

Solutions to Midterm

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

1. Suppose that your bank pays interest rate 6% per annum on a quarterly basis, and your goal is to have \$30,000 at the end of 5 years. How much do you have to deposit now in order to achieve the goal?

Answer: $C = \$30000 \left(1 + \frac{0.06}{4}\right)^{-20} = \$22,274.11$.

2. **For questions 2 and 3**, consider the AR(2) model

$$(1 - B + 0.3B^2)r_t = 0.3 + a_t,$$

where $\{a_t\}$ is a sequence of independent standard normal random variables. Does the model imply existence of business cycles? If yes, what is the average period of the cycles?

Answer: Yes, because the roots of polynomial $1 - B + 0.3B^2$ are $1.666667 \pm 0.745356i$. The average period is $2\pi / \cos^{-1}(1.666667/1.825742) \approx 14.94$.

Alternatively, write the equation as $1 - B - (-.3)B^2 = 0$. We have $\phi_1^2 + 4\phi_2 = 1 + 4(-.3) = -0.2 < 0$ so that there are business cycles. The average period is $2\pi / \cos^{-1}(1/(2\sqrt{3})) = 14.94$.

3. Suppose that the last two data points are $r_{100} = 1.2$ and $r_{99} = -0.6$. What is the 1-step ahead forecast $r_{100}(1)$? What is the 100-step ahead forecast $r_{100}(100)$?

Answer: $r_{100}(1) = r_{100} - .3r_{99} + 0.3 = 1.68$. Using mean-reverting, $r_{100}(100) = E(r_t) = 0.3/(1 - 1 + 0.3) = 1$.

4. **For questions 4 to 8**, consider the monthly log returns, including dividends, of Johnson and Johnson stock from January 1961 to December 2005 for 540 observations. Summary statistics of the log returns from R and S-Plus are attached. Is the mean of log returns significantly different from zero? Why?

Answer: Yes. From R output, the 95% C.I. of the mean does not include zero. From S-Plus output, $t = 0.01219/(0.06393/\sqrt{540}) = 4.43 > 1.96$.

5. Suppose one invested \$1 dollar on Johnson and Johnson stock at the very beginning of 1961 and held the investment until the end of 2005. What was the value of the investment at the market closing on December 30, 2005 (the last trading day)?

Answer: From R, $\exp(6.581) = \$721.26$. From S-Plus, $\exp(0.1219 \times 540) = \722.42 .

6. Test the null hypothesis that the log returns is normally distributed. Draw your conclusion. You may use the Jarque-Bera statistic.

Answer: $JB = \frac{(-.1193)^2}{6/540} + \frac{0.084^2}{24/540} = 1.44$. Using χ_2^2 distribution, the p-value is 0.49. Thus, cannot reject the normality assumption for the log returns.

7. Given that the observed simple return of Johnson and Johnson stock on December 30, 2005 was -2.672% , what is the corresponding log return?

Answer: $r_{540} = \ln(-0.02672 + 1) = -0.0271 = -2.71\%$.

8. What is the null hypothesis of the $Q(10)$ statistic of the log return series? What is the conclusion based on the $Q(10)$ statistic?

Answer: (1) $H_o : \rho_1 = \rho_2 = \dots = \rho_{10} = 0$. (2) $Q(10) = 10.26$ with p-value 0.42. Cannot reject H_o of no serial correlations.

9. Consider a GARCH(1,1) model

$$\sigma_t^2 = 0.15 + 0.07a_{t-1}^2 + 0.78\sigma_{t-1}^2$$

for $a_t = \sigma_t \epsilon_t$, where $\{\epsilon_t\}$ is a sequence of independent standard normal random variables. What is the unconditional variance of a_t ?

Answer: $\text{Var}(a_t) = 0.15 / (1 - 0.07 - 0.78) = 1.0$.

10. Give two possible reasons that introduce serial correlations in some daily U.S. stock or index returns.

Answer: Any two of (1) nonsynchronous trading, (b) bid-ask bounce, and (c) risk premium associated with volatility risk.

11. Consider an IGARCH model for $a_t = \sigma_t \epsilon_t$,

$$\sigma_t^2 = 0.145a_{t-1}^2 + .855\sigma_{t-1}^2.$$

Suppose that $a_{100} = 0.02$ and $\sigma_{100} = 0.2$. What is the 1-step ahead volatility forecast $\sigma_{100}(1)$? What is the 10-step ahead volatility forecast $\sigma_{100}(10)$? [Recall that volatility is the conditional standard deviation.]

Answer: $\sigma_{100}(1) = \sqrt{.145(0.02)^2 + .855(.2)^2} = 0.185$, and $\sigma_{100}(10) = \sigma_{100}(1) = 0.185$.

12. Consider the simple moving average model of order 2, i.e., MA(2),

$$r_t = 0.015 + a_t - 0.3a_{t-2},$$

where $\{a_t\}$ is a sequence of independent and identically distributed normal random variables with mean zero and variance 2. What is the mean of r_t [i.e., $E(r_t)$]? What is the lag-1 ACF of r_t ?

