

Solutions to Final Exam

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

1. Suppose that the price  $P_t$  of a stock follows the geometric Brownian motion

$$dP_t = \mu P_t dt + \sigma P_t dw_t,$$

where  $\mu$  and  $\sigma$  are constant and  $w_t$  is the standard Brownian motion. What is the model for the square-root of the price  $G(P_t) = P_t^{0.5}$ ?

Answer: The partial derivatives are

$$\frac{\partial G(P_t)}{\partial P_t} = 0.5P_t^{-0.5}, \quad \frac{\partial G(P_t)}{\partial t} = 0, \quad \frac{\partial G(P_t)^2}{\partial P_t^2} = -0.25P_t^{-1.5}.$$

By Ito's Lemma, we have

$$dP_t^{0.5} = 0.5(\mu - 0.25\sigma^2)P_t^{0.5}dt + 0.5\sigma P_t^{0.5}dw_t,$$

which remains a geometric Brownian motion.

2. Consider the daily log returns, in percentages, of the stock of Boeing Company from January 1997 to December 2006 for 2516 observations. Summary statistics of the returns are in the attached output. Let  $\mu$  be the mean of the log return. Test  $H_o : \mu = 0$  vs  $H_a : \mu \neq 0$ . What is the test statistic? Draw your conclusion.

Answer: The  $t$ -ratio for the sample mean is  $t = 0.02628/0.04273 = 0.615$ , which is less than 2 so that we cannot reject the null hypothesis. In other words, there is no sufficient evidence to reject the hypothesis that the mean of the return is zero.

3. Again, consider the log returns of Boeing stock and the output. Let  $\rho_i$  be the lag- $i$  autocorrelation of the return. Test  $H_o : \rho_1 = \rho_2 = \dots = \rho_{10} = 0$  vs  $H_a : \rho_i \neq 0$  for some  $1 \leq i \leq 10$ . What is the test statistic? Draw your conclusion.

Answer:  $Q(10) = 17.33$  with p-value 0.067, which is greater than 0.05 so that we cannot reject the null hypothesis of zero serial correlations.

4. Again, consider the log returns of Boeing stock and the output. Is there any ARCH effect in the return? Why?

Answer: The  $Q(10)$  statistic for the squared residuals is 145.02 with p-value close to zero. Thus, there is significant ARCH effect in the data.

5. Consider a non-dividend paying stock. Suppose that the current price is  $P_t = \$30$  and the risk-free interest rate is 6% per annum. If the price of a European call option of the stock is \$3.10 when the strike price is \$31 and the time to expiration is 3 months. What is the price of a European put option of the stock with the same duration and strike price?

Answer: From the put-call parity,  $p_t = c_t + K \exp[-r(T-t)] - P_t = \$(3.10 + 31e^{-0.06 \times 0.25} - 30) = \$3.64$ .

6. Describe two methods to identify the order of an AR model.

Answer: (a) PACF and (b) Information criteria.

7. Give two weaknesses of GARCH models in modeling asset volatility.

Answer: Any two of (a) symmetric response between past positive and negative returns (i.e., no leverage effect), (b) restrictive, (c) provides no explanation for volatility, and (d) slow response to an isolated jump (or drop) in stock price.

8. Suppose that the true price of a stock follows the random walk model

$$P_t^* = P_{t-1}^* + a_t, \quad a_t \sim N(0, 2).$$

Suppose also that the true price is the average of the bid and ask prices and the bid-ask spread is 0.25. What is the impact of the bid-ask bounce on the serial correlation of the simple return of the stock?

Answer:  $\rho_1 = \frac{-0.25^2/4}{2+0.25^2/2} = -0.0156/2.0313 = -0.008$ .

9. Consider the daily log returns, in percentages, of the FedEx stock from January 1997 to December 2006 for 2516 observations. A GARCH(1,1) model is fitted to the data. The result indicates that the following IGARCH(1,1) model seems reasonable for the log returns:

$$r_t = a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim N(0, 1), \quad \sigma_t^2 = 0.98\sigma_{t-1}^2 + 0.02a_{t-1}^2.$$

In addition, we have  $a_{2516} = 0.544$  and  $\sigma_{2516}^2 = 1.95$ . What is the VaR for the next trading day for a position that long the stock valued at \$1 million dollars?

Answer:  $\sigma_{2517}^2 = 0.98 \times 1.95 + 0.02 \times 0.544^2 = 1.917$ . The 1% quantile is  $-2.326 \times \sqrt{1.917} = -3.2204$ . Consequently,  $\text{VaR} = \$1000000 \times 3.2204/100 = \$32204$ .

10. The computer output also shows that the correlation coefficient between the daily log returns of FedEx and Boeing stocks is 0.23. If you hold both stocks each valued at \$1 million dollars and you know that the VaR for the Boeing stock is \$28,000 for the next trading day. What is the VaR of your combined portfolio for the next trading day?

Answer:  $\text{VaR} = \sqrt{32204^2 + 2 \times 0.23 \times 32204 \times 28000 + 28000^2} = \$47285$ .

