Booth School of Business, University of Chicago  
Business 41202, Spring Quarter 2010, Mr. Ruey S. Tsay

Midterm

ChicagoBooth Honor Code:  
I pledge my honor that I have not violated the Honor Code during this examination.

Signature: Name: ID:

Notes:  
- Open notes and books.  
- For each question, write your answer in the blank space provided.  
- Manage your time carefully and answer as many questions as you can.  
- The exam has 7 pages and the R output has 8 pages. Total is 15 pages. Please check to make sure that you have all the pages.  
- For simplicity, ALL tests use the 5% significance level.  
- Round your answer to 3 significant digits.

Problem A: (30 pts) Answer briefly the following questions. Each question has two points.

1. Give two reasons that the observed daily asset returns have significant serial correlations, even if the underlying true returns are serially uncorrelated.

2. Describe two methods to model seasonality in a financial time series.

3. For Questions 3 to 6, consider the daily closing values of the S&P 500 index from January 2, 2003 to April 19, 2010. Let sp5 denote the series of the logarithms of closing values. Is there a unit root in the sp5 series? Why?
4. Consider the daily log returns of the S&P 500 index. Based on the observed data, is the mean return significantly different from zero? Why?

5. Is the distribution of the daily log returns symmetric with respective to zero? Why?

6. Does the distribution of the daily log returns have heavy tails? Why?

7. For Questions 7 to 9, consider the simple AR(2) model
   \[ r_t = 0.01 + 1.3r_{t-1} - 0.4r_{t-2} + a_t, \quad \text{Var}(a_t) = 2.0. \]
   What is the mean of \( r_t \)? That is, find \( E(r_t) \).

8. Does the model imply the existence of business cycles? Why?

9. Suppose that \( r_{100} = 1.2 \) and \( r_{99} = 0.5 \). What is the 1-step ahead prediction of \( r_{101} \) at the forecast origin \( T = 100 \)? What is the 95% confidence interval of the prediction.

10. Describe two characteristics of the EGARCH models that are not available in the GARCH models.

11. For Questions 11 and 13, consider the following ARMA(0,0)-GARCH(1,1) model
   \[
   r_t = 0.03 + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim N(0, 1) \\
   \sigma_t^2 = 0.144 + 0.07a_{t-1}^2 + 0.83\sigma_{t-1}^2.
   \]
   What is the unconditional variance of \( a_t \)?
12. Suppose that $\sigma^2_{100} = 0.6$ and $a_{100} = -0.1$. Obtain the 95% confidence interval for the 1-step ahead prediction of $r_{101}$ at the forecast origin $T = 100$.

13. Obtain the 95% confidence interval for the 100-step ahead prediction of $r_{200}$ at the forecast origin $T = 100$.

14. Describe two methods to compare different forecasting models.

15. Describe briefly the difference between trend- and difference-stationarity in the long-term prediction.
Problem B. (37 pts) Consider the monthly log returns, in percentages, of the Coca-Cola stock from January 1960 to December 2009. The relevant R output is attached. Answer the following questions.

1. (2 points) Is the mean of the log return series equal to zero? Why?

2. (2 points) Is there any serial correlation in the monthly log returns? Why?

3. (2 points) Is there any ARCH effect in the monthly log returns? Why?

4. (3 points) A GARCH(1,1) model with Gaussian distribution is used for the volatility equation. Write down the fitted model, including the mean equation. Is the model adequate? Why?

5. (2 points) Except for the normality, is the fitted GARCH(1,1) model adequate? Why?

6. (3 points) A GARCH(1,1) model with Student-$t$ innovations is applied. Write down the volatility equation, including the degrees of freedom.

7. (3 points) A GARCH(1,1) model with skew Student-$t$ innovations is used. Write down the volatility equation, including the distribution parameters. Is the model adequate? Why?

8. (2 points) Let $\xi$ be the skew parameter. To check whether the distribution of the innovations is skewed, consider the test $H_o : \xi = 1$ versus $H_a : \xi \neq 1$. Perform the test and draw a conclusion.
9. (3 points) Focus on an IGARCH(1,1) model, write down the fitted model.

10. (3 points) A GJR model is also fitted. Write down the fitted volatility equation.

