

Solutions to Midterm

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

1. Describe two methods for choosing a time series model.

Answer: Any two of (a) Information criteria such as AIC or BIC, (b) Out-of-sample forecasts, and (c) ACF and PACF of the series.

2. Describe two applications of volatility in finance.

Answer: Any two of (a) derivative (option) pricing, (b) risk management, (c) portfolio selection or asset allocation.

3. Give two applications of seasonal time series models in finance.

Answer: (a) Earnings forecasts and (b) weather-related derivative pricing or risk management.

4. Describe two weaknesses of the ARCH models in modelling stock volatility.

Answer: Any two of (a) symmetric response to past positive and negative shocks, (b) restrictive, (c) Not adaptive, and (d) provides no explanation about the source of volatility clustering.

5. Give two empirical characteristics of daily stock returns.

Answer: any two of (a) heavy tails, (b) non-Gaussian distribution, (c) volatility clustering.

6. The daily simple returns of Stock A for the last week were 0.02, 0.01, -0.005, -0.01, and 0.025, respectively. What is the weekly log return of the stock last week? What is the weekly simple return of the stock last week? **Answer:** Weekly log return is 0.03938; weekly simple return is 0.04017.

7. Suppose the closing price of Stock B for the past three trading days were \$100, \$120, and \$100, respectively. What is the arithmetic mean of the simple return of the stock for the past three days? What is the geometric average of the simple return of the stock for the past three days?

Answer: Arithmetic mean = $\frac{1}{2} \left[\frac{120-100}{100} + \frac{100-120}{120} \right] = 0.017$. and the geometric mean is $\sqrt{\frac{120}{100} \times \frac{100}{120}} - 1 = 0$.

8. Consider the AR(1) model $r_t = 0.02 + 0.8r_{t-1} + a_t$, where the shock a_t is normally distributed with mean zero and variance 1. What are the variance and lag-1 autocorrelation function of r_t ?

Answer: $\text{Var}(r_t) = \frac{1}{1-0.8^2} = 2.78$ and the lag-1 ACF is 0.8.

9. **For problems 6 and 7**, suppose the daily return r_t , in percentages, of Stock A follows the model $r_t = 1.0 + a_t + 0.3a_{t-1}$, where $a_t = \sigma_t \epsilon_t$ with $\sigma_t^2 = 1.0 + 0.4a_{t-1}^2$ and ϵ_t being standard normal. What is the unconditional variance of a_t ? What is the variance of r_t ?

Answer: $\text{Var}(a_t) = \frac{1}{1-0.4} = 1.67$. $\text{Var}(r_t) = (1 + 0.3^2)\sigma_a^2 = 1.82$.

10. Suppose that $a_n = 3.0$, what is the 1-step ahead forecast for r_{n+1} at the forecast origin n ? What is the 1-step ahead volatility forecast of r_t at the forecast origin n ?

Answer: $r_n(1) = 1 + 0.3a_n = 1.9$, and $\sigma_n^2(1) = 1 + 0.4a_n^2 = 4.6$.

11. Consider the simple AR(1) model $r_t = 100 + 0.8r_{t-1} + a_t$, where a_t is normally distributed with mean zero and variance 10. Is the r_t series mean-reverting? If yes, what is the half-life of the series?

Answer: Yes, r_t series is mean-reverting. The half-life is $\ln(0.5)/\ln(0.8) = 3.11$.

12. Describe two test statistics for testing the ARCH effect of an asset return series. Write down the associated null hypotheses.

Answer: (a) The Ljung-Box statistic $Q(m)$ of the squared shocks, i.e. a_t^2 . The null hypothesis is $H_o : \rho_1 = \rho_2 = \dots = \rho_m = 0$, where ρ_i is the lag- i ACF of a_t^2 . (b) The Engle F-test for the regression $a_t^2 = \beta_0 + \beta_1 a_{t-1}^2 + \dots + \beta_m a_{t-m}^2 + e_t$. The null hypothesis is $H_o : \beta_1 = \beta_2 = \dots = \beta_m = 0$.

13. Consider the following two IGARCH(1,1) models for percentage log returns:

$$\text{Model A : } \sigma_t^2 = 1.0 + 0.1a_{t-1}^2 + 0.9\sigma_{t-1}^2$$

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

Suppose that $\sigma_{100}^2 = 20$ and $a_{100} = -2.0$. What are the 3-step ahead volatility forecasts for Models A and B?

Answer: For model A: 3-step ahead volatility forecast is $\sigma_{100}^2(3) = 2 + (1.0 + 0.1 \times (-2.0)^2 + 0.9 \times 20) = 21.4$. For model B, the 3-step ahead volatility forecast is $\sigma_{100}^2(3) = 0.1(-2.0)^2 + 0.9 \times 20 = 18.4$.

