

Solutions to Exercises of
An Introduction to Analysis of Financial Data
with R

Ruey S. Tsay
University of Chicago

February 26, 2013

Preface

The solutions provided are for instructors who adopt the book as a textbook. They are not for distributions to students or any third party. I did the solutions as best as I can over a period of time, but did not check them step-by-step. Omissions, typos, and mistakes are possible. Thus, I would appreciate any feedback. R commands and outputs are included for details. In some cases, I created R scripts to simplify the commands. These R scripts are posted on the web page of the book.

R. Tsay, Feb. 2013

Chapter 1

Financial Data and Their Properties

1. Problem 1.

- (a) The summary statistics of daily simple returns are as follows:

Asset	Mean $\times 100$	StDev $\times 100$	Skewness	Kurtosis	Min	Max
axp	0.053	2.64	0.46	9.59	-0.176	0.206
vw	0.022	1.37	-0.10	7.98	-0.09	0.11
ew	0.06	1.21	-0.25	8.11	-0.08	0.11
sp	0.01	1.38	0.01	8.53	-0.09	0.12

- (b) The summary statistics of daily log returns are as follows:

Asset	Mean $\times 100$	StDev $\times 100$	Skewness	Kurtosis	Min	Max
axp	0.02	2.63	0.02	9.02	-0.19	0.19
vw	0.01	1.37	-0.30	7.88	-0.09	0.11
ew	0.06	1.21	-0.43	8.02	-0.08	0.10
sp	-0.00	1.38	-0.21	8.32	-0.09	0.11

- (a) For log returns of AXP, the t-ratio is 0.36 with p-value 0.72. Thus, we cannot reject the null hypothesis of zero mean return at the 5% level.

2. Problem 2.

- (a) The summary statistics of monthly simple returns are as follows:

Asset	Mean	StDev	Skewness	Kurtosis	Min	Max
ge	0.011	0.07	0.05	1.24	-0.27	0.25
vw	0.009	0.04	-0.66	2.36	-0.23	0.17
ew	0.012	0.05	-0.31	3.14	-0.27	0.30
sp	0.006	0.04	-0.59	2.37	-0.24	0.16

- (b) The summary statistics of monthly log returns are

Asset	Mean	StDev	Skewness	Kurtosis	Min	Max
ge	0.008	0.07	-0.29	1.78	-0.32	0.22
vw	0.008	0.04	-0.94	3.52	-0.26	0.15
ew	0.011	0.06	-0.75	4.17	-0.32	0.26
sp	0.005	0.04	-0.88	3.61	-0.27	0.15

- (c) The t-ratio is 3.713 with p-value 0.0002. Therefore, the null hypothesis of zero mean is rejected at the 5% level. [The 95% confidence interval for the mean DOES not contain zero.]

3. Problem 3. Consider the monthly simple returns of S&P index.

- (a) The t-ratio is 4.24 with p-value 2.44×10^{-5} . The null hypothesis of zero being is rejected at the 5% level.
- (b) The t-ratio is -7.06 whose absolute value is greater than 1.96. Therefore, the null hypothesis is rejected. This is the distribution is skew.
- (c) The t-ratio is 14.18, which is greater than 1.96, implying that the null hypothesis of normal tail is rejected at the 5% level. That is, the returns have heavy tails.

4. Problem 4. Consider the daily log returns of AXP stock.

- (a) The t-ratio is 0.43, which is less than 1.96, implying that the null hypothesis cannot be rejected. That is, we cannot reject that the log returns have a symmetric distribution.
- (b) The t-ratio is 92.71, which is much greater than 1.96. Therefore, the null hypothesis is rejected. That is, the log returns have heavy tails.