11. (2 points) Based on the GJR model, is the leverage effect significant? Why?

12. (2 points) Among the volatility models entertained so far, which model is preferred? State the criterion used in your choice.

13. (4 points) An EGARCH(1,1) model is also fitted, but to the log returns instead of the percentage log returns. Write down the fitted volatility equation. For simplicity, you may ignore the ARCH parameter, which is insignificant.

14. (2 points) Based on the fitted EGARCH(1,1) model, is the leverage effect significant? Why?

15. (2 points) Is the EGARCH model adequate in modeling the serial dependence in $r_t$ and $a_t^2$? Why?
Problem C. (11 pts) Consider the quarterly earnings per share of the 3M Company from 1992 to 2009. We analyzed the logarithms of the earnings. That is, \( x_t = \ln(y_t) \), where \( y_t \) is the quarterly earnings per share.

1. (2 points) Test \( H_0 : \rho_4 = 0 \) versus \( H_a : \rho_4 \neq 0 \), where \( \rho_4 \) is the lag-4 ACF of the differenced series of \( x_t \). Compute the test statistic and draw a conclusion.

2. (5 points) Write down the final fitted model for \( x_t \), including the residual variance.

3. (2 points) Is the fitted model adequate? Why?

4. (2 points) Obtain the 1-step and 2-step ahead forecasts of \( y_t \) (not the log earnings \( x_t \)) at the forecast origin December 2009.

5. (2 points) Let \( \theta_1 \) be the MA(1) coefficient. Test \( H_0 : \theta_1 = 0 \) versus \( H_a : \theta_1 \neq 0 \). Compute the test statistic and draw a conclusion.

Problem D. (22 pts) Consider the monthly series of the U.S. 30-year fixed mortgage rate from April 1971 to March 2010. The rate is generally believed to be related to the bank prime rate. To understand the relationship between the two rates, we consider some simple analysis. The R output is attached. Answer the following questions:
1. (2 points) Let $Y_t$ and $X_t$ be the monthly mortgage and prime rate, respectively. A simple linear regression model $Y_t = \beta_0 + \beta_1 X_t + \epsilon_t$ is employed. Write down the fitted model, including R-square and the residual standard error.

2. (2 points) Is the simple linear regression model adequate? Why?

3. (3 points) Let $y_t$ and $x_t$ be the first differenced series of $Y_t$ and $X_t$, respectively. Is the correlation coefficient between $y_t$ and $x_t$ significantly different from zero at the 5% level? Why?

4. (4 points) The residuals of the regression $y_t = \beta x_t + \epsilon_t$ show certain serial dependence. An AR(2) model is identified for the residuals, resulting in a regression model with time series errors. Write down the fitted model, including residual variance.

5. (3 points) Is the model in part (4) adequate? Why?

6. (3 points) Does the mortgage rate depend on the prime rate? Why?

7. (3 points) If pure AR models are entertained, an AR(3) model is specified for the differenced mortgage rate series. Write down the fitted AR(3) model, including residual variance.

