Solutions to Midterm

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

1. Give two situations under which serial correlations exist in observed asset returns even though the true underlying returns are serially uncorrelated.
   Answer: Any two of (a) data smoothing, (b) bid-ask bounce, and (c) nonsynchronous trading.

2. (Questions 2 to 8): Consider the daily S&P 500 index for a certain period. Some analysis is attached. Let \( r_t \) be the daily log return of the index. Is the expected mean return \( E(r_t) \) zero? Why?
   Answer: The expected return is not significantly different from zero because the 95% confidence interval contains zero.

3. Does the daily log return of the S&P index follow a skew distribution? Why?
   Answer: Yes. \( t = -0.281 / \sqrt{6/1335} = -4.19 \), which is less than \(-1.96\) so that the hypothesis of symmetric distribution is rejected.

4. Does the daily log return of the S&P index have heavy tails? Why?
   Answer: Yes. \( t = 4.24 / \sqrt{24/1335} = 31.62 \), which is highly significant.

5. The sample ACF of \( r_t \), namely \( \hat{\rho}_0, \hat{\rho}_2, \ldots, \hat{\rho}_9 \), are given. Test the null hypothesis \( H_0: \rho_1 = 0 \) versus the alternative hypothesis \( H_a: \rho_1 \neq 0 \), where \( \rho_1 \) is lag-1 ACF of \( r_t \). Compute the test statistic and draw the conclusion.
   Answer: The t-ratio is \( t = -0.0855 / \sqrt{1/1335} = -3.124 \), which is greater than 1.96. Thus, the lag-1 serial correlation is statistically significant.

6. Turn to the daily log index \( p_t \). A model is fitted, called \( m2 \), in the R output. Write down the fitted model, including residual variance.
   Answer: \((1 - B)p_t = a_t - 0.0788a_{t-1}\), where \( \sigma^2 = 0.000144 \).

7. Use the fitted model \( m2 \) to forecast the log index at the forecast origin \( T = 1336 \). What is the 1-step ahead point forecast? Obtain a 95% interval forecast for \( p_{1337} \).
   Answer: \( p_{1336}(1) = p_{1336} - 0.0788a_{1336} = 7.53 - 0.0788(-0.008) = 7.531 \).
8. What is the 2-step ahead point forecast of $p_t$ at the forecast origin $T = 1336$? Use the model to derive the forecast.

Answer: $p_{1338} = p_{1337} + a_{1338} - 0.0788a_{1337} = p_{1336} + a_{1337} - 0.0788a_{1336} + a_{1338} - 0.0788a_{1337}$. Therefore, $p_{1336}^{(2)} = p_{1336} - 0.0788(0.008) = 7.531$.

9. Let $R_t$ and $r_t$ be the daily simple and log return, respectively, of an asset. What is the relationship between $R_t$ and $r_t$? Suppose further that $r_t$ follows a normal distribution with mean 0.05 and variance 0.04. What is the expected value of $R_t$ for the asset?

Answer: $r_t = \log(R_t + 1)$ or $R_t = \exp(r_t) - 1$. $E(R_t) \exp(0.05 + 0.5 \times 0.04) - 1 = 0.0725$.

10. Consider the monthly log return, in percentages, of the Decile 8 portfolio of Center for Research in Security Prices (CRSP) from January 1961 to December 2013 for 636 observations. A GARCH-M model is fitted to the series. Write down the fitted model.

Answer: The mean equation is $r_t = -1.263 + 0.0625\sigma_t^2 + a_t$ with $a_t = \sigma_t \epsilon_t$, where $\epsilon_t \sim iid N(0,1)$. The volatility equation is

$$\sigma_t^2 = 4.075 + 0.0701a_{t-1}^2 + 0.829a_{t-1}^2.$$ 

11. Consider again the monthly log returns, in percentages, of Decile 8 portfolio. Is the risk premium statistically significant at the 5% level? Why?

Answer: The $t$-ratio of $\gamma$ is 2.16 with $p$-value 0.031 so that the risk premium is statistically significant at the 5% level.