14. Consider the following two models for the log price of an asset:

$$\text{Model A : } p_t = p_{t-1} + a_t$$

$$\text{Model B : } p_t = 0.00001 + p_{t-1} + a_t$$

where the shock a_t is normally distributed with mean zero and variance $\sigma^2 > 0$. Suppose further that $p_{100} = 5$. Let $p_n(\ell)$ be the ℓ -step ahead forecast at the forecast origin n . What are the point forecasts $p_{100}(\ell)$ for both models as $\ell \rightarrow \infty$?

Answer: For model A, $p_{100}(\ell) = 5$ for all ℓ . For model B, $p_{100}(\ell)$ converges to infinity as $\ell \rightarrow \infty$.

15. Suppose that we have $T = 1000$ daily log returns for the Decile 1 portfolio. Suppose further that the sample autocorrelation at lag-12 is $\hat{\rho}_{12} = 0.15$. Test the hypothesis $H_o : \rho_{12} = 0$ against the alternative hypothesis $H_a : \rho_{12} \neq 0$. Compute the test statistic and draw your conclusion.

Answer: $t = \frac{0.15}{1/\sqrt{1000}} = \sqrt{1000} \times 0.15 = 4.74$, which is highly significant. Thus, the lag-12 ACF is not zero.

Problem B. (20 pts) It is well-known in economics that growth rate of the domestic gross product (GDP) is negatively correlated with the change in unemployment rate. Consider the U.S. quarterly real GDP and unemployment rate from the first quarter of 1948 to the first quarter of 2008. Let $\text{dgd}p_t$ be the growth rate of the GDP, i.e. $\text{dgd}p_t = \ln(\text{GDP}_t) - \ln(\text{GDP}_{t-1})$, and dun_t be the change in unemployment rate, i.e. $\text{dun}_t = U_t - U_{t-1}$ with U_t being the civilian unemployment rate. The data were seasonally adjusted and obtained from the Federal Reserve Bank at St. Louis. The sample size after the differencing is 240. Use the attached R output to answer the following questions.

1. (5 points) Write down the fitted linear regression model with $\text{dgd}p_t$ and dun_t representing the dependent and independent variable, respectively, including residual standard error. What is the R^2 of the linear regression? Is the fitted model adequate? Why?

Answer: The fitted linear regression is

$$\text{dgd}p_t = 0.017 - 0.017\text{dun}_t + e_t, \quad \hat{\sigma}_e = 0.0088.$$

The R^2 is 0.37. The model is not adequate because the $Q(m)$ statistics of the residuals show that the residuals have serial correlations, i.e. $Q(12) = 219.4$ with p-value close to zero.

2. (5 points) To take care of the serial correlations in the residuals, a linear regression model with time-series errors is built for the two variables. Write down the fitted model, including the residual variance.

Answer: The fitted linear regression model with time-series errors is

$$(1 - 0.21B - 0.12B^2)(1 - 0.86B^4)(\text{dgd}p_t - 0.017 + 0.018\text{dun}_t) = (1 - 0.72B^4)a_t, \quad \hat{\sigma}_a^2 = 6.01 \times 10^{-5}.$$

3. (2 points) Is the model in Question 2 adequate? Why?

Answer: Yes, the model is adequate. The $Q(m)$ statistics of the residuals fail to indicate the existence of any serial correlations. We have $Q(12) = 17.53$ with p-value 0.13.

4. (4 points) Based on the fitted model in Question 2, is the growth rate of GDP negatively correlated with the change in unemployment rate? Why?

Answer: Yes, the growth rate of GDP is negatively related to the change in unemployment rate. The estimated coefficient is -0.018 which is highly significant, because its standard error 0.0014 is small, resulting in a large t -ratio.

5. (4 points) To check the predictive power of the model, it was re-estimated using the first 236 data points. This re-fitted model is used to produce 1-step to 4-step ahead forecasts at the forecast origin $t = 236$. The actual value of the GDP growth rates are also given. Construct the 1-step ahead 95% interval forecast of the model. Is the actual growth rate in the forecasting interval?

Answer: The 95% interval forecast is $0.012 \pm 1.96 \times 0.0077$, i.e. $[-0.0031, 0.027]$. The actual value is 0.0159, which is in the interval.

Problem C. (16 pts) Consider the quarterly earnings per share of the Microsoft stock from the first quarter of 1992 to the first quarter of 2008. The data were obtained from First Call. To take the log transformation, **we add 0.5 to all data points**. The R output is attached. Let $x_t = \ln(y_t + 0.5)$ 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: The fitted model is

$$(1 - B)r_t = (1 - 0.70B + 0.39B^2)(1 + 0.39B^4)a_t, \quad \hat{\sigma}_a^2 = 0.0016,$$

where $r_t = \ln(x_t + 0.5)$ with x_t being the earnings per share.

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

Answer: No, the $Q(m)$ statistics of the residuals give $Q(12) = 9.65$ with p-value 0.65.

3. (4 points) Let $T = 65$ be the forecast origin, where T is the sample size. Based on the fitted model, and, for simplicity, use the relationship $y_t = \exp(x_t) - 0.5$, what are the 1-step and 2-step ahead forecasts of earnings per share for the Microsoft stock?

