i<j$. Note that, since $\{r_t : t=1,2,\ldots\}$ is weakly stationary,

$$\mathrm{Cov}(\tilde{r}_{21(t-1)+i},\tilde{r}_{21(t-1)+j})=\gamma(|i-j|)$$

where $\gamma(\cdot)$ is the autocovariance function of $\{r_t : t=1,2,\ldots\}$. It follows that $\mathrm{Var}(\tilde{r}_t)$ does not depend on t.

Now consider $\mathrm{Cov}(\tilde{r}_t,\tilde{r}_s)$ for $t\neq s$. Note that

$$\mathrm{Cov}(\tilde{r}_t,\tilde{r}_s)=\sum_{j=1}^{21}\sum_{i=1}^{21}\mathrm{Cov}(r_{21(t-1)+j},r_{21(s-1)+i}).$$

Since, for any i,j,

$$\mathrm{Cov}(r_{21(t-1)+j},r_{21(s-1)+i})=\gamma(|21(t-s)+j-i|)$$

for any $j = 1, 2, \ldots, 21$,

$$\mathrm{Cov}(\tilde{r}_t, \tilde{r}_s) = \sum_{j=1}^{21} \sum_{i=1}^{21} \gamma(|21(t-s) + j - i|),$$

which clearly depends on t, s only through $t - s$. By symmetry of the covariance operator,

$$\sum_{j=1}^{21} \sum_{i=1}^{21} \gamma(|21(t-s) + j - i|) = \sum_{j=1}^{21} \sum_{i=1}^{21} \gamma(|21(s-t) + j - i|)$$

so that $\mathrm{Cov}(\tilde{r}_t, \tilde{r}_s)$ depends on t, s only through $|t-s|$. It follows that $\{\tilde{r}_t : t = 1, 2, \ldots\}$ is weakly stationary.

2.11. Since $\mathrm{E}(X_j) = \mu$, $j = 1, \ldots, n$,

$$\mathrm{E}(Y_k) = \frac{1}{w} \sum_{j=k+1}^{k+w} \mathrm{E}(X_j) = \frac{1}{w} w\mu = \mu, \quad k = 1, \ldots, n-w.$$

Since $X_1, \ldots, X_n$ are independent with $\mathrm{Var}(X_j) = \sigma^2$, $j = 1, \ldots, n$,

$$\mathrm{Var}(Y_k) = \frac{1}{w^2} \sum_{j=k+1}^{k+w} \mathrm{Var}(X_j) = \frac{1}{w^2} w\sigma^2 = \frac{1}{w}\sigma^2, \quad k = 1, \ldots, n-w.$$

Consider $\mathrm{Cov}(Y_i, Y_k)$, where $i < k$. If $k > i + w$, then Y_i and Y_k have no terms in common so that $\mathrm{Cov}(Y_i, Y_k) = 0$. Otherwise, the sums

$$\sum_{j=i+1}^{i+w} X_j \quad \text{and} \quad \sum_{\ell=k+1}^{k+w} X_\ell$$

have terms $X_{k+1}, \ldots, X_{i+w}$ in common so that

$$\mathrm{Cov}(Y_i, Y_k) = \frac{i - k + w}{w^2}\sigma^2.$$

Since $\mathrm{E}(Y_k)$ and $\mathrm{Var}(Y_k)$ are constant and $\mathrm{Cov}(Y_i, Y_k)$ depends only on $k - i$, it follows that the process $Y_1, \ldots, Y_{n-w}$ is weakly stationary with mean function μ and variance function σ^2/w.