11. Describe two major difficulties in modeling the volatility series of multiple asset returns.

Answer: The dimension increases quickly and the volatility matrix must be positive definitely almost surely.

12. Suppose that the daily log returns, in percentages, of a stock follows the model

$$r_t = 0.2 + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim N(0, 1), \quad \sigma_t^2 = 0.94\sigma_{t-1}^2 + 0.06a_{t-1}^2.$$

Suppose that at the forecast origin  $T$ ,  $r_T = 0.54$  and  $\sigma_T^2 = 2.05$ . What is the 2-step ahead forecast  $r_T(2)$  of the return? What is the 2-step ahead volatility forecast  $\sigma_T(2)$ ?

Answer:  $r_T(2) = 0.2$ ,  $\sigma_T(1)^2 = 0.94 \times 2.05 + 0.06 \times 0.34^2 = 1.934$ . Since  $\sigma_T(2)^2 = \sigma_T(1)$ , we have  $\sigma_T(2) = 1.391$ .

13. Consider the growth rate of U.S. quarterly GDP from 1948 to 2004. An AR(3) model is fitted to the data. Write down the fitted model. Does the model imply the existence of business cycles? If yes, calculate the average length of business cycles.

Answer: The fitted model is

$$x_t = 0.31x_{t-1} + 0.12x_{t-2} - 0.11x_{t-3} + a_t,$$

where  $x_t = r_t - 0.085$ . Yes, the model implies the existence of business cycles. The average period is  $2\pi / \cos^{-1}(1.6098/2.0333) = 9.56$ .

14. Give two objectives of analysis of high-frequency financial data that cannot be achieved by using daily data.

Answer: Any two of (a) price discovery, (b) trading cost, (c) efficiency of market design, (d) effect of change in tick size, and (e) impact of price limit.

15. The square root of time rule of RiskMetrics is based on some critical assumptions. List two of these assumptions used.

Answer: Any two of (a) Gaussian distribution, (b) zero expectation of return, and (c) the specific IGARCH(1,1) model without a constant.

**Problem B.** (25 points) Consider the daily log returns, in percentages, of the Boeing stock and the output. Suppose that you hold a long position of the stock valued at \$1 million dollars. Answer the following questions.

1. Apply the traditional extreme value theory to the negative log returns with block size 63. What are the estimates of the three parameters  $k$ ,  $\alpha$ , and  $\beta$ ? Are these estimates statistically significant? Why?

Answer: The estimates and their standard errors (in parentheses) are  $-k = \xi = 0.6199(0.211)$ ,  $\alpha = \sigma = 1.42(0.28)$ ,  $\beta = \mu = 3.64(0.27)$ . The estimates are all significantly different from zero because their  $t$ -ratios are 2.94, 5.07, and 13.48, respectively, which are all greater than 1.96.

2. Based on the estimates of block size 63, what is the VaR of your position for the next trading day?

Answer: The 1% quantile of the log returns is

$$VaR = 3.64 - \frac{1.42}{0.62} [1 - (-63 \ln(0.99))^{-0.62}] = 4.39.$$

Consequently, VaR for the position is  $\$1000000 \times 4.39/100 = \$43900$ .

3. Turn to the approach of peaks over the threshold. Using threshold 2.0, we estimate the parameters of generalized Pareto distribution for the stock returns. Does the log returns have a heavy left tail? Why?

Answer: Yes, the log returns have a heavy left tail because the  $t$ -ratio of the estimated shape parameter is  $t = \frac{-0.215}{0.063} = -3.41$ , which is less than  $-2.326$ . Thus, we reject the null hypothesis of normal left tail at the 5% level.

4. Based on estimated generalized Pareto distribution, what is the VaR of your financial position for the next trading day? What is the VaR of your financial position for the next 10 trading days?

Answer: VaR for the position is  $\$1000000 \times 5.793/100 = \$57930$ . For the next ten trading day,  $VaR[10] = \$57930 \times (10)^{0.215} = \$95039$ .

5. (2 pts) Again, based on the fitted Pareto distribution with threshold 2.0, what is the expected shortfall when the 1% VaR is used?

Answer; The expected shortfall is  $\$1000000 \times 8.2665/100 = \$82665$ .

6. (3 pts) If empirical quantiles are used, what is the VaR of your financial position for the next trading day?

Answer:  $p_1 = \frac{25}{2516}$ ,  $p_2 = \frac{26}{2516}$ . Therefore,

$$VaR = \frac{p_2 - 0.01}{p_2 - p_1} r_{(25)} + \frac{0.01 - p_1}{p_2 - p_1} r_{(26)} = -5.601.$$

Therefore,  $VaR = \$1000000 \times 5.601/100 = \$56010$ .

**Problem C.** (25 pts) Again, consider the daily log returns, in percentages, of Boeing stock from January 1997 to December 2006. Suppose also that your position remains unchanged.

1. A GARCH(1,1) model with Gaussian distribution is fitted to the data. Based on the fitted model, what is the VaR of your financial position for the next trading day?

