Booth School of Business, University of Chicago
Business 41202, Spring Quarter 2015, Mr. Ruey S. Tsay

Solutions to Midterm

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

1. Give two situations under which returns of an assert follow an MA(1) model.

   Answer: (1) Bid-ask bounce in high-frequency trading and (2) data smoothing (or manipulation).

2. Describe two ways by which a GARCH(1,1) model can introduce heavy tails.

   Answer: (1) Use heavy-tailed distribution for ϵ_t and (2) the GARCH dynamic introduced by α_1.

3. **(Questions 3 to 6)**: Suppose that the asset return r_t follows the model

$$\begin{aligned} r_t &= 0.005 + a_t \\ a_t &= \sigma_t \epsilon_t, \quad \epsilon_t \sim \text{iid } t_6^* \\ \sigma_t^2 &= 0.09 + 0.10 a_{t-1}^2 + 0.855 \sigma_{t-1}^2. \end{aligned}$$

   Compute the mean and variance of r_t, i.e., $E(r_t)$ and $\text{Var}(r_t)$.

   Answer: (1) $E(r_t) = 0.005$ and (2) $\text{Var}(r_t) = \text{Var}(a_t) = \frac{0.09}{1-0.1-0.855} = 2$.

4. Express the volatility model in an ARMA(1,1) formation using $\eta_t = a_t^2 - \sigma_t^2$.

   Answer: $a_t^2 = 0.09 + 0.955 a_{t-1}^2 + \eta_t - 0.855 \eta_{t-1}$.

5. Describe two nice characteristics of the volatility model for r_t.

   Answer: (1) creates heavy tails ($\alpha_1 \neq 0$) and (2) describe volatility clustering.

6. Suppose further that $a_{100} = -0.02$ and $\sigma_{100}^2 = 0.16$. Compute the 1-step ahead prediction (both mean and volatility) of the return at the forecast origin $t = 100$.

   Answer: $r_{100}(1) = 0.005$. $\sigma_{101}^2 = 0.09 + 0.1(-0.02)^2 + 0.855(0.16) = 0.227$. Therefore, $\sigma_{100}(1) = 0.476$.

7. Give two nice features of the exponential GARCH model that the standard GARCH model does not have. Answer: (1) Use log volatility to relax parameter constraints and (2) model the leverage effect.

8. Give two empirical characteristics of daily asset returns.

   Answer: (1) Heavy tails or high excess kurtosis and (2) volatility clustering.

9. **(Questions 9 - 11)**: Let p_t be the log price of an asset at time t. Assume that the log price follows the model

$$p_t = 0.001 + p_{t-1} + a_t, \quad a_t \sim \text{iid } N(0, 0.16),$$

   where $N(\mu, \sigma^2)$ denotes normal distribution with mean μ and variance σ^2. Assume further that $p_{200} = 4.551$. Compute the 95% interval forecast for p_{201} at the forecast origin $t = 200$.

   Answer: Point forecast $p_{201}(1) = 0.001 + 4.551 = 4.552$. Variance of forecast error is 0.16. Therefore, 95% interval forecast is $4.552 \pm 1.96 \times 0.4$, i.e. (3.768, 5.336).

10. Compute the 2-step ahead point forecast and its standard error for p_{202} at the forecast origin $t = 200$.

    Answer: $p_{200}(2) = 0.001 + p_{200}(1) = 4.553$. Variance of forecast errors is $2 \times 0.16 = 0.32$ so that the standard error of forecast is 0.566.

11. What is the 100-step ahead forecast for p_{300} at the forecast origin $t =$ 200?

    Answer: $p_{200}(100) = 100 * \times 0.001 + p_{100} = 4.651$.

12. Describe two methods for comparing two different models for a time series z_t.

    Answer: (1) In sample comparison: Use information criterion and (2) out-of-sample: back-testing with root mean squares errors or mean absolute errors.

13. **(Questions 13-15)**: Suppose that the quarterly growth rates r_t of an economy follows the model

$$r_t = 0.006 + 0.168 r_{t-1} + 0.338 r_{t-2} - 0.189 r_{t-3} + a_t, \quad a_t \sim \text{iid } N(0, .0016).$$

    What is the expected growth rate of r_t?

