Graduate School of Business, University of Chicago
Business 41202, Spring Quarter 2004, Mr. Ruey S. Tsay

Midterm – Solution

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

1. For problems 1 to 6, consider the daily log return of the exchange rate between U.S. Dollar and Japanese Yen from January 3, 2000 to March 26, 2004 for 1062 observations. The returns are in percentages. Summary statistics of the log returns from SCA and Splus are attached. See Page 1 of output. What is the sample skewness of the log returns?

Answer: $\hat{S}(x) = -0.2920$.

2. Is the mean of log returns significantly different from zero? Why?

Answer: No, $t = 0.2072 < 1.96$, which is the critical value of a normal distribution $z_{0.025}$.

3. Test the null hypothesis that the excess kurtosis of the log returns is zero. Draw your conclusion.

Answer: `SCA` reports the excess kurtosis. The null hypothesis is $H_0 : K(x) - 3 = 0$. The t-test is

$$t = \frac{\hat{K}(x) - 3}{\sqrt{24/T}} = \frac{1.4874}{0.15} = 9.916 > 1.96;$$

therefore, we reject the null hypothesis that the excess kurtosis is zero.

4. The output also provides the last 3 data points of the percentage log returns. Transform the last data point -0.123 into percentage simple return.

Answer: Since the return is in percentage, $R_t = [\exp(r_t/100) - 1] \times 100 = -0.1229\%$.

5. Write down the null hypothesis of the Ljung-Box statistic $Q(12)$ for testing that the daily log returns have no serial correlations.

Answer: $H_0 : \rho_1 = \rho_2 = \cdots = \rho_{12} = 0$, where ρ_i is the lag-i ACF of the log returns.

6. Are the log returns of exchange rate serially correlated? Answer this question using the Q(12) statistic.

Answer: No, $Q(12)$ =7.0 with *p-value* 0.858 > 0.05 (the 5% significance level.)

7. Give two possible applications of asset return volatility.

Answer: Any two of the followings: (a) derivative pricing, (b) asset allocation,(c) interval forecasts, or (d) riak management.

8. Give two approaches that researchers commonly used to construct or estimate the volatility of an asset return.

Answer: Any two of the following: (a) Econometric modeling, (b) realized volatility, (c) implied volatility.

9. Define the lag-3 sample partial autocorrelation function (PACF) of the return series $\{r_1, r_2, \cdots, r_T\}$.

Answer: Perform the AR regression $r_t = \phi_0 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \phi_3 r_{t-3} + a_t$. The least squares estimate of ϕ_3 is the lag-3 sample PACF.

10. Describe a method to check the serial correlations of a linear regression model with time series errors.

Answer: Use the Ljung-Box Q(m) statistics of the residuals, including DW statistics as a special case.

11. Consider the seasonal MA model $x_t = 0.5 + (1 - 0.4B)(1 - 0.6B^{12})a_t$, where $\{a_t\}$ is a white noise series. Besides the lag-0 serial correlation, how many autocorrelation functions (ACF) of x_t are non-zero? Provide the list of non-zero ACFs.

Answer: There are 4. Namely, ρ_1, ρ_{11}, ρ_{12}, and ρ_{13}.

12. Consider the simple ARCH(1) model

$$r_t = 0.021 + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = 0.116 + 0.42 a_{t-1}^2.$$

What is the unconditional variance of a_t? That is, what is $\text{Var}(a_t)$?

Answer: $\text{Var}(a_t) = \alpha_0/(1 - \alpha_1) = 0.116/(1 - 0.42) = 0.2$.

13. For problems 13-15, consider the following model for the log return r_t of an asset

$$r_t = 0.016 + 0.2 r_{t-1} + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = 0.1 + 0.1 a_{t-1}^2 + 0.8 \sigma_{t-1}^2,$$

where $\{\epsilon_t\}$ are *iid* standard Gaussian random variables. What is the expected value of the log return r_t?

Answer: $E(r_t) = \phi_0/(1 - \phi_1) = 0.016/(1 - 0.2) = 0.02$.

