

Solutions to Midterm

**Problem A:** (30 pts) Answer briefly the following questions.

1. Describe a situation under which the  $R^2$  defined as

$$R^2 = \frac{(\text{Sum of squares of total}) - (\text{Sum of squares of residuals})}{\text{Sum of squares of total}},$$

is not informative in evaluating a fitted time series model.

Answer: The time series is unit-root nonstationary.

2. Consider a linear regression model with time-series errors. Why is the Durbin-Watson statistics not sufficient in model checking?

Answer: DW statistic only checks the lag-1 serial correlation.

3. **For questions 3 to 5**, consider the AR(1)-IGARCH(1,1) model

$$r_t = 0.02 + 0.2r_{t-1} + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim N(0, 1).$$

$$\sigma_t^2 = 0.1a_{t-1}^2 + 0.9\sigma_{t-1}^2.$$

What is the expected value of  $r_t$ ?, i.e.,  $E(r_t) = ?$

Answer:  $E(r_t) = \frac{0.02}{1-0.2} = 0.025$ .

4. Suppose that  $r_{h-1} = 0.04$ , what are the 1-step ahead forecast and its forecast error of  $r_t$  at the forecast origin  $h - 1$ ?, i.e.  $r_{h-1}(1) = ?$  and  $e_{h-1}(1) = ?$

Answer:  $r_{h-1}(1) = 0.02 + 0.2r_{h-1} = 0.02 + 0.2 \times 0.04 = 0.028$ .  $e_{h-1}(1) = a_h$ .

5. In addition to the information of the prior question, suppose we also observe that  $r_h = -0.012$  and  $\sigma_h^2 = 0.25$ . What are the 1-step and 2-step ahead volatility forecasts of the model at time origin  $h$ ? That is, what are  $\sigma_h^2(1)$  and  $\sigma_h^2(2)$ ?

Answer:  $a_h = r_h - r_{h-1}(1) = -0.012 - 0.028 = -0.04$ . Thus,  $\sigma_h^2(1) = 0.1a_h^2 + 0.9\sigma_h^2 = 0.1 \times (-0.04)^2 + 0.9 \times 0.25 = 0.225$ . For the particular IGARCH model considered,  $\sigma_h^2(2) = \sigma_h^2(1) = 0.225$ .

6. Give two advantages of EGARCH models over the GARCH models.

Answer: (1) Allows for leverage effect, i.e. asymmetric responses in volatility for prior positive and negative returns. (2) Uses log volatility to relax the positiveness constraint.

7. **For problems 7 to 9**, consider the daily exchange rate between U.S. dollar and U.K. pound from January 2001 to April 26, 2007. Descriptive statistics of the daily log returns are given in the attached output. Is the mean of the log return different from zero? Why?

Answer: No. The 95% confidence interval for the mean is  $(-0.0001, 0.0004)$ , which contains zero.

8. Is the distribution of the log return symmetric with respect to its mean? Why?

Answer: No, the t-ratio for skewness is  $t = -0.142/\sqrt{6/1587} = -2.31$ . The p-value is 0.021. Thus, the null of symmetry with respect to its mean is rejected at the 5% level.

9. Does the distribution of the log return have heavy tails? Why?

Answer: Yes. The t-ratio for excess kurtosis is  $t = \frac{0.597}{\sqrt{24/1587}} = 4.85$  with p-value close to zero. Reject the null hypothesis of normal tails.

10. Suppose that the monthly time series  $r_t$  follows the model

$$r_t = (1 - \theta_2 B^2)(1 - \theta_{12} B^{12})a_t, \quad a_t \sim N(0, \sigma_a^2),$$

where  $\theta_2$  and  $\theta_{12}$  are non-zero real numbers satisfying  $|\theta_2| < 1$  and  $|\theta_{12}| < 1$ , and  $\sigma_a^2 > 0$ . List all non-zero autocorrelations of  $r_t$ .