    Answer: $E(r_t) = 0.006/(1 - .168 - .338 + .189) = 0.00878$.

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

    Answer: Yes, the equation $1 - .168x - .338x^2 + .189x^3 = 0$ contains complex roots.

15. What is the average length of business cycles of the economy, if any?

    Answer: $k = \frac{2\pi}{\cos^{-1}(1.608/1.925)} = 10.79$ quarters.

Problem B. (45 points) Consider the daily log returns r_t, in percentages, of the NASDAQ index for a certain period of time with 1841 observations. Answer the following questions, using the attached R output.

1. (2 points) Let μ be the expected value of r_t. Test $H_0 : \mu = 0$ versus $H_a : \mu \neq 0$. Obtain the test statistic and draw your conclusion.

   Answer: $t = 0.0361/0.035 = 1.031$, which is less than 1.96 so that $\mu = 0$ cannot be rejected.

2. (2 points) Is the distribution of r_t skew? Why?

   Answer: $t = \frac{-0.245}{\sqrt{6/1841}} = -4.292$, which is less than -1.96 so that the distribution is skewed.

3. (2 points) Does the distribution of r_t have heavy tails? Why?

   Answer: Yes, it has heavy tails because $t = 6.832/sqrt(24/1841) = 59.837$, a large value.

4. (2 points) Let ρ_1 be the lag-1 ACF of r_t. Test $H_0 : \rho_1 = 0$ versus $H_a : \rho_1 \neq 0$. The sample lag-1 ACF is -0.086. Obtain the test statistic and draw your conclusion.

   Answer: $t = -0.086/sqrt(1/1841) = -3.69$, which is less than -1.96. Yes, the lag-1 ACF is not zero.

5. (2 points) An MA(1) model is fitted. Write down the fitted model, including σ^2 of the residuals.

   Answer: $r_t = a_t - 0.948a_{t-1}$ with $\sigma^2 = 2.241$.

6. (2 points) Is there any ARCH effect in the residuals of the fitted MA(1) model? Why?

   Answer: Yes, the Ljung-Box statistics of the squared residuals show Q(10) = 1507.77 with p-value close to zero.

7. (2 points) An ARMA(0,1)+GARCH(1,1) model with Gaussian innovations was entertained for r_t. See **m2**. Write down the fitted model.

   Answer: $r_t = 0.0915 - 0.0366a_{t-1}$, $a_t = \sigma_t\epsilon_t$ with ϵ_t being iid N(0,1), and $\sigma_t^2 = 0.0282 + 0.104a_{t-1}^2 + 0.88\sigma_{t-1}^2$.

8. (2 points) Is the fitted ARMA(0,1)+GARCH(1,1) model adequate? Why?

   Answer: No, the MA(1) coefficient is not significant, and the normality assumption is highly rejected.

9. (3 points) A GARCH(1,1) model with standardized Student-t innovations was fitted. See **m3**. Write down the fitted model.

   Answer: $r_t = 0.119 + a_t$, $a_t = \sigma_t\epsilon_t$ with ϵ_t being iid $t^*_{7.03}$, and $\sigma_t^2 = 0.0253 + 0.106a_{t-1}^2 + 0.882\sigma_{t-1}^2$.

10. (3 points) To further improve the model, a GARCH model with skew standardized Student-t innovations was considered. See **m4**. Write down the fitted model.

    Answer: $r_t = 0.0905 + a_t$, $a_t = \sigma_t \epsilon_t$ with ϵ_t being iid $t^*_{8.12}(0.873)$, where 0.873 is the skew parameter. The volatility equation is $\sigma_t^2 = 0.0233 + 0.104 a_{t-1}^2 + 0.884 \sigma_{t-1}^2$.

11. (2 points) Based on the fitted model **m4**. Does the return r_t follow a skew distribution? Why? Perform proper test and draw the conclusion.

    Answer: $t = (0.873 - 1)/.0275 = -4.618$, which is less than -1.96. Therefore, the distribution is skewed.