5. Problem 5. Daily dollar-yen exchange rates.

- (a) Simply take the `diff` on the `log(rate)`.
- (b) The mean, standard deviation, skewness, excess kurtosis, minimum and maximum for US-JP exchange rate are -0.00 , 0.008 , -0.42 , 4.84 , -0.05 and 0.03 , respectively. Those for US-UK exchange rate are -0.0002 , 0.007 , -0.36 , 5.47 , -0.05 , and 0.04 , respectively.
- (c) The density plot is given in Figure 1.1.
- (d) The t-ratio is -1.61 with p-value 0.11. Therefore, the null hypothesis cannot be rejected. That is, the log return of the exchange rate has zero mean.

R commands and output: edited to save space

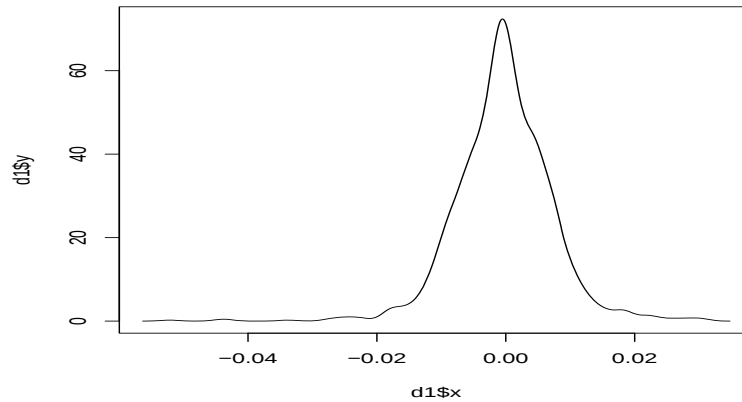


Figure 1.1: Empirical density function of daily log returns of dollar-yen exchange rate from January 02, 2007 to November 30, 2011.

```
#### Problem 1 ####
> library(fBasics)
> da=read.table("d-axp3dx-0111.txt",header=T)
> head(da)
      date      axp      vw      ew      sp
1 20010904  0.000824 -0.001658 -0.005708 -0.000565
2 20010905  0.007682 -0.003242 -0.008926 -0.001059
3 20010906 -0.039477 -0.020733 -0.014187 -0.022390
4 20010907 -0.019274 -0.017769 -0.011476 -0.018637
5 20010910  0.011850  0.003513 -0.007368  0.006226
6 20010917 -0.135961 -0.050714 -0.042398 -0.049216
> basicStats(da$axp)
      X..da.axp
nobs      2535.000000
NAs        0.000000
Minimum    -0.175949
Maximum     0.206485
1. Quartile -0.009672
3. Quartile  0.010540
Mean         0.000534
Median       0.000000
Sum          1.353560
SE Mean      0.000524
LCL Mean     -0.000493
UCL Mean     0.001561
Variance     0.000695
Stdev        0.026368
```

```

Skewness      0.459773
Kurtosis      9.592053
> basicStats(da$vw)
      X..da.vw
nobs      2535.000000
NAs        0.000000
Minimum    -0.089762
Maximum     0.114889
1. Quartile -0.005473
3. Quartile  0.006212
Mean        0.000224
Median      0.000848
Sum         0.567996
SE Mean     0.000271
LCL Mean    -0.000308
UCL Mean     0.000756
Variance    0.000186
Stdev       0.013652
Skewness    -0.098318
Kurtosis    7.982134
> basicStats(da$ew)
      X..da.ew
nobs      2535.000000
NAs        0.000000
Minimum    -0.078240
Maximum     0.107422
1. Quartile -0.004630
3. Quartile  0.006402
Mean        0.000626
Median      0.001429
Sum         1.586462
SE Mean     0.000240
LCL Mean    0.000155
UCL Mean     0.001096
Variance    0.000146
Stdev       0.012080
Skewness    -0.247410
Kurtosis    8.108428
> basicStats(da$sp)
      X..da.sp
nobs      2535.000000
NAs        0.000000
Minimum    -0.090350
Maximum     0.115800
1. Quartile -0.005798
3. Quartile  0.006117
Mean        0.000094
Median      0.000700
Sum         0.238869
SE Mean     0.000274