12. (For questions 12-15). Consider the growth rates of the real quarterly gross domestic product (GDP) of Canada from the second quarter of 1980 to the second quarter of 2011 for 125 data points. Figure ?? shows the PACF of the GDP growth rates. Specify two possible AR models for the growth rate series and briefly justify your choices.

Answer: an AR(1) or an AR(4) model because lag-1 PACF is significant whereas lag-4 PACF is marginally significant.

13. The order selection via AIC is also given. The criterion selects an AR(4) model. An AR(4) model is estimated. Write down the fitted model, including the residual variance.

Answer: $(1 - 0.649B + 0.176B^2 - 0.233B^3 + 0.207B^4)(r_t - 0.006) = a_t$, where $\sigma^2 = 3.898 \times 10^{-5}$.

14. Consider the fitted AR(4) model. Does it imply the existence of business cycles in the Canadian economy? Why?

Answer: Yes, because there are two pairs of complex solutions.
15. If business cycles exist, compute the periods of all possible cycles.

Answer: The average periods are 13.08 and 3.07 quarters, respectively.

**Problem B.** (27 pts) Consider the daily log returns of Apple stock starting from January 3, 2004 for 2517 observations. Let \( r_t \) be the log return series. Based on the attached R output, answer the following questions.

1. (3 points) What is the mean equation for \( r_t \)? Why?

Answer: The mean equation is \( r_t = \mu + a_t \) because there is no serial correlation in the log return.

2. (2 points) Is there any ARCH effect in \( r_t \)? Why?

Answer: The ARCH test shows \( Q(10) = 432.37 \) for the \( a_t^2 \) process. The associated \( p \)-value is close to zero.

3. (2 points) A simple volatility model, called \( m1 \) in R, is entertained for \( r_t \). Is Model \( m1 \) adequate for the log return series? Why?

Answer: No, the model \( m1 \) is not adequate because the normality assumption is rejected.

4. (3 points) A refined model, called \( m2 \) in R, is fitted. Write down the model, including the distribution of the innovations.

Answer: The model is \( r_t = 0.00188 + a_t, \ a_t = \sigma_t \epsilon_t, \ \epsilon_t \ are \ iid \ standardized \ Student-t \ with \ 5.32 \ degrees \ of \ freedom \). The volatility model is \( \sigma_t^2 = 5.97 \times 10^{-6} + 0.052a_{t-1}^2 + 0.938\sigma_{t-1}^2 \).

5. (3 points) Let \( \xi \) be the skew parameter in Model \( m3 \). Based on the model, is the distribution of \( r_t \) skew? Perform a statistical test to support your answer.

Answer: The \( t \)-ratio is \( t = (1.033 - 1)/0.0285 = 1.16 \), which is less than 1.96 so that the null hypothesis of symmetric distribution cannot be rejected at the 5% level.

6. (2 points) Compare the three models \( m1, m2, m3 \). Which model is preferred? Why?

Answer: \( m2 \) is preferred because it has the smallest AIC value.

7. (3 points) To estimate the potential leverage effect in \( r_t \), we consider an APARCH(1,1) model with \( \delta = 2.0 \). Write down the fitted model, including the innovation distribution.

Answer: The fitted model is

\[
\begin{align*}
    r_t &= 0.00176 + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim t^{*}_{5.46} \\
    \sigma_t^2 &= 1.087 \times 10^{-5} + 0.0649(|a_{t-1}| - 0.305a_{t-1})^2 + 0.912\sigma_{t-1}^2.
\end{align*}
\]
8. (4 points) The average volatility of $r_t$ via the APARCH model is 0.02248 and the approximate 99th quantile of $r_t$ is 0.061758 resulting in an $a_t = 0.06$. To see the impact of leverage effect, (a) compute the volatility $\sigma_t$ if $a_{t-1} = 0.06$ and $\sigma_{t-1} = 0.02248$, (b) compute the volatility $\sigma_t$ if $a_{t-1} = -0.06$ and $\sigma_{t-1} = 0.02248$, and finally, (c) compute volatility ratio $[(b)/(a)]$.