14. What is the unconditional variance of a_t?

Answer: $\text{Var}(a_t) = \alpha_0/(1 - \alpha_1 - \beta_1) = 0.1/(1 - 0.1 - 0.8) = 1$.

15. Suppose that $r_{100} = 0.015$, $a_{100} = -0.05$ and $\sigma_{100}^2 = 0.1$. Use $t = 100$ as the forecast origin. What is the 1-step ahead forecast of r_t? What is the 1-step ahead prediction of the volatility?

Answer: The 1-step ahead prediction of the return is

$$\hat{r}_{100}(1) = 0.016 + 0.2(r_{100}) = 0.016 + 0.2(0.015) = 0.019,$$

or 0.19%

The 1-step ahead prediction of the volatility is

$$\hat{\sigma}_{100}^2(1) = 0.1 + 0.1(-0.05)^2 + 0.8(0.1) = 0.18025.$$

The standard deviation is $\sqrt{0.1802} = 0.4246$, or 42.46%.

Problem B. (20 pts) This problem is concerned with the analysis of daily average temperature at the O'Hare International Airport (ORD) from 1979 to 2001. The data for February 29 were deleted resulting in 8395 observations. Because of the strong seasonal pattern, we took the seasonal difference $(1 - B^{365})$ in the analysis. SCA output is attached. See page 2 of output.(Splus took forever to estimate the model. Output is not included.) Due to long seasonality, output is edited to simplify the attachment. Answer the following questions.

1. (2 points) For the temperature data, seasonal differencing is sufficient, i.e. no further first differencing is necessary, why?

Answer: Temperature in ORD should be stationary after removing the seasonal pattern.

2. From the fitted final model, a regular AR(2) factor is $(1 - 0.88B + 0.13B^2)$. Does this factor contain stochastic cycle? Why?

Answer: Given that we have an AR(2) component there will be an stochastic cycle if and only if $\sqrt{\phi_1^2 + 4\phi_2} < 0$. Using the estimates obtained from SCA, we get $\hat{\phi}_1^2 + 4\hat{\phi}_2 = 0.2526$, which is positive. Therefore, there is no stochastic cycle.

3. (6 points) Write down the final fitted model for the daily temperature at ORD Airport.

Answer: Let X_t represent the temperature at time t. The fitted model is

$$(1 - 0.88B + 0.13B^2)(1 - B^{365})X_t = (1 - 0.15B^2)(1 - 0.79B^{365})a_t$$

where $a_t \overset{\text{iid}}{\sim} N(0, \sigma^2)$ with $\hat{\sigma}_a = 0.068$.

4. Is there any serial correlation in the residuals of the fitted model? Use Q(12) and lag-365 ACF of the residuals to answer the question.

Answer: No, the Ljung-Box statistic with 12 lags is $Q(12) = 12.6$, which compared to a χ^2 with $12 - 4 = 8$ d.f.'s gives a *p-value* of $0.126 > 0.05$.

The autocorrelation for lag-365 is 0.01 which is within the ± 2 confidence bands.

5. Is there any evidence of ARCH effects in the residuals of the fitted final model? Why?

Answer: Yes, The Ljung-Box statistics for the squared residuals are huge, and all the individual lags (up to 12) are also significant.

Problem C. (20 pts) This problem is a continuation of the analysis of HW#2 concerning monthly simple returns of Decile 1 from 1960 to 2003. It turns out that the seasonal effect shown in ACF is due to the *January effect*. S-plus output is attached. See page 5. Answer the following questions.

1. (4 points) Write down the linear regression model for the Decile 1 returns with January indictor as the explanaroty variable. Is the model adequate? Why?

Answer: Let R_t be the monthly simple return of the Decile 1 at time t; and let Jan_t be the indicator (dummy) variable that takes the value 1 if R_t corresponds to a January return, and 0 otherwise. The simple linear model is

$$R_t = 0.0066 + 0.1365\text{Jan}_t + a_t,$$

where $a_t \overset{\text{iid}}{\sim} N(0, \sigma^2)$ with $\hat{\sigma}_a = 0.06981$.