```

```

LCL Mean      -0.000442
UCL Mean      0.000631
Variance      0.000190
Stdev         0.013779
Skewness      0.008152
Kurtosis      8.532667
> dim(da)
[1] 2535      5
> da1=log(da[,2:5]+1)
> head(da1)
      axp      vw      ew      sp
1  0.0008236607 -0.001659376 -0.005724353 -0.0005651597
2  0.0076526437 -0.003247267 -0.008966075 -0.0010595611
3 -0.0402773512 -0.020950946 -0.014288598 -0.0226444615
4 -0.0194621653 -0.017928764 -0.011542357 -0.0188128573
5  0.0117803385  0.003506844 -0.007395278  0.0062066985
6 -0.1461373723 -0.052045156 -0.043323036 -0.0504683716
> basicStats(da1$axp)
      X..da1.axp
nobs      2535.000000
NAs        0.000000
Minimum    -0.193523
Maximum     0.187711
1. Quartile -0.009719
3. Quartile  0.010484
Mean        0.000188
Median      0.000000
Sum         0.476584
SE Mean     0.000522
LCL Mean    -0.000836
UCL Mean     0.001212
Variance    0.000691
Stdev       0.026294
Skewness    0.020992
Kurtosis    9.020499
> basicStats(da1$vw)
      X..da1.vw
nobs      2535.000000
NAs        0.000000
Minimum    -0.094049
Maximum     0.108755
1. Quartile -0.005489
3. Quartile  0.006193
Mean        0.000131
Median      0.000848
Sum         0.331452
SE Mean     0.000272
LCL Mean    -0.000402
UCL Mean     0.000663
Variance    0.000187

```

```

Stdev          0.013670
Skewness       -0.300352
Kurtosis       7.880082
> basicStats(dal$ew)
      X..dal.ew
nobs      2535.000000
NAs        0.000000
Minimum    -0.081470
Maximum     0.102035
1. Quartile -0.004641
3. Quartile  0.006382
Mean        0.000553
Median      0.001428
Sum         1.400780
SE Mean     0.000240
LCL Mean    0.000081
UCL Mean    0.001024
Variance    0.000146
Stdev       0.012100
Skewness    -0.427315
Kurtosis    8.017712
> basicStats(dal$sp)
      X..dal.sp
nobs      2535.000000
NAs        0.000000
Minimum    -0.094695
Maximum     0.109572
1. Quartile -0.005815
3. Quartile  0.006098
Mean        -0.000001
Median      0.000700
Sum         -0.001898
SE Mean     0.000274
LCL Mean    -0.000538
UCL Mean     0.000536
Variance    0.000190
Stdev       0.013790
Skewness    -0.206357
Kurtosis    8.322826
> t.test(dal$axp)
      One Sample t-test
data:  dal$axp
t = 0.36, df = 2534, p-value = 0.7189
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.0008360686  0.0012120714
sample estimates:
mean of x
0.0001880014