Answer: (a) $\sigma_t^2 = 1.087 \times 10^{-5} + 0.0649(0.06 - 0.305 \times 0.06)^2 + 0.912 \times 0.02248^2 = 0.000585$. Therefore, $\sigma_t = 0.0242$. (b) $\sigma_t^2 = 1.087 \times 10^{-5} + 0.0649(0.06 + 0.306 \times 0.06)^2 + 0.912 \times 0.02248^2 = 0.00087$. Therefore, $\sigma_t = 0.0295$. (c) The ratio is 1.22.

9. (2 points) An IGARCH model with normal innovations is also fitted for the Apple log return $r_t$. Write down the fitted model.

Answer: $r_t = 0.00212 + a_t$, $a_t = \sigma_t \epsilon_t$ with $\epsilon_t$ being iid N(0,1). The volatility equation is $\sigma_t^2 = (1 - 0.961)a_{t-1}^2 + 0.961\sigma_{t-1}^2$.

10. (3 points) Using the fitted IGARCH(1,1) model and the information provided, compute the volatility $\sigma_{2518}$ for the Apple log return.

Answer: The prediction is $\sigma_{2517}^2(1) = (1 - 0.961)(0.0117 - 0.00212)^2 + 0.961(0.0232)^2 = 0.00171$ so that $\sigma_{2518}(1) = 0.0131$.

Problem C. (17 points) Consider the monthly log return of Decile 1 portfolio of CRSP from January 1961 to December 2013 for 636 observations. Let $d1_t$ denote the monthly log return. Several volatility models were fitted. Use the attached R output to answer the following questions.

1. (4 points) Both the GARCH(1,1) model with Gaussian innovations, $g_1$, and the GARCH(1,1) model with Student-t innovations, $g_2$, were rejected based on model checking. A refined model, called $g_3$, is entertained. Write down the fitted model, including the mean equation and the innovation distribution.

$$ r_t = 0.00797 + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim t^*_{7.07,0.784} $$

$$ \sigma_t^2 = 8.50 \times 10^{-6} + 0.138a_{t-1}^2 + 0.824\sigma_{t-1}^2, $$

where $t^*_{v,s}$ denotes a skew standardized Student-t distribution with degrees of freedom $v$ and skew parameter $s$.

2. (2 points) Based on the fitted model $g_3$, is the distribution of the log returns skew? Why?

Answer: Yes. The $t$-ratio is $t = (1 - 0.784)/0.0471 = 4.59$, which is highly significant.
3. (3 points) Based on the model $g_3$, compute a 95% 4-step ahead interval forecast for the log return of Decile 1 portfolio at the forecast origin December 2013.

Answer: The 95% interval forecast is $0.00797 \pm 1.96 \times 0.0347$. That is, $(-0.0600, 0.0760)$.

4. (3 points) To study the leverage effect, a TGARCH or GJR-type of model is entertained. Denote the model by $g_4$. Based on the model, is the leverage effect significant? State the null and alternative hypotheses, obtain the test statistic, and draw the conclusion.

Answer: $H_0 : \gamma_1 = 0$ versus $H_a : \gamma_1 \neq 0$. The $t$-ratio is 2.018 with $p$-value 0.0436. Thus, $H_0$ is rejected at the 5% level. The leverage effect is statistically significant.

5. (3 points) Based on the model $g_4$, compute a 95% 4-step ahead interval forecast for the log return of Decile 1 portfolio at the forecast origin December 2013.

Answer: The 95% interval forecast is $0.00737 \pm 1.96 \times 0.0339$. That is, $(-0.0591, 0.0738)$.

6. (2 points) Compare the two 95% interval forecasts. Briefly state the impact of leverage effect?

Answer: Modeling the leverage effect shortens the length of the interval forecast.