The correlation of Y_i and Y_k is

$$\frac{((i-k+w)/w^2)\sigma^2}{\sigma^2/w} = 1 - \frac{k-i}{w}$$

so that the correlation function of the process is

$$\rho(h) = 1 - \frac{|h|}{w}, \quad h = 1, 2, \ldots.$$

2.12. (a) Let $\mu_X = \mathrm{E}(X_t)$, $\sigma_X^2 = \mathrm{Var}(X_t)$, $\mu_Y = \mathrm{E}(Y_t)$, and $\sigma_Y^2 = \mathrm{Var}(Y_t)$. Then

$$\mathrm{E}(X_t + Y_t) = \mathrm{E}(X_t) + \mathrm{E}(Y_t) = \mu_X + \mu_Y,\ t = 1, 2, \ldots,$$

$$\mathrm{Var}(X_t + Y_t) = \mathrm{Var}(X_t) + \mathrm{Var}(Y_t) = \sigma_X^2 + \sigma_Y^2,\ t = 1, 2, \ldots$$

and for $t \neq s$,

$$\mathrm{Cov}(X_t + Y_t, X_s + Y_s) = \mathrm{Cov}(X_t, X_s) + \mathrm{Cov}(X_t, Y_s) + \mathrm{Cov}(Y_t, X_s) + \mathrm{Cov}(Y_t, Y_s) = 0.$$

It follows that $\{X_t + Y_t : t = 1, 2, \ldots\}$ is a weak white noise process.

(b) Using the same notation as in part (a),

$$\mathrm{E}(X_t Y_t) = \mathrm{E}(X_t)\mathrm{E}(Y_t) = \mu_X \mu_Y,\ t = 1, 2, \ldots;$$

note that

$$\mathrm{E}\{(X_t Y_t)^2\} = \mathrm{E}(X_t^2)\mathrm{E}(Y_t^2) = (\mu_X^2 + \sigma_X^2)(\mu_Y^2 + \sigma_Y^2)$$

so that

$$\mathrm{Var}(X_t Y_t) = (\mu_X^2 + \sigma_X^2)(\mu_Y^2 + \sigma_Y^2) - \mu_X^2 \mu_Y^2,\ t = 1, 2, \ldots.$$

Similarly, for $t \neq s$,

$$\mathrm{E}(X_t Y_t X_s Y_s) = \mu_X \mu_Y \mu_X \mu_Y = \mu_X^2 \mu_Y^2$$

so that

$$\mathrm{Cov}(X_t Y_t, X_s Y_s) = \mu_X^2 \mu_Y^2 - (\mu_X \mu_Y)^2 = 0.$$

It follows that $\{X_t Y_t : t = 1, 2, \ldots\}$ is a weak white noise process.

2.13. (a) The following R commands may be used to download the necessary price data.

```
> library(tseries)
> x<-get.hist.quote(instrument="PZZA", start="2012-12-31", end="2015-12-31",
+     quote="AdjClose", compression="d")
> pzza0<-as.vector(x)
```

(b) The returns corresponding to the prices downloaded in part (a) may be calculated using the commands

```
> length(pzza0)
[1] 757
> pzza<-(pzza0[-1]-pzza0[-757])/pzza0[-757]
```

(c) The summary statistics for the returns are

```
> summary(pzza)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-0.1210 -0.0074  0.0010  0.0011  0.0097  0.0804
```