```

```
#### Problem 2 ####
> da=read.table("m-ge3dx-4011.txt",header=T)
> head(da)
      date      ge      vw      ew      sp
1 19400131 -0.061920 -0.024020 -0.019978 -0.035228
2 19400229 -0.009901  0.013664  0.029733  0.006639
3 19400330  0.049333  0.018939  0.026168  0.009893
4 19400430 -0.041667  0.001196  0.013115 -0.004898
5 19400531 -0.197324 -0.220314 -0.269754 -0.239541
6 19400629  0.061667  0.066664  0.066550  0.076591
> basicStats(da$ge)
      X..da.ge
nobs      861.000000
NAs        0.000000
Minimum    -0.272877
Maximum     0.251236
1. Quartile -0.030648
3. Quartile  0.048684
Mean        0.010519
Median      0.007117
Sum         9.056533
SE Mean     0.002249
LCL Mean    0.006104
UCL Mean    0.014933
Variance    0.004356
Stdev       0.065998
Skewness    0.051618
Kurtosis    1.239488
> basicStats(da$vw)
      X..da.vw
nobs      861.000000
NAs        0.000000
Minimum    -0.225363
Maximum     0.165585
1. Quartile -0.016655
3. Quartile  0.038534
Mean        0.009316
Median      0.013354
Sum         8.021486
SE Mean     0.001477
LCL Mean    0.006418
UCL Mean    0.012215
Variance    0.001877
Stdev       0.043328
Skewness    -0.660821
Kurtosis    2.355320
> basicStats(da$ew)
      X..da.ew
nobs      861.000000
NAs        0.000000
```

```

Minimum      -0.272248
Maximum      0.299260
1. Quartile  -0.018789
3. Quartile   0.043145
Mean         0.012179
Median       0.014968
Sum          10.486069
SE Mean      0.001867
LCL Mean     0.008514
UCL Mean     0.015844
Variance     0.003002
Stdev        0.054793
Skewness     -0.306964
Kurtosis     3.138812
> basicStats(da$sp)
      X..da.sp
nobs      861.000000
NAs        0.000000
Minimum    -0.239541
Maximum     0.163047
1. Quartile -0.018349
3. Quartile  0.035138
Mean        0.006171
Median      0.008965
Sum         5.312834
SE Mean     0.001454
LCL Mean    0.003317
UCL Mean    0.009024
Variance    0.001820
Stdev       0.042665
Skewness    -0.589276
Kurtosis    2.366863
> dim(da)
[1] 861  5
> da1=log(da[,2:5]+1)
> basicStats(da1$ge)
      X..da1.ge
nobs      861.000000
NAs        0.000000
Minimum    -0.318660
Maximum     0.224132
1. Quartile -0.031127
3. Quartile  0.047536
Mean        0.008318
Median      0.007092
Sum         7.161840
SE Mean     0.002240
LCL Mean    0.003921
UCL Mean    0.012715
Variance    0.004321

```

```

Stdev      0.065735
Skewness   -0.290782
Kurtosis    1.778316
> basicStats(dal$vw)
      X..dal.vw
nobs      861.000000
NAs        0.000000
Minimum    -0.255361
Maximum     0.153223
1. Quartile -0.016795
3. Quartile  0.037810
Mean        0.008331
Median      0.013266
Sum         7.172567
SE Mean     0.001491
LCL Mean    0.005405
UCL Mean    0.011256
Variance    0.001913
Stdev       0.043740
Skewness    -0.943052
Kurtosis    3.517912
> basicStats(dal$ew)
      X..dal.ew
nobs      861.000000
NAs        0.000000
Minimum    -0.317795
Maximum     0.261795
1. Quartile -0.018968
3. Quartile  0.042240
Mean        0.010611
Median      0.014857
Sum         9.136445
SE Mean     0.001876
LCL Mean    0.006930
UCL Mean    0.014293
Variance    0.003030
Stdev       0.055043
Skewness    -0.745712
Kurtosis    4.169659
> basicStats(dal$sp)
      X..dal.sp
nobs      861.000000
NAs        0.000000
Minimum    -0.273833
Maximum     0.151043
1. Quartile -0.018519
3. Quartile  0.034535
Mean        0.005234
Median      0.008925
Sum         4.506303

```

```

SE Mean      0.001470
LCL Mean     0.002349
UCL Mean     0.008119
Variance     0.001860
Stdev        0.043130
Skewness     -0.877898
Kurtosis     3.611268
> t.test(da1$ge)
      One Sample t-test
data:  da1$ge
t = 3.713, df = 860, p-value = 0.000218
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.003921037 0.012715061
sample estimates:
 mean of x
0.008318049

#### Problem 3 ####
> sp=da$sp
> t.test(sp)
      One Sample t-test
data:  sp
t = 4.2438, df = 860, p-value = 2.437e-05
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.003316703 0.009024375
sample estimates:
 mean of x
0.006170539

> T=length(sp)
> s=skewness(sp)/sqrt(6/T)
> s
[1] -7.059027
attr(,"method")
[1] "moment"
> k=kurtosis(sp)/sqrt(24/T)
> k
[1] 14.1765
attr(,"method")
[1] "excess"
#### Problem 4 ####
> da=read.table("d-axp3dx-0111.txt",header=T)
> axp=log(da$axp+1)
> T=length(axp)
> s=skewness(axp)/sqrt(6/T)
> s
[1] 0.4314821
attr(,"method")

