Solutions to Midterm

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

1. **For Questions 1 to 5**, consider the monthly simple returns of CRSP Decile 10 portfolio from January 1970 to December 2008. The basic statistics of the data are given in the attached R output. Is the mean return of Decile 10 portfolio significantly different from zero? Why?
   
   Answer: Yes, the 95% C.I. for the mean returns does not cover zero.

2. What is the unconditional standard error of the simple monthly returns?
   
   Answer: 0.0431.

3. Are the monthly simple returns symmetric with respect to their mean value? Why?
   
   Answer: Perform the skewness test. The \( t \)-ratio is \( \frac{-0.421}{\sqrt{6/468}} = -3.718 \) with p-value 0.0002. No, the simple returns are not symmetric. They are skewed to the left.

4. Do the monthly simple returns have heavy tails? Why?
   
   Answer: Yes, because the \( t \)-ratio fo the excess kurtosis is 9.45 with p-value close to zero.

5. Are the monthly simple returns serially correlated? Why?
   
   Answer: No, the Ljung-Box statistic shows \( Q(10) = 10.687 \) with p-value 0.56.

6. **For Questions 6 and 7**, consider the monthly simple returns of the Decile 9 portfolio. As shown in homework assignment, the series follows the model \( r_t = 0.0099 + a_t + 0.155a_{t-1} \), \( \sigma_a^2 = 0.00242 \). Based on the model, what is the expected mean return of the Decile 9 portfolio?
   
   Answer: 0.0099, i.e., 0.99% per month.

7. Again, based on the model, what is the unconditional variance of the monthly simple returns of Decile 9 portfolio?
   
   Answer: \( \text{Var}(r_t) = (1 + 0.155^2)\sigma_a^2 = 0.00248 \).

8. What is the null hypothesis when one applies the Ljung-Box test \( Q(15) \) to a time series \( x_t \)?
   
   Answer: \( H_0 : \rho_1 = \rho_2 = \cdots = \rho_{15} = 0 \), where \( \rho_i \) is the lag-i ACF of \( x_t \).

9. **For Questions 9 and 10**, consider the quarterly earnings per share of Intel Corporation from the first quarter of 1992 to the first quarter of 2009. The series follows approximately the AR(1) model \( (1 - 0.81B)(r_t - 0.19) = a_t \), \( \sigma_a^2 = 0.0044 \). It is also known
that the company earned $0.04 in the first quarter of 2009. Based on the model, obtain the forecast of the earnings per share for the second quarter of 2009.

Answer: From the model, the 1-step ahead forecast is \( r_T(1) = 0.81r_T + 0.19(1 - 0.81) = 0.0685 \approx 0.07 \). That is, $0.07.

10. What is the standard deviation of the associated forecast error for the Intel earnings prediction?
Answer: standard error is \( \sqrt{0.0044} = 0.0663 \approx 0.07 \).

11. Both the EGARCH and GJR models can handle asymmetric responses of asset volatility to prior positive and negative innovations. Give a main difference between these two models.
Answer: Any one of (a) EAGARCH employs \( \ln(\sigma_t^2) \) whereas GJR uses \( \sigma_t^2 \) directly, (b) GJR is simpler in estimation, and (c) EGARCH is easier to assess the impact of asymmetry.

12. **For Questions 12 to 13**, consider a the GARCH(1,1) model

\[
egin{align*}
    r_t &= 0.0075 + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim iid(0, 1) \\
    \sigma_t^2 &= 0.001 + 0.853 \sigma_{t-1}^2 + 0.125 a_{t-1}^2.
\end{align*}
\]

Suppose that at the forecast origin \( t = h \), \( \sigma_h = 0.053 \) and \( a_h = -0.012 \). Compute the volatility forecast for time \( t = h + 1 \).

Answer: \( \sigma_h^2(1) = 0.001 + 0.853 \times 0.053^2 + 0.125(-0.012)^2 = 0.00341 \) so that \( \sigma_h(1) = 0.0584 \).