8. (2 points) Use out-of-sample forecasts to compare the two models for the mortgage rates. Based on the output provided, select a model for the mortgage rate series. Explain your choice.
**Problem A: S&P 500 index**

```r
> da=read.table("d-sp55010.txt",header=T)
> x=da[1:1836,]  # select the time span.
> tail(x)

<table>
<thead>
<tr>
<th>Mon</th>
<th>Day</th>
<th>Year</th>
<th>Open</th>
<th>High</th>
<th>Low</th>
<th>Close</th>
<th>Volume</th>
<th>Adjclose</th>
</tr>
</thead>
<tbody>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1836</td>
<td>1</td>
<td>2</td>
<td>2003</td>
<td>879.82</td>
<td>909.03</td>
<td>879.82</td>
<td>909.03</td>
<td>1229200000</td>
</tr>
</tbody>
</table>
> spp=log(x[,9])
> sp5=rev(spp)  # reverse the series
> m1=ar(diff(sp5),method='mle')
> m1$order
[1] 12
> adfTest(sp5,lag=13)

Title: Augmented Dickey-Fuller Test

Test Results:

<table>
<thead>
<tr>
<th>PARAMETER:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lag Order: 13</td>
</tr>
<tr>
<td>STATISTIC:</td>
</tr>
<tr>
<td>Dickey-Fuller: 0.5828</td>
</tr>
<tr>
<td>P VALUE: 0.8021</td>
</tr>
</tbody>
</table>

> rtn=diff(sp5)
> basicStats(rtn)

<table>
<thead>
<tr>
<th>rtn</th>
</tr>
</thead>
<tbody>
<tr>
<td>nobs</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
<tr>
<td>Maximum</td>
</tr>
<tr>
<td>1. Quartile</td>
</tr>
<tr>
<td>3. Quartile</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Median</td>
</tr>
<tr>
<td>Sum</td>
</tr>
<tr>
<td>SE Mean</td>
</tr>
<tr>
<td>LCL Mean</td>
</tr>
<tr>
<td>UCL Mean</td>
</tr>
<tr>
<td>Variance</td>
</tr>
<tr>
<td>Stdev</td>
</tr>
<tr>
<td>Skewness</td>
</tr>
<tr>
<td>Kurtosis</td>
</tr>
</tbody>
</table>
```
**Problem B: Monthly returns of KO stock**

```r
> da=read.table("m-kosp6009.txt",header=T)
> ko=log(da$ko+1)*100 # Percentage log returns
> t.test(ko)
   One Sample t-test
   data: ko
   t = 4.3938, df = 599, p-value = 1.318e-05
   alternative hypothesis: true mean is not equal to 0

> Box.test(ko,lag=24,type='Ljung')
   Box-Ljung test
   data: ko
   X-squared = 25.5163, df = 24, p-value = 0.3782

> at=ko-mean(ko)
> Box.test(at^2,lag=24,type='Ljung')
   Box-Ljung test
   data: at^2
   X-squared = 194.0843, df = 24, p-value < 2.2e-16

> library(fGarch)
> m1=garchFit(~garch(1,1),data=ko,trace=F)
> summary(m1)
   Call:
   garchFit(formula = ~garch(1, 1), data = ko, trace = F)
   Mean and Variance Equation:
   data ~ garch(1, 1)
   [data = ko]
   Conditional Distribution:
   norm
   Estimate  Std. Error  t value Pr(>|t|)
   mu 1.29278 0.23313 5.545 2.94e-08 ***
   omega 2.65291 0.90419 2.934 0.00335 **
   alpha1 0.09407 0.02195 4.285 1.82e-05 ***
   beta1 0.83508 0.03420 24.420 < 2e-16 ***

---

Standardised Residuals Tests:

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jarque-Bera Test</td>
<td>R Chi^2</td>
<td>76.74663 0</td>
</tr>
<tr>
<td>Ljung-Box Test</td>
<td>R Q(10)</td>
<td>9.61664 0.4747453</td>
</tr>
<tr>
<td>Ljung-Box Test</td>
<td>R Q(20)</td>
<td>21.31853 0.3786019</td>
</tr>
</tbody>
</table>
```

---

9
> m2=garchFit(~garch(1,1),data=ko,trace=F,cond.dist="std")
> summary(m2)

Call:
garchFit(formula = ~garch(1,1),data=ko,cond.dist="std",trace = F)

Mean and Variance Equation:
data ~ garch(1, 1)
[data = ko]

Conditional Distribution:
std

Coefficient(s):

|                | Estimate | Std. Error | t value | Pr(>|t|) |
|----------------|----------|------------|---------|----------|
| mu             | 1.31748  | 0.21787    | 6.047   | 1.47e-09 *** |
| omega          | 2.30627  | 0.95826    | 2.407   | 0.016096 *   |
| alpha1         | 0.09990  | 0.02794    | 3.575   | 0.000350 *** |
| beta1          | 0.84071  | 0.03858    | 21.791  | < 2e-16 *** |
| shape          | 6.92073  | 1.82420    | 3.794   | 0.000148 *** |

---

Standardised Residuals Tests:

<table>
<thead>
<tr>
<th></th>
<th>Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jarque-Bera</td>
<td>78.30792</td>
<td>0</td>
</tr>
<tr>
<td>Ljung-Box Test</td>
<td>10.0325</td>
<td>0.4376465</td>
</tr>
<tr>
<td>Ljung-Box Test</td>
<td>21.46689</td>
<td>0.3701242</td>
</tr>
<tr>
<td>Ljung-Box Test</td>
<td>10.85525</td>
<td>0.3689062</td>
</tr>
<tr>
<td>Ljung-Box Test</td>
<td>12.69702</td>
<td>0.8900018</td>
</tr>
</tbody>
</table>

LM Arch Test R^2 9.3179 0.6755644

Information Criterion Statistics:

<table>
<thead>
<tr>
<th>AIC</th>
<th>BIC</th>
<th>SIC</th>
<th>HQIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.398308</td>
<td>6.427621</td>
<td>6.398220</td>
<td>6.409719</td>
</tr>
</tbody>
</table>

> m3=garchFit(~garch(1,1),data=ko,trace=F,cond.dist="sstd")
> summary(m3)
Call:
garchFit(formula = ~garch(1,1),data=ko,cond.dist="sstd",trace=F)

Mean and Variance Equation:
data ~ garch(1, 1)
[data = ko]

Conditional Distribution:
sstd

Coefficient(s):

|       | Estimate | Std. Error | t value | Pr(>|t|) |
|-------|----------|------------|---------|----------|
| mu    | 1.26669  | 0.23141    | 5.474   | 4.4e-08  *** |
| omega | 2.32235  | 0.95860    | 2.423   | 0.015408 * |
| alpha1| 0.09843  | 0.02750    | 3.579   | 0.000345 *** |
| beta1 | 0.84083  | 0.03854    | 21.814  | < 2e-16 *** |
| skew  | 0.96197  | 0.05720    | 16.817  | < 2e-16 *** |
| shape | 7.15816  | 1.97750    | 3.620   | 0.000295 *** |

Standardised Residuals Tests:

<table>
<thead>
<tr>
<th>Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jarque-Bera Test</td>
<td>R Chi^2 79.22959 0</td>
</tr>
<tr>
<td>Ljung-Box Test</td>
<td>R Q(10) 10.05494 0.4356862</td>
</tr>
<tr>
<td>Ljung-Box Test</td>
<td>R Q(20) 21.54316 0.3658051</td>
</tr>
<tr>
<td>Ljung-Box Test</td>
<td>R^2 Q(10) 10.98003 0.3590763</td>
</tr>
<tr>
<td>Ljung-Box Test</td>
<td>R^2 Q(20) 12.84643 0.8838753</td>
</tr>
<tr>
<td>LM Arch Test</td>
<td>R TR^2 9.461733 0.6630679</td>
</tr>
</tbody>
</table>

Information Criterion Statistics:

<table>
<thead>
<tr>
<th></th>
<th>AIC</th>
<th>BIC</th>
<th>SIC</th>
<th>HQIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6.363256</td>
<td>6.407225</td>
<td>6.363058</td>
<td>6.380372</td>
</tr>
</tbody>
</table>

> source("garchoxfit_R.txt")
> m4=garchOxFit(formula.mean=~arma(0,0),formula.var=~igarch(1,1),series=ko, cond.dist='t')

***********************
** SPECIFICATIONS **
***********************

Dependent variable : X
Mean Equation : ARMA (0, 0) model.
Variance Equation : IGARCH (1, 1) model.
The distribution is a Student distribution, with 5.35014 degrees of freedom.
Maximum Likelihood Estimation (Std.Errors based on Second derivatives)

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std.Error</th>
<th>t-value</th>
<th>t-prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cst(M)</td>
<td>1.318261</td>
<td>0.21560</td>
<td>6.114</td>
</tr>
<tr>
<td>Cst(V)</td>
<td>1.152686</td>
<td>0.57519</td>
<td>2.004</td>
</tr>
<tr>
<td>ARCH(Alpha1)</td>
<td>0.140729</td>
<td>0.031861</td>
<td>4.417</td>
</tr>
<tr>
<td>Student(DF)</td>
<td>5.350140</td>
<td>1.2196</td>
<td>4.387</td>
</tr>
<tr>
<td>GARCH(Beta1)</td>
<td>0.859271</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