Answer: The 1-step and 2-step earnings forecasts are 0.56 and 0.54, respectively.

4. (2 points) 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 $t = \frac{0.3912}{0.1442} = 2.71$ with two-sided p-value 0.0067. Thus, the seasonal MA coefficient θ_4 is significantly different from zero.

5. (3 points) Consider the regular (i.e., non-seasonal) part of the MA model. Is it invertible? Why?

Answer: Yes, it is invertible, because the polynomial $1 - 0.6953x + 0.3889x^2$ has roots $0.89 \pm 1.33i$ so that the absolute value of the roots (Mod in R) is 1.6, which is greater than 1. [If you compute the roots of $x^2 - 0.6953x + 0.3889$, the the absolute value of the roots is less than 1.]

Problem D. (34 pts) Consider the daily log returns of the Starbucks stock, in percentages, from January 1993 to December 2007. The relevant R output is attached. Answer the following questions.

1. (2 points) Is the mean log return significant different from zero? Why?

Answer: No, the basic statistics show the 95% confidence interval of the mean is $[-0.0103, 0.1624]$, which contains zero.

2. (2 points) Is there any serial correlation in the log return series? Why?

Answer: Yes, the $Q(m)$ statistics show $Q(15) = 38.39$ with p-value 0.0008.

3. (2 points) An MA model is used to handle the mean equation, which appears to be adequate. Is there any ARCH effect in the return series? Why?

Answer: Yes, because the $Q(m)$ statistics of the squared residuals show $Q(15) = 112.61$ with p-value close to zero.

4. (6 points) A GARCH(1,1) model with Student- t distribution is used for the volatility equation. Write down the fitted model, including the degrees of freedom of the Student- t innovations and mean equation.

Answer: The fitted model is

$$r_t = 0.037 + a_t - 0.043a_{t-1} - 0.048a_{t-2}, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim t_{5.27}.$$

$$\sigma_t^2 = 0.012 + 0.026a_{t-1}^2 + 0.973\sigma_{t-1}^2.$$

5. (4 points) Since the constant term of the GARCH(1,1) model is not significantly different from zero at the 1% level, an IGARCH(1,1) model is used. Write down the fitted IGARCH(1,1) model, including the mean equation.

Answer: The fitted IGARCH(1,1) model is

$$r_t = 0.077 + a_t - 0.029a_{t-1} - 0.044a_{t-2}, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim N(0, 1).$$

$$\sigma_t^2 = 0.022a_{t-1}^2 + 0.978\sigma_{t-1}^2.$$

6. (3 points) Is the IGARCH(1,1) model adequate? Why? What is the 3-step ahead volatility forecast with the last data point as the forecast origin?

Answer: Yes, the $Q(m)$ statistics for the standardized residuals give $Q(10) = 1.77$, $Q(15) = 10.79$, and $Q(20) = 18.66$. The p-values of these statistics are all greater than 0.05. In addition, the $Q(m)$ statistics of the squared standardized residuals also have large p-values.

The 3-step ahead volatility forecast is $\sqrt{3.779} = 1.94$.

7. (5 points) A GJR (or TGARCH) model with Student- t distribution is also fitted to the log return series. Write down the fitted model, including the mean equation and all parameters.

Answer: The fitted GJR model is

$$r_t = 0.032 + a_t - 0.043a_{t-1} - 0.048a_{t-2}, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim t_{5.31}.$$

$$\sigma_t^2 = 0.015 + (0.021 + 0.017N_{t-1})a_{t-1}^2 + 0.970\sigma_{t-1}^2,$$

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

8. (2 points) Is the fitted GJR (or TGARCH) model adequate? Why?

Answer: Yes, the $Q(m)$ statistics of the standardized residuals and those of the squared standardized residuals all have large p-values.

9. (2 points) Among the GARCH(1,1), IGARCH(1,1) and GJR(1,1) models, which one is preferred? Why?

Answer: The GJR(1,1) model because it has the smallest AIC value.

10. (2 points) Is the leverage effect of the GJR model significant? Why?

Answer: Yes, the t -ratio of the leverage parameter is 2.01, which is significant at the 5% level.

11. (4 points) To better understand the leverage effect, use the fitted GJR to calculate the ratio $\frac{\sigma_t^2(a_{t-1}=-5.10)}{\sigma_t^2(a_{t-1}=5.10)}$, assuming $\sigma_{t-1}^2 = 7.5$.

Answer:

$$\begin{aligned} \frac{\sigma_t^2(a_{t-1} = -5.10)}{\sigma_t^2(a_{t-1} = 5.10)} &= \frac{0.0154 + 0.0379 \times (-5.10)^2 + 0.97 \times 7.5}{0.0154 + 0.0208 \times (5.10)^2 + 0.97 \times 7.5} \\ &= 1.057. \end{aligned}$$