13. Compute the volatility forecast for time \( t = h + 2 \) made at the forecast origin \( t = h \).
Answer: \( \sigma_h^2(2) = 0.001 + (0.853 + 0.125)\sigma_h^2(1) = 0.00434 \) so that \( \sigma_h(2) = 0.0659 \).

14. Describe a consequence of overlooking the serial correlations in a linear regression model.
Answer: Not reliable estimates of the variances of the parameter estimates.

15. Describe two differences between the empirical density function of an asset return series and the normal density that has the same mean and standard deviation.
Answer: (a) The empirical distributions have heavy tails. (b) The empirical density tends to have a higher peak.

**Problem B.** (37 pts) Consider the daily log returns, in percentages, of the 3M stock from January 1999 to December 2008. The relevant R output is attached. Answer the following questions.

1. (2 points) Is there any serial correlation in the log return series? Why?
Answer: Yes, the Ljung-Box statistic gives \( Q(10) = 28.73 \) with p-value 0.0014.
2. (2 points) An MA model is used to handle the mean equation, which appears to be adequate. Write down the fitted MA model, including the residual variance.

Answer: \( r_t = a_t - 0.0218a_{t-1} - 0.0773a_{t-2} \), with \( \sigma^2_a = 2.661 \).

3. (6 points) A GARCH(1,1) model with Gaussian distribution is used for the volatility equation. It turns out the MA coefficients become insignificant so that the mean equation is refined. Write down the fitted model, including the mean equation.

Answer:

\[
\begin{align*}
r_t &= 0.046 + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim N(0,1) \\
\sigma^2_t &= 0.110 + 0.090a^2_{t-1} + 0.869\sigma^2_{t-1}.
\end{align*}
\]

4. (2 points) Except for the normality, is the fitted GARCH(1,1) model adequate? Why?

Answer: Yes, the fitted model is adequate as all \( p \)-values involved in the model checking are greater than 0.05.

5. (4 points) To check the risk premium of volatility, a GARCH(1,1)-M model is used. Is the risk premium statistically significant? Why?

Answer: No, the ARCH-in-mean parameter is 0.0294 with \( p \)-value 0.23.

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

Answer: \( \sigma^2_t = 0.00689 + 0.0373a^2_{t-1} + 0.9619\sigma^2_{t-2} \), where the Student \( t \) distribution has 4.70 degrees of freedom.

7. (3 points) Focus on an IGARCH(1,1) model, write down the fitted model.

Answer:

8. (2 points) Why are the volatility forecasts of the IGARCH(1,1) model not constant for all forecast horizons.

Answer: Because of the existence of the constant term 0.0585.

9. (3 points) A GJR model is also fitted. Write down the fitted volatility equation.

Answer: \( \sigma^2_t = 0.102 + (0.0533 + 0.0734N_{t-1})a^2_{t-1} + 0.873\sigma^2_{t-1} \), where \( N_{t-1} \) is the indicator variable for the negative value of \( a_{t-1} \).

10. (2 points) Based on the GJR model, is the leverage effect significant? Why?

Answer: Yes, the leverage parameter has a \( t \)-ratio 3.695 with \( p \)-value 0.0002.

11. (4 points) An EGARCH(1,1) model is also fitted. Write down the fitted volatility equation. You may ignore the insignificant parameter, if any.

Answer: \( \ln(\sigma^2_t) = 1.143 + \frac{1.497}{1.048}g(\epsilon_{t-1}) \), where

\[
g(\epsilon_t) = -0.04888\epsilon_t + 0.146(|\epsilon_t| - 0.8).
\]
12. (4 points) Based on the EGARCH model, calculate the ratio $\frac{\sigma_t^2(\epsilon_{t-1}=-2)}{\sigma_t^2(\epsilon_{t-1}=2)}$ to see the leverage effect.