No. Observations : 600  No. Parameters : 5
Mean (Y) : 1.12225  Variance (Y) : 39.07788
Skewness (Y) : -0.51509  Kurtosis (Y) : 5.69349

** TESTS **

<table>
<thead>
<tr>
<th>Statistic</th>
<th>t-Test</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness</td>
<td>-0.34891</td>
<td>3.4978</td>
</tr>
<tr>
<td>Excess Kurtosis</td>
<td>1.8038</td>
<td>9.0564</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>93.517</td>
<td>NaN</td>
</tr>
</tbody>
</table>

Information Criterium (to be minimized)

Akaike       6.366363  Shibata       6.366275
Schwarz      6.395675  Hannan-Quinn  6.377774

Q-Statistics on Standardized Residuals
Q( 10) = 12.2181  [0.2707240]
Q( 20) = 22.9868  [0.2894423]
H0 : No serial correlation ==> Accept H0 when prob. is High [Q < Chisq(lag)]

Q-Statistics on Squared Standardized Residuals
--> P-values adjusted by 2 degree(s) of freedom
Q( 10) = 9.08732  [0.3349834]
Q( 20) = 10.8474  [0.9007201]
H0 : No serial correlation ==> Accept H0 when prob. is High [Q < Chisq(lag)]

ARCH 1-2 test:  F(2,593) = 1.5588  [0.2112]
ARCH 1-10 test: F(10,577)= 0.69136  [0.7330]

> m5=garchOxFit(formula.mean="arma(0,0),formula.var="gjr(1,1),series=ko,cond.dist='t')

Dependent variable : X
Mean Equation : ARMA (0, 0) model.
Variance Equation: GJR (1, 1) model.
The distribution is a Student distribution, with 7.05477 degrees of freedom.

Maximum Likelihood Estimation (Std.Errors based on Second derivatives)

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std.Error</th>
<th>t-value</th>
<th>t-prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cst(M)</td>
<td>1.294115</td>
<td>0.21935</td>
<td>5.900</td>
</tr>
<tr>
<td>Cst(V)</td>
<td>2.628110</td>
<td>1.1362</td>
<td>2.313</td>
</tr>
<tr>
<td>ARCH(Alpha1)</td>
<td>0.072563</td>
<td>0.037686</td>
<td>1.925</td>
</tr>
<tr>
<td>GARCH(Beta1)</td>
<td>0.832876</td>
<td>0.042561</td>
<td>19.57</td>
</tr>
<tr>
<td>GJR(Gamma1)</td>
<td>0.047660</td>
<td>0.053171</td>
<td>0.8963</td>
</tr>
<tr>
<td>Student(DF)</td>
<td>7.054772</td>
<td>1.9079</td>
<td>3.698</td>
</tr>
</tbody>
</table>

No. Observations: 600 No. Parameters: 6

Mean (Y): 1.12225 Variance (Y): 39.07788
Skewness (Y): -0.51509 Kurtosis (Y): 5.69349

************
** TESTS **
************

<table>
<thead>
<tr>
<th>Statistic</th>
<th>t-Test</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness</td>
<td>-0.33265</td>
<td>3.3348</td>
</tr>
<tr>
<td>Excess Kurtosis</td>
<td>1.4462</td>
<td>7.2611</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>63.354</td>
<td>.NaN</td>
</tr>
</tbody>
</table>

Information Criterium (to be minimized)

Akaike 6.362520 Shibata 6.362322
Schwarz 6.406489 Hannan-Quinn 6.379636

Q-Statistics on Standardized Residuals

Q(10) = 9.73300  [0.4642223]
Q(20) = 21.0652  [0.3933024]
H0: No serial correlation ==> Accept H0 when prob. is High [Q < Chisq(lag)]

Q-Statistics on Squared Standardized Residuals

--> P-values adjusted by 2 degree(s) of freedom

Q(10) = 9.00421  [0.3419413]
Q(20) = 11.1132  [0.8894894]
H0: No serial correlation ==> Accept H0 when prob. is High [Q < Chisq(lag)]

ARCH 1-2 test: F(2,593) = 1.4585  [0.2334]
ARCH 1-10 test: F(10,577) = 0.74364  [0.6834]

> source("garchoxfit_R_w.txt")
> ko=ko/100
> m6=garch0xFit(formula.mean=~arma(0,0),formula.var=~egarch(1,1),series=ko)

********************
** SPECIFICATIONS **
********************
Dependent variable : X
Mean Equation : ARMA (0, 0) model.
Variance Equation : EGARCH (1, 1) model.
The distribution is a Gauss distribution.