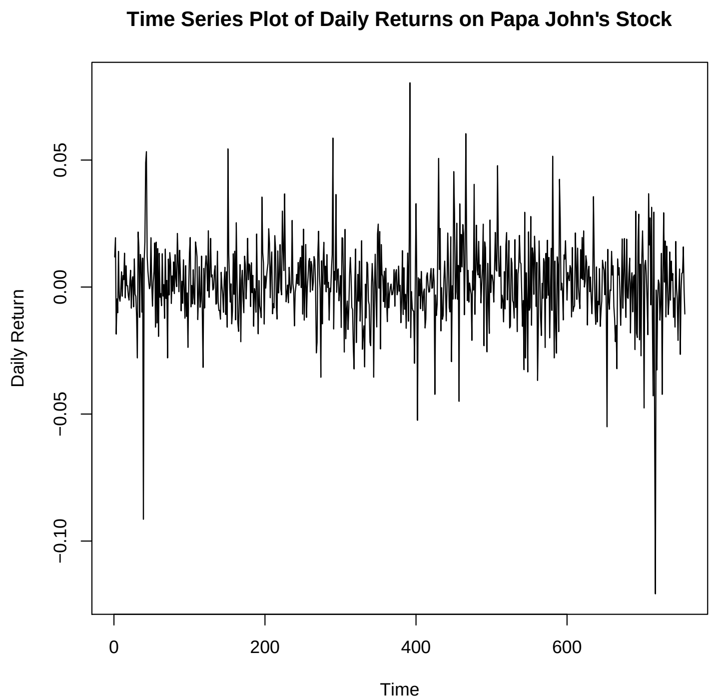

FIGURE 2.1
Plot in Exercise 2.13

(d) The time series plot of the returns may be constructed using the commands

```
> plot(pzza, type="l", xlab="Time", ylab="Daily Return")
> title(main="Time Series Plot of Daily Returns on Papa John's Stock")
```

The plot is given in Figure 2.1.

2.14. (a) The following R commands may be used to download the necessary price data.

```
> library(tseries)
> x<-get.hist.quote(instrument="PZZA", start="2010-12-31", end="2015-12-31",
+     quote="AdjClose", compression="m")
> pzza0<-as.vector(x)
```

(b) The returns corresponding to the prices downloaded in part (a) may be calculated using the commands

```
> length(pzza0)
[1] 61
> pzza.m<-(pzza0[-1]-pzza0[-61])/pzza0[-61]
```

(c) The summary statistics are

```
> summary(pzza.m)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-0.1780 -0.0072  0.0216  0.0265  0.0696  0.1890
```

(d) The time series plot of the returns may be constructed using the commands

```
> plot(pzza.m, type="l", xlab="Time", ylab="Monthly Return")
> title(main="Time Series Plot of Monthly Returns on Papa John's Stock")
```

The plot is given in Figure 2.2.

FIGURE 2.2
Plot in Exercise 2.14

2.15. The running means may be calculated and the plot constructed using the following commands.

```
> library(gtools)
>  pzza.rmean<-running(pzza.m, fun=mean, width=12)
>  mean(pzza.m) + 2*sd(pzza.m)/(12^.5)
[1] 0.0677
>  mean(pzza.m) - 2*sd(pzza.m)/(12^.5)
[1] -0.0147
>  plot(pzza.rmean, type="l", ylim=c(-.02, .07), xlab="Time", ylab="Return")
```

```
> title(main="Running Means for Monthly Returns on Papa John's Stock")
> lines(1:49, rep(0.0677,49), lty=2)
> lines(1:49, rep(-0.0147,49), lty=2)
```

The plot is given in Figure 2.3. According to this plot, there is no evidence of non-stationarity in the returns.