12. (2 points) Based on the model **m4**. Obtain 1-step and 2-step 95% interval forecasts for r_t at the forecast origin $t = 1841$.

    Answer: 1-step: $0.0905 \pm 1.96(0.78)$; 2-step: $0.0905 \pm 1.96(0.791)$. That is, $(-1.438, 1.619)$ and $(-1.46, 1.641)$.

13. (2 points) An IGARCH model is fitted to r_t. See **m5**. Write down the fitted model, including mean equation.

    Answer: $r_t = a_t$, $a_t = \sigma_t \epsilon_t$ with ϵ_t being iid N(0,1). The volatility equation is $\sigma_t^2 = (1 - 0.922) a_{t-1}^2 + 0.922 \sigma_{t-1}^2$.

14. (2 points) Based on the output provided, is the IGARCH model adequate? Why?

    Answer: No, because the Ljung-Box statistics of the squared residuals show Q(10) = 21.45 with p-value 0.0181.

15. (2 points) From the IGARCH model, we have $r_{1841} = -0.0955$ and $\sigma_{1841} = 0.745$. Compute the 1-step ahead forecast of r_t and its volatility at the forecast origin $t = 1841$?

    Answer: $r_{1841}(1) = 0$ and $\sigma_{1841}^2(1) = 0.078(-0.0955)^2 + 0.922(0.745)^2 = 0.512$ so that $\sigma_{1841}(1) = 0.716$.

16. (3 points) A GARCH-M model is entertained for the r_t series. See **m6**. Write down the fitted model.

    Answer: $r_t = 0.076 + 0.014 \sigma_t^2 + a_t$, $a_t = \sigma_t \epsilon_t$ with ϵ_t being iid N(0,1). The volatility equaiton is $\sigma_t^2 = 0.0285 + 0.105 a_{t-1}^2 + 0.879 \sigma_{t-1}^2$.

17. (2 points) Based on the fitted model **m6**, is the risk premium statistically significant? Why?

    Answer: No, the t-ratio of risk premium is 0.65 with p-value 0.51.

18. (4 points) A Threshold GARCH model with standard Student-t innovations is considered. See **m7**. Write down the fitted model.

Answer: $r_t = 0.0785 + a_t$, $a_t = \sigma_t \epsilon_t$ with ϵ_t being iid $t^*_{7.961}$. The volatility equation is $\sigma_t^2 = 0.0324 + (0.0524 + 0.992 N_{t-1}) a_{t-1}^2 + 0.874 \sigma_{t-1}^2$, where $N_{t-1} = 1$ if and only if $a_{t-1} < 0$.

19. (2 points) Based on the fitted TGARCH model **m7**, is the leverage effect statistically significant? Why?

    Answer: Yes, the t-ratio is 3.542 with p-value 0.0004.

20. (2 points) Among models { **m7, m4, m2** }, which one is preferred? Why?

    Answer: Model m7, because it has the smallest AIC criterion.

Problem C. (12 points) The consumption of natural gas in northern American cities depends heavily on temperature and daily activities. Let y be the daily sendout of natural gas and DHD be the degrees of heating days defined as DHD = $65^o F$ minus daily average temperature. Also, let x_1 = DHD, x_2 = lag-1 DHD, x_3 = windspeed, and x_4 be the indicator variable for weekend. The data were collected for 63 days. Statistical analysis is included in the attached R output. Answer the following questions.

1. (2 points) A multiple linear regression is applied. Write down the fitted model, including the R^2.

   Answer: The multiple linear regression is

   $$y = 1.858 + 5.875 x_1 + 1.405 x_2 + 1.315 x_3 - 15.857 x_4 + e, \quad \sigma = 18.32, \quad R^2 = 0.9521.$$

2. (3 points) Model checking shows that the residuals have serial correlations, and the AIC selects an AR(8) model. After removing insignificant parameters, we have the model **n5**. Write down the fitted model. Is the model adequate? Why?

   Answer: The model is

   $$(1 - .536B - .368B^7 - .261B^8)(y_t - 5.767 x_{1t} - 1.475 x_{2t} - 1.319 x_{3t} + 10.13 x_{4t}) = a_t,$$

   where $\sigma^2 = 188.1$. Model checking shows that the residuals have no serial correlations with $Q(10) = 7.28$ with p-value 0.698. The model is adequate.