Maximum Likelihood Estimation (Std.Errors based on Second derivatives)

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std.Error</th>
<th>t-value</th>
<th>t-prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cst(M)</td>
<td>0.011263</td>
<td>0.0023541</td>
<td>4.784</td>
</tr>
<tr>
<td>Cst(V)</td>
<td>-56141.0888</td>
<td>914.23</td>
<td>-61.41</td>
</tr>
<tr>
<td>ARCH(Alpha1)</td>
<td>0.767725</td>
<td>0.72315</td>
<td>1.062</td>
</tr>
<tr>
<td>GARCH(Beta1)</td>
<td>0.829261</td>
<td>0.058293</td>
<td>14.23</td>
</tr>
<tr>
<td>EGARCH(Theta1)</td>
<td>-0.050112</td>
<td>0.036360</td>
<td>-1.378</td>
</tr>
<tr>
<td>EGARCH(Theta2)</td>
<td>0.156142</td>
<td>0.059092</td>
<td>2.642</td>
</tr>
</tbody>
</table>

No. Observations : 600 No. Parameters : 6
Mean (Y) : 0.01122 Variance (Y) : 0.00391
Skewness (Y) : -0.51509 Kurtosis (Y) : 5.69349

Warning : To avoid numerical problems, the estimated parameter Cst(V), and its std.Error have been multiplied by 10^-4.

** TESTS **

<table>
<thead>
<tr>
<th>Statistic</th>
<th>t-Test</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness</td>
<td>-0.35709</td>
<td>3.5798</td>
</tr>
<tr>
<td>Excess Kurtosis</td>
<td>1.3494</td>
<td>6.7748</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>58.271</td>
<td>.NaN</td>
</tr>
</tbody>
</table>

Information Criterium (to be minimized)
Akaike                   -2.777973
Shibata                  -2.778170
Schwarz                  -2.734004
Hannan-Quinn             -2.760857

Q-Statistics on Standardized Residuals
Q( 10) = 8.56962 [0.5733761]
Q( 20) = 21.8249 [0.3500881]
H0 : No serial correlation ==> Accept H0 when prob. is High [Q < Chisq(lag)]

Q-Statistics on Squared Standardized Residuals
--> P-values adjusted by 2 degree(s) of freedom
Q( 10) = 9.87575 [0.2738554]
Q(20) = 15.4734 [0.6292488]
H0: No serial correlation ==> Accept H0 when prob. is High [Q < Chisq(lag)]

ARCH 1-2 test: F(2,593) = 0.55584 [0.5739]
ARCH 1-10 test: F(10,577) = 0.87431 [0.5572]

***** Problem C: 3M earnings ***
> da=read.table("q-3m-earns9209.txt")
> mmm=da[,4]
> mmm=log(mmm) # Take log transformation
> dm=diff(diff(mmm),4)

> f1=acf(dm,lag=20)
> f1$acf[1:9]
[1] 1.00  0.0946 -0.2625 -0.1517 -0.1188  0.0044 -0.0904 -0.0790 -0.0458
> length(dm)
[1] 67

> mml=arima(mmm,order=c(0,1,2),seasonal=list(order=c(0,1,0),period=4))
> tsdiag(mml)
> Box.test(mml$residuals,lag=20,type='Ljung')
Box-Ljung test
data: mml$residuals
X-squared = 16.4748, df = 20, p-value = 0.6868

> Box.test(mml$residuals^2,lag=20,type='Ljung')
Box-Ljung test
data: mml$residuals^2
X-squared = 26.2933, df = 20, p-value = 0.1563

> mml
Call:
arima(x=mmm,order=c(0,1,2),seasonal=list(order=c(0,1,0),period=4))