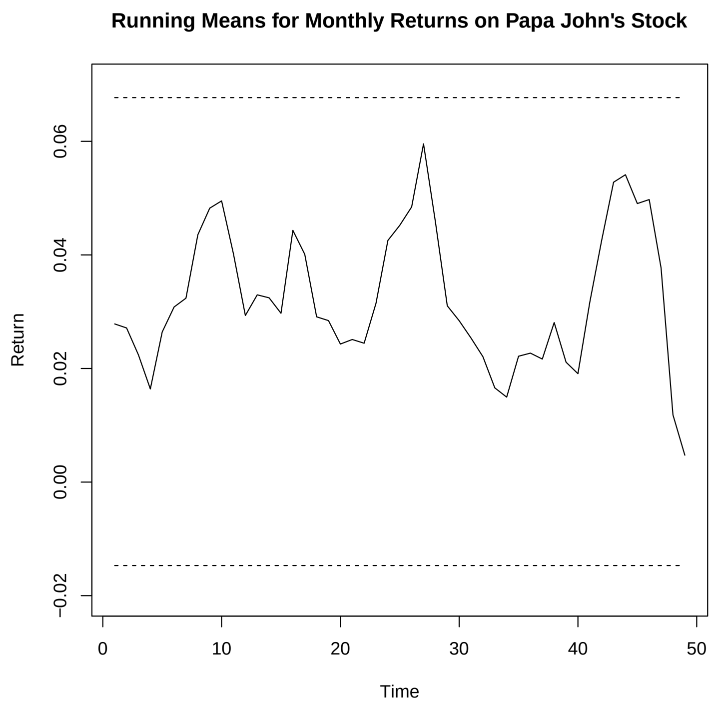

FIGURE 2.3
Plot in Exercise 2.15

2.16. The running standard deviations may be calculated and the plot may be constructed using the following commands.

```
> pzza.rsd<-running(pzza.m, fun=sd, width=12)
> log(sd(pzza.m)) + (2/11)^.5
[1] -2.21
> log(sd(pzza.m)) - (2/11)^.5
[1] -3.07
> plot(log(pzza.rsd), type="l", ylim=c(-3.6, -2), ylab="log of running sd",
+  xlab="time")
> title(main="Log of Running SDs of Returns on Papa John's Stock")
> lines(1:49, rep(-2.21, 49), lty=2)
> lines(1:49, rep(-3.07, 49), lty=2)
```

The plot is given in Figure 2.4. According to this plot, there is some evidence of non-stationarity of the returns. There is a relatively long period of relatively small variability, as well as brief periods of relatively large variability.

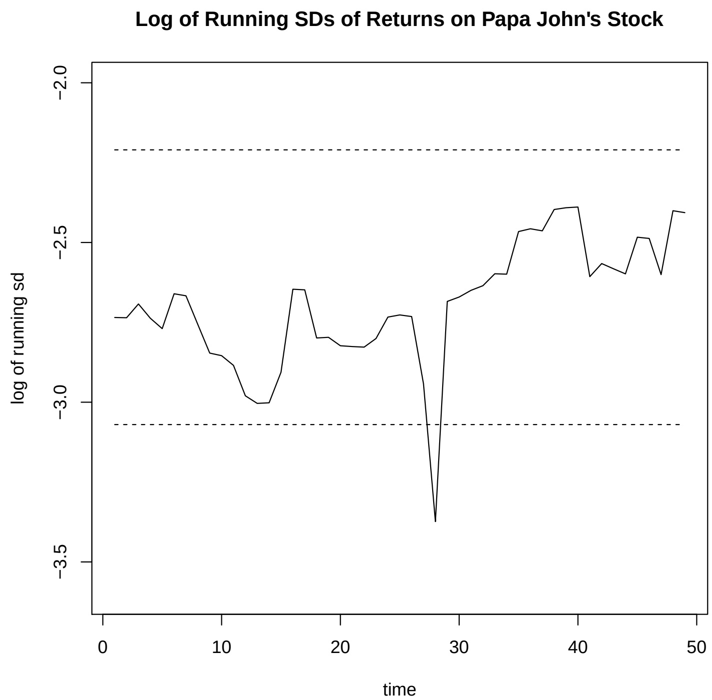

FIGURE 2.4
Plot in Exercise 2.16

2.17. The autocorrelation function based on the daily returns may be calculated using the command

```
>  print(acf(pzza, lag.max=20))

Autocorrelations of series pzza, by lag

     0      1      2      3      4      5      6      7      8      9     10
 1.000 -0.011 -0.013  0.050 -0.013 -0.061  0.005 -0.036  0.062  0.052 -0.007
    11     12     13     14     15     16     17     18     19     20
-0.027  0.044 -0.016 -0.023  0.027  0.029 -0.018  0.032  0.008 -0.028
```

The plot is given in Figure 2.5.

The estimated autocorrelation function based on the monthly returns is given by

```
> print(acf(pzza.m, lag.max=12))

Autocorrelations of series pzza.m, by lag

     0      1      2      3      4      5      6      7      8      9     10
 1.000  0.167 -0.009 -0.061 -0.057 -0.179 -0.146 -0.053  0.103 -0.218 -0.064
    11     12
 0.049  0.015
```