Answer:  $\rho_2, \rho_{10}, \rho_{12}$  and  $\rho_{14}$ . (Of course,  $\rho_0 = 1$  is non-zero.)

11. Give two reasons that observed daily returns of an asset are serially correlated even though the true underlying returns are serially uncorrelated.

Answer: Any two of (a) nonsynchronous trading, (b) bid-ask bounce, (c) risk premium.

12. To test for ARCH effect, one often employs the Ljung-Box statistics  $Q(m)$  of the squared residuals of the mean equation. Write down the null and alternative hypotheses for  $Q(10)$  statistic in ARCH-effect testing.

Answer:  $H_o : \rho_1 = \dots = \rho_{10} = 0$  versus  $H_a : \text{Some } \rho_i \neq 0 \text{ for some } i \text{ with } 1 \leq i \leq 10$ , where  $\rho_i$  is the lag- $i$  ACF of the squared residuals.

13. Assume that time series  $x_t$  and  $y_t$  follow the following models,

$$\begin{aligned} x_t &= 0.5x_{t-1} + a_t, \\ y_t &= 1.3y_{t-1} - 0.4y_{t-2} + a_t, \end{aligned}$$

where  $\{a_t\}$  are *iid*  $N(0, \sigma_a^2)$  with  $\sigma_a^2 > 0$ . Both series are mean reverting. What is the half-life for  $x_t$ ? What is the half-life of  $y_t$ ?

Answer: For  $x_t$ , it is  $k = \frac{\ln(0.5)}{\ln(0.5)} = 1$ . For  $y_t$ , since the characteristic equation is  $(1 - 1.3x + 0.4x^2) = (1 - 0.8x)(1 - 0.5x)$ , the half-life is approximately  $k = \frac{\ln(0.5)}{\ln(0.8)} = 3.11$ . (A more careful approximation is  $k = \frac{\ln(0.5)}{\ln(1.3-0.4)}$ .)

14. Suppose that your average daily balance of a credit card is \$1000. Suppose also that the card charges an interest rate of 22.5% per annum (daily compounding). How much is your financial charge in a 30-day billing cycle?

Answer:  $1000(1 + \frac{.225}{365})^{30} - 1000 = 18.66$ . You can also use  $1000 \exp(\frac{0.225}{365} \times 30) - 1000 = 18.93$ . (You may use 360 days for a year.)

15. Suppose that the monthly log returns of an asset are normally distributed with mean 0.08 and standard deviation 0.12. What is the mean of the monthly simple return of the asset?

Answer:  $E(R_t) = \exp(0.08 + \frac{0.12^2}{2}) - 1 = 0.091$ .

**Problem B.** (20 pts) Consider Moody's seasoned AAA and BAA corporate bond yields from January 5, 1962 to April 20, 2007. The data are averages of daily yields and obtained from the Federal Reserve Bank of St. Louis. Denote the bond yields by AAA and BAA, respectively. To find the relationship between the two bond yields, we conduct certain analysis. The output is attached. Answer the following questions.

1. Write down the fitted linear regression with BAA and AAA representing the dependent and independent variable, respectively. What is the  $R^2$  of the linear regression? Is the fitted model adequate? Why?

Answer:  $BAA = -0.03 + 1.129AAA + \epsilon$  with  $\sigma_\epsilon = 0.279$ . The  $R^2$  is 98.97%. The model is inadequate because the residuals have strong serial correlations.

2. Let  $Y_t = BAA_t - BAA_{t-1}$  and  $X_t = AAA_t - AAA_{t-1}$  be the differenced series. Consider the linear regression  $Y_t = \beta_0 + \beta_1 X_t + \epsilon_t$ . What is the fitted model? What is the residual standard deviation of the model?

Answer:  $Y_t = 0.0002 + 0.719X_t + \epsilon_t$  with  $\sigma_\epsilon = 0.044$ .

3. The residuals of the prior linear regression show certain serial correlations. A linear regression model with time series errors is employed. Write down the fitted model. Based on the available output, is this model adequate? Why?