Answer: (1) $E(r_t) = 0.015$, (2) $\rho_1 = 0$.

13. Standardized residuals, denoted by $\hat{\epsilon}_t$, are often used to check a fitted GARCH model. What are the purposes of checking the Ljung-Box statistics $Q(10)$ of $\hat{\epsilon}_t$ and $\hat{\epsilon}_t^2$?

Answer: $Q(10)$ of $\hat{\epsilon}_t$ is to check the mean equation and $Q(10)$ of $\hat{\epsilon}_t^2$ is for checking the volatility equation.

14. Describe two advantages of EGARCH models over GARCH models.

Answer: (1) Allows for asymmetric responses to prior positive and negative returns. (2) Uses log transformation to relax the positiveness constraint.

15. Let ρ_k be the lag- k ACF of a monthly time series r_t . Describe a method to test the null hypothesis $H_0 : \rho_{12} = 0$ versus the alternative hypothesis $\rho_{12} \neq 0$. Your method should include a decision rule for making inference.

Answer: Use t-ratio $t = \sqrt{T}\hat{\rho}_{12}$, where T is the sample size. For large T , the t-ratio follows a standard normal random variable. Decision rule: reject the null hypothesis of $\rho_{12} = 0$ if $|t| > 1.96$ at the 5% level.

Problem B. (20 pts) Power companies must supply sufficient natural gas to heat all of their customers' homes and businesses, particularly during the coldest days in winter. To forecast the sendouts of natural gas, data on some factors were collected for a major northern city of the U.S. in one winter. The dependent variable Y is the sendout and the explanatory variables are

1. X_1 : degree of heating days (= 65 degree - daily average temperature)
2. X_2 : lag-1 value of X_1 , i.e., $X_{2,i} = X_{1,i-1}$.
3. X_3 : wind speed, a 24 hour average,
4. X_4 : indicator variable for weekends. [$X_4 = 1$ for weekends and = 0, otherwise.]

There are 63 observations, representing 63 consecutive days. R and S-Plus outputs are attached. Use the output to answer the following questions.

1. (4 points) Write down the fitted multiple linear regression model,

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + e.$$

Answer: The fitted model is

$$Y = 1.858 + 5.874X_1 + 1.405X_2 + 1.315X_3 - 15.857X_4 + e.$$

All estimates, but the constant term, are significant at the 5% level.

2. (4 points) What are the residual variance and R^2 of the fitted multiple linear regression?

Answer: The residual variance is $18.32^2 = 335.62$ and the R^2 is 95.21%.

3. (4 points) Is the fitted multiple linear regression model adequate? Why?

Answer: No, the model is inadequate because the residuals have strong serial correlations. For instance, $Q(10) = 48.17$ with zero p-value for the residuals.

4. (6 points) A seasonal time series model is specified for the residuals with period 7, indicating the weekly cycle of the demand of natural gas. Write down the fitted model.

Answer: From R, the fitted model is

$$Y_i = 7.721 + 5.707X_{1i} + 1.409X_{2i} + 1.201X_{3i} - 10.059X_{4i} + e_i$$

$$(1 - 0.541B)(1 - 0.375B^7)e_i = a_i$$

where the variance of a_i is 188.2. The residuals of the fitted model show $Q(14) = 8.14$ with p-value 0.88. Thus, the model fits the data well.

From S-Plus, the fitted model is

$$Y_i = 8.265 + 5.810X_{1i} + 1.314X_{2i} + 1.309X_{3i} - 11.454X_{4i} + e_i$$

$$(1 - 0.552B)(1 - 0.368B^7)e_i = a_i,$$

where the variance of a_i is 200.10.

5. (2 points) The fitted regression model with time series errors is adequate. Why is the coefficient of Weekend indicator (X_4) negative? It suffices to provide a simple explanation.

Answer: Most businesses were closed over the weekends so that the demand of natural gas was less over the weekends.

Problem C. (30 pts) Consider the daily log returns of Qualcomm stock from January 1992 to December 2005 with 3530 observations. R and S-Plus outputs are attached. Because ACFs of the returns are small, the mean equation consists of a constant term only. Answer the following questions. **Note that a model consists of both mean and volatility equations.**

- (4 points) A GARCH(1,1) model with Student t distribution is fitted to the return series. Write down the fitted model.

Answer: From R, the fitted model is

$$r_t = 0.0003 + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim t_6$$

$$\sigma_t^2 = 0.018a_{t-1}^2 + 0.982\sigma_{t-1}^2.$$

From S-Plus, the fitted model is

$$r_t = 0.0006 + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim t_{6.05},$$

$$\sigma_t^2 = 0.036a_{t-1}^2 + 0.964\sigma_{t-1}^2.$$

- (8 points) A GJR (or TGARCH) model is also fitted to the return series. Write down the fitted model. Is the model adequate? Why?