3. (3 points) Let ϕ_i be the lag-i AR coefficient. It is seen that $\phi_1 \times \phi_7 \approx 0.2$, which is not far away from $-\phi_8$, especially in view of its standard error. This is indicative of a multiplicative model. Therefore, we fit a seasonal model. Denoted by **n6**. Write down the fitted seasonal model. Is the model adequate? Why?

Answer: The model is

$$(1-0.536B)(1-0.368B^7)(y_t-5.765x_{1t}-1.473x_{2t}-1.282x_{3t}+9.595x_{4t}) = a_t,$$

where $\sigma^2 = 189.5$. The model is also adequate as its residuals have no serial correlations; see Q(10) = 6.579 with p-value 0.76.

4. (2 points) Compare models **n5** and **n6**. Which one is preferred? Why?

   Answer: For in-sample fit, model n6 is preferred as it has lower AIC value.

5. (2 points) The weekend effect is rather significant in the multiple linear regression model of Question 1, but it is not so in the seasonal model. Why?

   Answer: Part of the weekend effects belong to the seasonality as the period is 7.

Problem D. (13 points) Consider the monthly log return of Decile 10 portfolio of CRSP from 1961,1 to 2014.12 with $T = 648$. The returns include dividends. Let r_t denote the monthly log return. Answer the following questions based on the attached R output.

1. (2 points) The ACF of r_t shows $\hat{\rho}_1 = 0.203$ and $\hat{\rho}_{12} = 0.127$. Test $H_0 : \rho_{12} = 0$ versus $H_a : \rho_{12} \neq 0$. Compute the test statistic and draw the conclusion.

   Answer: $t = \frac{0.127}{\sqrt{1/648}} = 3.23$, which is greater than 1.96. Therefore, ρ_{12} is not zero.

2. (2 points) The $\hat{\rho}_{12}$ is likely due to the *January effect*. To remove ρ_{12}, one can use a simple linear regression with January dummy variable. The fitted model is $r_t = 0.0064 + 0.06926\text{Jan}_t + \epsilon_t$. Let $\tilde{r}_t = r_t - 0.06928\text{Jan}_t$ be the adjusted log returns of Decile 10 portfolio. Several models were entertained for $\tilde{r}_t$; see model **g1, g2, g3** and **g4**. Which model is preferred? Why?

   Answer: Models g3 and g4 are identical as expected. They are preferred over g1 and g2 based on the AIC criterion.

3. (2 points) Write down the fitted model **g4**.

   Answer: $r_t = 0.006 + a_t + 0.222a_{t-1}$, $a_t = \sigma_t\epsilon_t$ with ϵ_t being iid $t^2_{5.729}$. The volatility equation is

$$\sigma_t^2 = 0.00031 + 0.129(|a_{t-1}| - 0.319a_{t-1})^2 + 0.79\sigma_{t-1}^2.$$

4. (2 points) Based on the fitted model **g4**. Is the leverage effect significant? Why?

   Answer: Yes, the t-ratio is 2.317 with p-value 0.0205.

5. (3 points) The average of fitted volatility is 0.0614 and the 1% quantile of the residuals of model **g4** is -0.171. Compute the ratio $\frac{\sigma_t^2(-0.171)}{\sigma_t^2(0.171)}$ of model **g4**, where $\sigma_t^2(a_t)$ denotes the conditional variance of the series when innovation is a_t.

   Answer:

$$\frac{\sigma_t^2(-.171)}{\sigma_t^2(.171)} = \frac{0.00031 + 0.129(.171 - 0.319(-.171))^2 + 0.79(0.0614)^2}{0.00031 + 0.129(.171 - 0.319(.171))^2 + 0.79(.0614)^2} = 1.955.$$

6. (2 points) Consider the forecasts of model **g4**. Why is the 1-step ahead forecast of $\tilde{r}_t$ different from multi-step ahead forecasts?

   Answer: Because of the MA(1) model of the mean equation.