Answer:  $Y_t = 0.0002 + 0.692X_t + \epsilon_t$ , where  $\epsilon_t = a_t + 0.248a_{t-1} + 0.097a_{t-2}$  with  $\sigma_a^2 = 0.0018$ . Yes, the model is adequate. The Q(10) of the residual ACF is 9.48 with p-value 0.487, indicating no residual serial correlations.

4. Consider the above linear regression model with time-series error. One way to confirm that the MA(2) model is needed is to test the lag-2 MA coefficient. Write down the null and alternative hypotheses for such a test. What is the test statistic? Drawn your conclusion.

Answer:  $H_o : \theta_2 = 0$  versus  $H_a : \theta_2 \neq 0$ . The t-ratio is  $\frac{0.0974}{0.0199} = 4.89$ , which is greater than 1.96. Thus,  $H_o$  is rejected, confirming that  $\theta_2$  is not zero.

5. Construct a 95% confidence interval for the coefficient  $\beta_1$  (the slope parameter of the linear regression model with time-series errors). Is the estimate 0.719 (see Question 2) in the 95% confidence interval? Discuss the implication of the result.

Answer:  $0.692 \pm 1.96 \times 0.010$ , i.e.,  $[0.672, 0.712]$ . The estimate 0.719 is not in the 95% confidence interval, indicating that overlooking the residual serial correlations may lead to erroneous conclusion about the strength of dependence between the two variables.

**Problem C.** (30 pts) Consider the daily closing values of the VIX index (which is an implied volatility for the S&P 500 index) of CBOE from January 2, 2004 to April 5, 2007. The index appears to have a unit root so that we analyze its log return series. The relevant compute output is attached. Answer the following questions.

1. (4 points) Write down the fitted mean equation for the log return series, including the residual variance. Is the model adequate in handling the serial correlations? Why?

Answer: The model is  $r_t = -0.0004 + a_t - 0.103a_{t-1} - 0.117a_{t-2}$  with  $\sigma_a^2 = 0.0033$ . The model is adequate for the mean equation because Q(5) and Q(10) of the residuals cannot reject the null hypothesis of no serial correlations.

2. Is there any ARCH effect in the log return series? Why?

Answer: Yes, the Q(10) statistic of the squared residuals (of the mean equation) is 58.32 with p-value close to zero.

3. A GARCH(1,1) model is used in the volatility equation. Write down the fitted model, including the degrees of freedom of the Student- $t$  innovations.

Answer: The model consists of the mean equation  $r_t = -0.0029 + a_t - 0.102a_{t-1} - 0.111a_{t-2}$ ,  $a_t = \sigma_t \epsilon_t$ , where  $\epsilon$  follows a normalized Student- $t$  distribution with 4.96 degrees of freedom, and the volatility equation  $\sigma_t^2 = 2.094 \times 10^{-4} + 0.087a_{t-1}^2 + 0.845\sigma_{t-1}^2$ .

4. Based on the output, what are the estimated standard errors of ARCH ( $\alpha_1$ ) and GARCH ( $\beta_1$ ) coefficients?

Answer: The standard error of  $\alpha_1$  and  $\beta_1$  are  $\frac{0.0869}{3.156} = 0.028$  and  $\frac{0.845}{19.07} = 0.044$ , respectively.

5. (8 points) A GJR (or TGARCH) model is also fitted to the log return series. Write down the fitted model.

Answer: The model consists of the mean equation

$$r_t = -0.0024 + a_t - 0.097a_{t-1} - 0.099a_{t-2}, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim t_{5,10},$$

and the volatility equation

$$\sigma_t^2 = 2.03 \times 10^{-4} + (0.12 - 0.139N_{t-1})a_{t-1}^2 + 0.867\sigma_{t-1}^2,$$

where  $N_{t-1} = 1$  if  $a_{t-1} < 0$ , and  $= 0$  otherwise.