Answer: From R, the fitted model is

$$r_t = 0.0003 + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim t_{6.29}$$

$$\sigma_t^2 = 0.048 + (0.031 + 0.045N_{t-1})a_{t-1}^2 + 0.947\sigma_{t-1}^2,$$

where $N_{t-1} = 0$ if $a_{t-1} \geq 0$ and $= 1$ if $a_{t-1} < 0$. Based on the checking statistics $Q(m)$ provided, the fitted model is adequate. Basically, all $Q(m)$ statistics for the standardized residuals and their squared series cannot reject the adequacy of the model.

From S-Plus, the fitted model is

$$r_t = 0.0003 + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim t_{6.31},$$

$$\sigma_t^2 = 0.8 \times 10^{-5} + (0.039 + 0.061N_{t-1})a_{t-1}^2 + 0.932\sigma_{t-1}^2,$$

where N_{t-1} is defined as before. The model fits the data well.

- (4 points) Based on the fitted GJR or TGARCH model, test the null hypothesis that the leverage effect is zero, i.e., $H_o : \gamma = 0$ versus $H_a : \gamma \neq 0$. Draw your conclusion.

Answer: From R, the t-ratio of γ is 3.87 with p-value 0.0001. From S-Plus, the t-ratio is 4.29 with p-value 0.00002. Thus, the leverage effect is significant at the 5% level.

4. (5 points) To better understand the leverage effect, use the fitted GJR or TGARCH model to calculate the ratio $\frac{\sigma_t^2(\epsilon_{t-1}=-3)}{\sigma_t^2(\epsilon_{t-1}=3)}$, where $\{\epsilon_t\}$ denotes a sequence of innovations that are independent and identically distributed. For simplicity, you may ignore the constant term of the volatility equation.

Answer: From R, we obtain

$$\frac{\sigma_t^2(-3)}{\sigma_t^2(3)} = \frac{\sigma_{t-1}^2[(0.947 + 0.076 \times 3^2)]}{\sigma_{t-1}^2[0.947 + 0.031 \times 3^2]} = \frac{1.631}{1.226} = 1.33.$$

From S-Plus,

$$\frac{\sigma_t^2(-3)}{\sigma_t^2(3)} = \frac{\sigma_{t-1}^2[(0.932 + 0.10 \times 3^2)]}{\sigma_{t-1}^2[0.932 + 0.039 \times 3^2]} = \frac{1.832}{1.283} = 1.43.$$

5. (4 points) Based on the fitted GJR or TGARCH model, what are the 1-step ahead forecasts of the log return and its volatility at the forecast origin $T = 3530$, the last data point?

Answer: From R, log return: 0.00032; volatility is $\sqrt{0.0003102} = 0.0176$.

From S-Plus, log-return 0.00031, volatility is 0.0179.

6. (4 points) An EGARCH model is also fitted to the log return series. The θ_1 in R output and LEV parameter in S-Plus denote the leverage effect. Is the leverage effect significant? Why?

Answer: From R, the t-ratio of θ_1 is

$$t = \frac{-0.093}{0.024} = -3.93.$$

From S-Plus, the t-ratio is $t = \frac{-0.288}{0.064} = -4.51$. Thus, the leverage effect is significant.

7. (3 points) Based on the AIC criterion, compare the fitted GJR (or TGARCH in S-Plus) model against EGARCH model. Which model is preferred?

Answer: From R output, GJR model is preferred. From the S-Plus, EAGRCH model is better.

Problem D. (20 pts) Consider the quarterly earnings per share of the Carterpillar (CAT) Company from December 1991 to September of 2005 for 56 data points. The data were obtained from First Call. R and S-Plus outputs are attached. Let x_t be the quarterly earnings per share and $w_t = (1 - B)(1 - B^4)x_t$.

1. (4 points) Write down the sample ACFs of w_t at lag 1, 3, 4, and 5.

Answer: $\rho_1 = -0.27$, $\rho_3 = 0.20$, $\rho_4 = -0.379$ and $\rho_5 = .107$.

2. (5 points) Write down the fitted model for x_t , including the variance of the residuals.

Answer: The fitted model is

$$(1 - B)(1 - B^4)x_t = (1 - 0.216B)(1 - 0.474B^4)a_t,$$

where the variance of a_t is 0.0086.

3. (4 points) Is there any significant serial correlation in the residuals of the fitted model? Why? [Hint: use the checking statistics provided in the output.]

Answer: No, the Ljung-Box statistics show $Q(10) = 4.15$ with p-value 0.94.

4. (4 points) Let $T = 56$ be the forecast origin. Based on the fitted model, what are the 1-step and 2-step ahead forecasts of earnings per share for CAT stock?

Answer: 1-step and 2-step ahead forecasts of quarterly earnings are 1.048 and 1.043, respectively.

5. (3 points) Obtain a 95% interval forecast for x_{57} at the forecast origin $T = 56$.

Answer: The 95% interval forecast is $1.048 \pm 1.96 \times 0.093$. That is, $[0.866, 1.230]$.