Coefficients:
ma1    ma2
0.1831 -0.8169
s.e. 0.1059 0.1021

sigma^2 estimated as 0.006934: log likelihood = 69.4, aic = -132.8
> predict(mml,8)
$pred

15
Time Series:
Start = 73
End = 80
Frequency = 1
[1] -0.2306 -0.3048 0.0883 0.2207 -0.3247 -0.3989 -0.0058 0.1267

$se$
Time Series:
Start = 73
End = 80
Frequency = 1
[1] 0.0839 0.1290 0.1326 0.1360 0.1777 0.2193 0.2276 0.2356

**** Problem D: Mortgage rate *****
> da=read.table("m-mortg.txt",header=T)
> da1=read.table("m-prime.txt",header=T)
> mort=da$rate
> prim=da1$rate
> plot(prim,mort)
> m1=lm(mort~prim)
> summary(m1)

Call:
  lm(formula = mort ~ prim)

Coefficients:  
            Estimate Std. Error t value  Pr(>|t|)  
(Intercept) 2.56445    0.17322   14.80 <2e-16 ***  
 prim        0.75734    0.01907   39.71 <2e-16 ***  
---
Residual standard error: 1.35 on 466 degrees of freedom  
Multiple R-squared: 0.7719,   Adjusted R-squared: 0.7714  

> Box.test(m1$residuals,lag=12,type='Ljung')

Box-Ljung test
data: m1$residuals
X-squared = 3457.820,  df = 12,  p-value < 2.2e-16  

> dm=diff(mort)
> dp=diff(prim)
> m2=lm(dm~-1+dp)
> summary(m2)

Call:
lm(formula = dm ~ -1 + dp)

Coefficients:
   Estimate Std. Error t value Pr(>|t|)   
  dp 0.36018     0.02146   16.79 <2e-16 ***  
---  
Residual standard error: 0.2371 on 466 degrees of freedom
Multiple R-squared: 0.3768,    Adjusted R-squared: 0.3755

> mm2 = ar(m2$residuals, method='mle')
> mm2$order   # identify the order
[1] 2

> m3 = arima(dm, order=c(2,0,0), xreg=dp, include.mean=F)
> m3
Call:
  arima(x = dm, order = c(2, 0, 0), xreg = dp, include.mean = F)

Coefficients:
      ar1     ar2     dp
    0.3988 -0.2570  0.3311
s.e.  0.0451  0.0453  0.0233

sigma^2 estimated as 0.04727:  log likelihood = 49.83, aic = -91.65
> tsdiag(m3, gof=36)
> Box.test(m3$residuals, lag=24, type='Ljung')

Box-Ljung test

data:  m3$residuals
X-squared = 26.7422, df = 24, p-value = 0.3166

> mm3 = ar(dm, method='mle')
> mm3$order
[1] 3
> t.test(dm)

One Sample t-test
data:  dm
t = -0.3605, df = 466, p-value = 0.7186
alternative hypothesis: true mean is not equal to 0

> m4 = arima(dm, order=c(3,0,0), include.mean=F)
> m4
Call:
  arima(x = dm, order = c(3, 0, 0), include.mean = F)
Coefficients:
\[
\begin{array}{ccc}
ar1 & ar2 & ar3 \\
0.5791 & -0.3856 & 0.1101 \\
s.e. & 0.0460 & 0.0501 & 0.0459 \\
\end{array}
\]

\[
sigma^2 \text{ estimated as } 0.06607: \text{ log likelihood } = -28.44, \text{ aic } = 64.88
\]

\[
> \text{Box.test(m4$residuals,lag=24,type='Ljung')}
\]

Box-Ljung test
\[
data: \ m4$residuals
X-squared = 22.9252, df = 24, p-value = 0.5242
\]

\[
> \text{source("backtest.R")}
\]

\[
> \text{ backtest(m3, dm, 400, 1, xre=dp)}
\]

[1] "RMSE of out-of-sample forecasts"
[1] 0.1977504

[1] "Mean absolute error of out-of-sample forecasts"
[1] 0.1408496

\[
> \text{backtest(m4, dm, 400, 1)}
\]

[1] "RMSE of out-of-sample forecasts"
[1] 0.2030052

[1] "Mean absolute error of out-of-sample forecasts"
[1] 0.1435424