Answer: Based on the model, $g(-2) = -0.0488(-2) + 0.146(2 - 0.8) = 0.2728$ $g(2) = -0.0488(2) + 0.146(2 - 0.8) = 0.0776$. Also, $\ln(\sigma_t^2) = 0.971 \ln(\sigma_{t-1}^2) + 0.0331 + g(\epsilon_{t-1})$. Consequently,

$$\frac{\sigma_t^2(\epsilon_{t-1}=-2)}{\sigma_t^2(\epsilon_{t-1}=2)} = \frac{e^{g(-2)}}{e^{g(2)}} = 1.216.$$ 

The effect is about 21.6% for a shock of magnitude two standard deviations.

**Problem C.** (13 pts) Consider the monthly unemployment rate for the State of Illinois from 1976 to March 2009. The data are obtained from the Federal Reserve Bank at St Louis and are seasonally adjusted. Let $x_t$ be the monthly unemployment rate.

1. (5 points) Write down the final fitted model for $x_t$, including the residual variance.

Answer: $(1 - 1.259B + 0.138B^2 + 0.133B^3)(x_t - 7.030) = a_t$, where $\sigma_a^2 = 0.0295$.

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

Answer: Yes, the residuals fail to reject the null hypothesis of no serial correlations. For instance, $Q(24) = 22.38$ with $p$-value 0.56 for the residuals.

3. (2 points) Based on the fitted model, are there business cycles in the Illinois economy? Why?

Answer: Yes, because the AR polynomial contains a pair of complex roots.

4. (2 points) Calculate the average length of the business cycles implied by the fitted model.

Answer: $k = 2\pi / \cos^{-1}(-0.105095/1.783618) = 3.855$. That is, average cycle length is about 4 months.

5. (2 points) Use the data and fitted model, obtain a 95% interval forecast for the May unemployment rate of Illinois.

Answer: $9.753 \pm 1.96(0.276)$.

**Problem D.** (20 pts) As mentioned in class, VIX of CBOE is widely used as a volatility index for the U.S. stock market. This problem is concerned with the relationship between the VIX index and the volatility series estimated from the daily log returns of the S&P 500 index. The sample period is from January 2, 2004 to April 16, 2009 for 1331 observations. Note that I used a Gaussian AR(2)-GARCH(1,1) model to estimate the volatility. In addition, the GARCH daily volatility is annualized (multiplied by $\sqrt{252}$) to match with VIX measure. Let $Y_t = \ln(VIX_t)$ and $X_t = \ln(V_t)$, where $V_t$ is the volatility obtained from the GARCH(1,1) model.
1. (2 points) Is the fitted GARCH(1,1) model for the daily S&P 500 index highly persistent? Why?
   Answer: Yes, because \( \alpha + \beta = 0.993 \), which is close to 1.

2. (3 points) Write down the linear regression model between \( Y_t \) and \( X_t \).
   Answer: \( Y_t = 0.528 + 0.859X_t + e_t \).

3. (3 points) Let \( y_t \) and \( x_t \) be the first differenced series of \( Y_t \) and \( X_t \), respectively. Is the correlation between \( y_t \) and \( x_t \) significant at the 5% level? Why?
   Answer: Yes, because the slope coefficient of regressing \( y_t \) on \( x_t \) is significant at the 5% level.

4. (6 points) Write down the final fitted model between \( y_t \) and \( x_t \).
   Answer: \( y_t = -0.0108 x_t + e_t, \ e_t = (1 - 0.166B - 0.094B^2)(1 - 0.089B^{10})a_t, \) where \( \sigma_a^2 = 0.00413 \).

5. (2 points) Is there any serial correlation in the residuals of the final fitted model? Why?
   Answer: No, the Ljung-Box statistic gives \( Q(20) = 14.96 \) with p-value 0.78.

6. (4 points) Does the VIX index depend on the GARCH volatility? Why?
   Answer: No, the final regression model fails to have a significant coefficient for the differedenced GARCH(1,1) volatility.