6. Is the fitted GJR (or TGARCH) model adequate? Why?

Answer: Yes, the model is adequate based on the information provided. The Q statistics of the standardized residuals and the squared standardized residuals fail to reject the hypothesis of no serial correlations and no ARCH effects.

7. (4 points) Between the GARCH(1,1) and GJR(1,1) models, which one is preferred? Why?

Answer: The GJR(1,1) model because its has lower AIC value.

8. Is the leverage effect of the GJR model significant? Why? Why is the leverage parameter negative?

Answer: Yes, the t-ratio of the leverage effect is  $-2.927$  with p-value  $0.0035$ . We reject the null hypothesis of no leverage effect at the 5% level. The leverage effect is negative because a drop in volatility means the market is less volatile so that the volatility of volatility is less affected. An unexpected increase in volatility would create uneasiness in the market so that it has a greater impact on the volatility of volatility.

9. (5 points) To better understand the leverage effect, use the fitted GJR or TGARCH model to calculate the ratio  $\frac{\sigma_t^2(\epsilon_{t-1}=-2)}{\sigma_t^2(\epsilon_{t-1}=2)}$ , where  $\{\epsilon_t\}$  denotes the standardized innovation. For simplicity, you may ignore the constant term of the volatility equation.

Answer: Ignoring the constant term, we have  $\frac{\sigma_t^2(\epsilon_{t-1}=-2)}{\sigma_t^2(\epsilon_{t-1}=2)} = \frac{(.1201-.1386)(-2)^2+.8671}{.1201*2^2+.8671} = 0.589$ .

10. (4 points) Based on the fitted GJR or TGARCH model, what are the 1-step and 5-step ahead forecasts of the log return and its volatility at the forecast origin  $T = 820$ , the last data point?

Answer: Forecasts of the log return are  $r_{820}(1) = 0.00053$  and  $r_{820}(5) = -0.0024$ . The volatility forecasts are  $\sigma_{820}(1) = \sqrt{0.003373}$  and  $\sigma_{820}(5) = \sqrt{0.002393}$ , i.e.,  $0.0581$  and  $0.0489$ , respectively.

**Problem D.** (20 pts) Consider the quarterly earnings per share of the FedEx stock from the fourth quarter of 1991 to the last quarter of 2006. The data were obtained from First Call. To take the log transformation, we add one to all data points. Compute output is attached. Let  $x_t = \ln(y_t + 1)$  be the transformed earnings, where  $y_t$  is the actual earnings per share.

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

Answer:  $(1 - B)(1 - B^4)x_t = (1 - 0.722B)(1 - 0.382B^4)a_t, \quad \sigma_a^2 = 0.0072$ .

2. (4 points) Is there any significant serial correlation in the residuals of the fitted model? Why?

Answer: No, the Q(12) statistic of the residuals is  $9.95$  with p-value  $0.62$ . If we adjusted the degrees of freedom of the Chi-square distribution, the p-value of  $0.445$  remains high.

3. (4 points) Let  $T = 62$  be the forecast origin. Based on the fitted model, and, for simplicity, use the relationship  $y_t = \exp(x_t) - 1$ , what are the 1-step and 2-step ahead forecasts of earnings per share for the FedEx stock?

Answer: The forecasts are  $x_{62}(1) = 1.64$  and  $x_{62}(2) = 2.20$ .

4. (3 points) Obtain a 95% interval forecast for  $x_{63}$  at the forecast origin  $T = 62$ .

Answer:  $0.972 \pm 1.96 \times 0.085$ , i.e.,  $[0.805, 1.139]$ .

5. Test the null hypothesis  $H_o : \theta_4 = 0$  vs  $H_a : \theta_4 \neq 0$ . What is the test statistic? Draw your conclusion.

Answer: The test statistic is  $t = \frac{-0.382}{0.116} = -3.293$  with p-value 0.0016. Thus, we reject the null hypothesis at the 5% level. The seasonal MA parameter is significantly different from zero.