```

```

[1] "moment"
> k=kurtosis(axp)/sqrt(24/T)
> k
[1] 92.7073
attr(,"method")
[1] "excess"

### Problem 5 ###
> da=read.table("d-jpus-0711.txt",header=T)
> head(da)
  year mon day   yen
1 2007   1   2 118.83
2 2007   1   3 119.58
...
6 2007   1   9 119.45
> dim(da)
[1] 1238   4
> rt=diff(log(da$yen))
> require(fBasics)
> basicStats(rt)

              rt
nobs          1237.000000
NAs            0.000000
Minimum       -0.052156
Maximum        0.030593
1. Quartile   -0.004517
3. Quartile    0.004075
Mean          -0.000345
Median        -0.000370
Sum           -0.426384
SE Mean        0.000214
LCL Mean      -0.000764
UCL Mean       0.000074
Variance       0.000056
Stdev          0.007515
Skewness      -0.416500
Kurtosis       4.835848
> t.test(rt)

      One Sample t-test

data:  rt
t = -1.6133, df = 1236, p-value = 0.1069
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -7.638670e-04  7.448262e-05
sample estimates:
mean of x
-0.0003446922
> d1=density(rt)
> plot(d1$x,d1$y,type='l')
>

```

```
> da=read.table("d-usuk-0711.txt",header=T)
> uk=diff(log(da$uk))
> basicStats(uk)
```

	uk
nobs	1237.000000
NAs	0.000000
Minimum	-0.049662
Maximum	0.044349
1. Quartile	-0.004204
3. Quartile	0.004058
Mean	-0.000185
Median	0.000000
Sum	-0.228465
SE Mean	0.000213
LCL Mean	-0.000602
UCL Mean	0.000233
Variance	0.000056
Stdev	0.007485
Skewness	-0.359374
Kurtosis	5.467822

Chapter 2

Linear Models for Financial Time Series

1. Problem 1: U.S. monthly unemployment rates

- (a) Yes, based on augmented Dickey-Fuller test, the monthly unemployment rates have a unit root. To perform the test, one first selects the AR order for the series. AIC selects $p = 11$. Using $p = 11$ so that the error-correction part has 10 lags, the ADF test statistic is -0.3522 with p -value 0.5041. Consequently, the unit-root hypothesis is not rejected.
- (b) Let r_t be the monthly unemployment rate and x_t be the first-differenced series of r_t . Figure 2.1 shows the sample ACF and PACF of x_t . Some features emerge. First, ACF has several significant lags at lower orders, which decay slowly with peak at lag 2. Second, PACF also has significant lags at lower orders with peak at lag 2. Third, both ACF and PACF show significant correlations at lags 12 and 24 (seasonal lags). Combining these features, we specified an multiplicative seasonal ARIMA(3,1,2)×(1,0,1)₁₂ model for the unemployment rates. The use of $p = 3$ is to allow for exponential decay and damping sine and cosine serial dependence. The fitted model is

$$(1 - 1.468B - 0.454B^2 - 0.107B^3)(1 - B)(1 - 0.585B^{12})r_t = (1 - 1.483B + 0.664B^2)(1 - 0.83B^{12})a_t, \quad (2.1)$$

where the residual variance is $\hat{\sigma}_a^2 = 0.0363$. All coefficient estimates are statistically significant at the 5% level. Figure 2.2 shows the model checking of the fitted model. Except for some possible outliers at the beginning of the series, the fitted model is adequate.