Answer: From the output, the forecasts for the log return is 0.089 and that for variance is 1.454. Therefore, the 1% quantile is  $0.089 - 2.326 \times \sqrt{1.454} = -2.716$ . Thus,  $VaR = \$1000000 \times 2.716/100 = \$27160$ .

2. A GARCH(1,1) model with Student- $t$  innovations is also fitted to the data. Write down the fitted model.

Answer: The fitted model is

$$r_t = 0.06 + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim t_{6,0.2}.$$

$$\sigma_t^2 = 0.054 + 0.045a_{t-1}^2 + 0.942\sigma_{t-1}^2.$$

3. Based on the fitted GARCH(1,1) model with Student- $t$  innovations, what is the VaR of your position for the next trading day? [The 99% quantile of  $t_6$  is 3.14.]

Answer: Using the output, the 1% quantile is

$$VaR = 0.06 - 3.14/\sqrt{6/4} \times \sqrt{1.674} = -3.233.$$

Therefore, VaR for the position is  $VaR = \$1000000 \times 3.233/100 = \$32330$ .

4. A GJR model is also fitted to the daily log returns of Boeing stock. Write down the fitted model. Is the model adequate? Why?

Answer: The fitted GJR model is

$$r_t = 0.051 + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim t_{6,14}.$$

$$\sigma_t = 0.061 + (0.028 + 0.037N_{t-1})a_{t-1}^2 + 0.939\sigma_{t-1}^2,$$

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

The model is adequate because the Ljung-Box statistics of the standardized residuals and the squared standardized residuals fail to indicate model inadequacy as the p-values of these test statistics are all greater than 0.05.

5. Based on the fitted GJR model and ignoring the constant term of the volatility equation, compute the leverage impact  $\frac{\sigma_t^2(\epsilon=-3)}{\sigma_t^2(\epsilon=3)}$ , where  $\epsilon$  is the standardized innovation.

$$\text{Answer: } \frac{\sigma_t^2(\epsilon=-3)}{\sigma_t^2(\epsilon=3)} = \frac{[(0.027+0.037)9+0.939]\sigma_{t-1}^2}{(0.027 \times 9+0.939)\sigma_{t-1}^2} = 1.28.$$

**Problem D.** (20 points) To study the direction of price movement of FedEx stock, we define the dependent variable  $y_t$  as

$$y_t = \begin{cases} 1 & \text{if } r_t > 0 \\ 0 & \text{otherwise,} \end{cases}$$

where  $r_t$  is the daily simple return of the stock. The independent variables are  $r_{t-2}$  and  $r_{t-3}$ . [We started with five lags, but only kept those that are statistical significant.] Two methods are discussed in class to model the direction of price movement. The first method is the neural network and the second method is linear logistic regression. For the neural network, we use a simple 2-2-1 network with input variables  $r_{t-2}$  and  $r_{t-3}$ . Computer output is attached. Answer the following questions.

1. Write down the model for the two nodes in the hidden layer.

Answer: The fitted equations are

$$h_{1t} = \frac{\exp(1.34 - 1.13r_{t-2} - 0.75r_{t-3})}{1 + \exp(1.34 - 1.13r_{t-2} - 0.75r_{t-3})},$$
$$h_{2t} = \frac{\exp(3.10 + 0.12r_{t-2} + 0.07r_{t-3})}{1 + \exp(3.10 + 0.12r_{t-2} + 0.07r_{t-3})}.$$

2. Write down the model for the output node.

Answer: The fitted function is

$$h(z_t) = \begin{cases} 1 & \text{if } z_t > 0 \\ 0 & \text{if } z_t \leq 0 \end{cases}$$

where  $z_t = -0.71 + 1.06h_{1t} - 0.17h_{2t} - 5.61r_{t-2} - 3.94r_{t-3}$ .

3. Write down the model for the linear logistic regression.

Answer:  $\text{logit}(p_i) = -0.031 - 5.80r_{t-2} - 4.11r_{t-3}$  or

$$p_i = \frac{\exp(-0.031 - 5.80r_{t-2} - 4.11r_{t-3})}{1 + \exp(-0.031 - 5.80r_{t-2} - 4.11r_{t-3})}.$$

4. Suppose  $x_{n-2} = -0.018$  and  $x_{n-1} = 0.015$ . Based on the fitted logistic regression, what is  $P(y_{n+1} = 1)$ ?

Answer:  $P(y_{n+1} = 1) = p_i = \frac{\exp(-0.031 - 5.80(0.015) - 4.11(-0.018))}{1 + \exp(-0.031 - 5.80(0.015) - 4.11(-0.018))} = \frac{0.957}{1 + 0.957} = 0.489$ .

5. Consider the fitted values of the neural network and the logistic regression. Let “dif” be the difference between the fitted values of the two methods. The summary statistics of “dif” are given. Is there any major difference between the two methods? Why?

Answer: There is no major difference between the two methods. The maximum difference of the fitted probabilities is 0.0012 and the average difference is  $-0.00004$ .