Problem D. (14 points) Consider the monthly U.S. heating oil price and the natural gas price from November 1993 to August 2012. Use the attached R output to answer the following questions:

1. (2 points) Focus on the logarithm of the heating oil price. Preliminary analysis shows that the log price has a unit root so that the growth rate is used in model specification. The AIC selects an AR(1) for the growth rate. Therefore, an ARIMA(1,1,0) model is entertained for the log heating price. Write down the fitted model, including residual variance.

Answer: $(1 - 0.203B)(1 - B)h_t = a_t$, where $\sigma^2 = 0.00706$ and $h_t$ denotes the logarithm of heating oil price.

2. (2 points) Since the fitted AR(1) coefficient is not large, we also entertained an exponential smoothing model. Write down the fitted model, including the residual variance.

Answer: $(1 - B)h_t = a_t + 0.183a_{t-1}$ with $\sigma^2 = 0.00709$. 

5
3. (2 points) Model checking shows that the prior two models fit the data reasonably well. Based on in-sample fit, which model is preferred? Why?

Answer: The ARIMA(1,1,0) model is preferred because it has smaller AIC value.

4. (2 points) The two models were used in out-of-sample forecasting. Based on the out-of-sample performance, which model is preferred? Why?

Answer: Again, the ARIMA(1,1,0) model is preferred because it has lower RMSE and MAE in out of sample forecasts.

5. (3 points) Next, to make use of the information in the natural gas price, we consider a simple linear regression between the log heating oil price and log natural gas price. The residuals of the regression model shows strong serial correlations. To avoid spurious regression, let $y_t$ and $x_t$ be the growth rate of heating oil price and natural gas price, respectively. Write down the simple linear regression for the two growth rate series. What is the $R^2$ of the model?

Answer: $y_t = 0.21x_t + \epsilon_t$ with standard error of $\epsilon_t$ being 0.0805. The $R^2$ of the model is 12.4%.

6. (3 points) The residuals of the prior simple linear regression contains significant lag-1 serial correlation so that a regression model with time series errors is fitted. Write down the fitted model.

Answer: $(1 - 0.192B)(y_t - 0.202x_t) = a_t$ with $\sigma^2 = 0.00622$.

**Problem E.** (12 points) Consider the quarterly earnings per share of Procter & Gamble from 1983.II to 2012.III. Figure 2 shows the time pot of the earnings. From the plot, there was a negative earnings in the 80s and two large jumps occurred around 2010. For simplicity, we analyze the earnings $x_t$ directly. Sample autocorrelations of differenced data suggest the Airline model.

1. (2 points) Write down the fitted time series model $m1$ for the $x_t$ series, including the residual variance.

   Answer: $(1 - B)(1 - B^4)x_t = (1 - 0.71B)(1 - 0.52B^4)a_t$ with $\sigma^2 = 0.00673$.

2. (2 points) The fitted model show a large outlier at $t = 104$. Define an indicator variable for this particular data point.

   Answer: $I_t^{(104)} = \begin{cases} 
   1 & \text{if } t = 104 \\
   0 & \text{otherwise.} 
\end{cases}$
3. (2 points) As a matter of fact, there are several outliers. The model m4 contains three large outliers. Model checking shows that the ACF of the residuals has a significant correlation at lag 3 so that a refined model is entertained. The resulting model is denoted by m5. Is the lag-3 MA coefficient $\theta_3$ of Model m5 significantly different from zero? Why?

Answer: Yes, the $t$-ratio is $t = -0.0743/0.101 = -7.36$, which is highly significant.

4. (6 points) Finally, an additional outlier is found and an insignificant parameter is also detected. The final model for $x_t$ is Model m7. Write down the fitted model, including residual variance.

Answer: The fitted model is

$$(1 - B)(1 - B^4)[x_t - 0.539I_t^{(104)} - 0.42I_t^{(108)} + 0.17I_t^{(18)} - 0.15I_t^{(102)}] =$$

$$(1 - 0.276B^2 - 0.323B^3)(1 - 0.335B^4)a_t,$$

where $\sigma^2 = 0.000688$. 