The 1-step to 4-step ahead predictions at November 2011 are 8.70, 8.72, 8.63, and 8.56, respectively. The standard errors of the predictions are 0.191, 0.268, 0.349, and 0.434, respectively.

- (c) Focus on the regular AR polynomial $\phi(x) = 1 - 1.468x + 0.454x^2 + 0.107x^3$. This polynomial has a real root and a pair of complex roots. Thus, there exist business cycles in the unemployment rates. The average period of the cycle is about 25 months (i.e., 2 years).
2. Problem 2: Monthly simple returns of Decile 2 and 10 portfolios from 1961 to September 2011.
- (a) Applying the Ljung-Box statistics with 12 lags, we have $Q(12) = 17.18$ with p -value 0.14 for Decile 2 returns and $Q(12) = 47.71$ with p -value 3.5×10^{-6} for Decile 10 returns. Therefore, Decile 2 returns have no serial correlations, but Decile 10 returns have significant serial correlations.
- (b) Even though part (a) with $Q(12)$ indicates no significant serial correlations in the returns of Decile 2 portfolio, a care examination of the sample ACF indicates that the lag-1 ACF is marginally significant at the 5% level. Therefore, we employ a MA(1) model for the returns. The fitted model is

$$r_t = 0.0093 + (1 + 0.131B)a_t, \quad \hat{\sigma}_a^2 = 0.00222. \quad (2.2)$$

Standard errors of the two parameters are 0.0022 and 0.0425, respectively. Thus, the two estimates are significant at the 5% level.

- (c) Based on the fitted MA(1) model, the 1-step to 12-step ahead forecasts are, respectively,

$$-0.0013, 0.0093, 0.0093, 0.0093, \dots, 0.0093,$$

and the associated standard errors of prediction are

$$0.0471, 0.0475, 0.0475, 0.0475, \dots, 0.0475.$$

Clearly, the predictions and their standard errors follow the property of an MA(1) model. Specifically, the predictions are mean-reverting after 1-step. The standard errors of forecasts approach the variance of the series, which is 0.0475.

3. Problem 3: Daily range of Apple stock. We downloaded the data via `quantmod` from Yahoo Finance. The range series has long-range dependence. This is evident from the sample ACF of the series shown in Figure 2.3. The ACF decays slowly and is statistically significant for large lags.

Using the package `fracdiff`, we obtain the model

$$(1 - B)^{0.338} r_t = a_t, \quad \hat{\sigma}_a = 2.862.$$

The fitted value of d is highly significant (differs from zero).

4. Problem 4: Monthly yields of Moody's Aaa bonds. Let x_t be the logarithm of the monthly Aaa bond yield at time index t . The time plot of x_t , not shown, seems to indicate the possibility of non-stationarity. To verify, we adopt the augmented Dickey-Fuller test. The `ar` command with Gaussian likelihood selects an AR(5) model for x_t . Using this model, the ADF test is -0.5895 with p -value 0.43 . Thus, the unit-root hypothesis is not rejected. In addition, the simple one-sample t -test shows that the mean of the differenced series $(1 - B)x_t$ is not significantly different from zero.

Figure 2.4 shows the sample ACF and PACF of $(1 - B)x_t$. From the plots, we make the following observations. First, the ACF suggests an MA(1) or MA(3) model. The lag-1 ACF is highly significant, but the lag-3 ACF is only marginal. Second, the PACF suggests an AR(2) model for the differenced series $(1 - B)x_t$. Consequently, we entertain three possible models below.

The fitted IMA(1) model is

$$(1 - B)x_t = (1 + 0.370B)a_t, \quad \hat{\sigma}_a^2 = 4.67 \times 10^{-4}. \quad (2.3)$$

Figure 2.5 shows the model checking of the IMA(1,1) model in Equation (2.3). The p -value plots indicates some minor serial correlations exist at lags 3 or 4.

The fitted IMA(3) model is

$$(1 - B)x_t = (1 + 0.377B - 0.013B^2 - 0.076B^3)a_t, \quad \hat{\sigma}_a^2 = 4.637 \times 10^{-4}.$$

Since the lag-2 coefficient is not statistically significant at the 5% level, we simplify the model as

$$(1 - B)x_t = (1 + 0.377B - 0.072B^3)a_t, \quad \hat{\sigma}_a^2 = 4.637 \times 10^{-4}. \quad (2.4)$$

Figure 2.6 shows the model checking of the IMA(1,3) model in Equation (2.4). From the plots, the fitted model is adequate except for a few possible outliers.

The fitted AR model is

$$(1 - 0.373B + 0.160B^2)(1 - B)x_t = a_t, \quad \hat{\sigma}_a^2 = 4.645 \times 10^{-4}. \quad (2.5)$$

Figure 2.7 shows the model checking of the integrated AR(2) model. From the plots, the model is also adequate. The possibility of outliers remains.

The three models are similar, but the AIC selects the IMA(3) model in Equation (2.4) as the best model, because the model has the smallest AIC, which is -5383.77 .

5. Problem 5: Moody's Aaa bond yield: exponential smoothing. Again, let x_t be the logarithm of the monthly Aaa bond yield. Since exponential

smoothing is a special case of the IMA(0,1,1) model, we fit the model to obtain the smoothing parameter. The fitted model is

$$(1 - B)x_t = (1 + 0.367B)a_t, \quad \hat{\sigma}_a^2 = 4.529 \times 10^{-4}.$$

This model is fitted using the first 1103 observations based on the specified forecast origin, which is November 2010. [**Note:** Strictly speaking, $\hat{\theta}$ should be positive for a conventional exponential smoothing model.] The 1-step to 12-step ahead forecasts, associated standard errors, and the observed logarithm of Aaa bond yield are given below:

	steps					
Variable	1	2	3	4	5	6
Forecast	1.594	1.594	1.594	1.594	1.594	1.594
StdError	0.021	0.036	0.046	0.055	0.062	0.068
Obs	1.613	1.617	1.652	1.635	1.641	1.601
Variable	7	8	9	10	11	12
Forecast	1.594	1.594	1.594	1.594	1.594	1.594
StdError	0.074	0.080	0.085	0.090	0.095	0.099
Obs	1.607	1.595	1.475	1.409	1.381	1.353

6. Problem 6: Aaa and Baa bond yields. If we entertain a linear regression model, we obtain

$$\ln(\text{Aaa}_t) = -0.359 + 1.081 \ln(\text{Baa}_t) + e_t,$$

where the standard error of the residuals is 0.1285 and the R^2 of the model is 91.78%. However, the sample ACF and time plot of the residuals indicate that the residual series is unit-root nonstationary. See Figure 2.8. Therefore, we consider the first differenced series.

Let $y_t = \text{diff}(\ln(\text{Baa}_t))$ and $z_t = \text{diff}(\ln(\text{Aaa}_t))$. The linear regression model is

$$z_t = 0.642y_t + \epsilon_t, \quad (2.6)$$

where the residual standard error is 0.0157 and the R^2 of the regression is 53.63%. Figure 2.9 shows the sample ACF and PACF of the residuals of Equation (2.6). From the PACF plot, an AR(2) is specified for the residuals. Consequently, we employ a linear regression model with an AR(2) residuals to study the relationship between the two bond yields. The fitted model is

$$(1 - 0.313B + 0.158B^2)(z_t - 0.623y_t) = a_t, \quad \hat{\sigma}_a^2 = 2.231 \times 10^{-4}, \quad (2.7)$$

where the coefficient estimates are all significant at the 5% level. Figure 2.10 shows the model checking of the model in Equation (2.7). From the plots, the linear regression model with AR(2) errors in Equation (2.7) is adequate. The model states that the log returns of the two bond yields