Both parameters are significant, implying that there is indeed a January effect. However, the *p-value* for the Ljung-Box statistic is very small; thus, there is serial correlation in the residuals.

2. (4 points) Consider the ACF of the residuals of the fitted regression model. Is there significant evidence of seasonal effect? Briefly justify your answer.

Answer: No, there is no significant seasonal serial correlations at lags 12 or 24.

3. Adding an AR(1) term to the linear regression. Is the AR(1) coefficient estimate significant at the 5% level? Do the residuals of the modified regression show any evidence of ARCH effects? Why?

Answer: Yes, the AR(1) coefficient 0.2476 is significant at the 5% level, since the *p-value* is basically zero. Thus, we reject the null hypothesis that $\phi_1 = 0$.

Yes, the residuals show an ARCH effect, since the *p-value* of the Ljung-Box statistic with 5 lags is $0.013 < 0.05$.

4. (6 points) Write down the fitted mean equation and volatility equation of the joint model with generalized error distribution.

Answer: The full model is

$$\begin{aligned} R_t &= 0.00122 + 0.1125\text{Jan}_t + 0.336R_{t-1} + a_t, \quad a_t = \sigma_t \varepsilon_t \\ \sigma_t^2 &= 0.00046 + 0.1054a_{t-1}^2 + 0.7905\sigma_{t-1}^2 \end{aligned}$$

where ε_t follows the *generalized error distribution* (ged) with parameter 1.141 (std err 0.073).

5. Is the fitted model adequate? Briefly justify your answer.

Answer: Yes, it is adequate. All parameters are significant (with exception of the constant term in the mean, which corresponds to the average return from February to December.) The *p-values* for the standardized and squared standardized residuals are considerably larger than 0.05. And the estimated shape parameter of the "ged" is less than 2, which corresponds to a distribution with tails heavier than normal.

Problem D. (30 pts) This problem is concerned with daily log returns of Yahoo stock from January 1997 to December 2003 with 1761 observations. Splus output is attached. See page 7 of output. Answer the following questions.

1. (4 points) Is there any ARCH effect in the log return series? Why?

Answer: Yes, there appears to be an ARCH effect. The autocorrelation test indicates no serial correlations in the data, but the ARCH test with 12 lags shows a *p-value* of being0. Thus, we reject the null of no ARCH effects.

2. Write down the fitted GARCH(1,1) model with conditional Gaussian distribution (both mean and volatility equations).

Answer: Let r_t be the log return at time t. The fitted model is

$$\begin{aligned} r_t &= 0.00324 + a_t, \quad a_t = \sigma_t \epsilon_t \\ \sigma_t^2 &= 3.185 \times 10^{-5} + 0.0668 a_{t-1}^2 + 0.9228 \sigma_{t-1}^2 \end{aligned}$$

where ϵ_t are *iid* $N(0,1)$.

3. (4 points) Is the fitted GARCH(1,1) model adequate? Why?

Answer: It is adequate, except for the normality assumption. All parameters are significant. The standardized and squared standardized residuals do not show serial correlations. However, the test for normality fails. This could be due to skewness and/or heavy tails.

4. (6 points) Write down the fitted EGARCH(1,1) model with leverage effect (both mean and volatility equations and the parameter of the conditional distribution used).

Answer: The fitted model is

$$\begin{aligned} r_t &= 0.0012 + a_t, \quad a_t = \sigma_t \epsilon_t \\ \ln(\sigma_t^2) &= -0.4108 + 0.2037(|\epsilon_{t-1}| - 0.243\epsilon_{t-1}) + 0.9583 \ln(\sigma_{t-1}^2), \end{aligned}$$

where ϵ_t follows the ged with shape paramter 1.1458.

5. Is the leverage effect statistically significant? Why?

Answer: Yes. Both parameters ω_1 and δ are significant, with t-statistics of 6.6 and -3.1, respectively. These are greater than 2 in absolut value.

Indeed, the estimated leverage effect is $\hat{\theta} = 0.204(-0.243) \approx -0.05$.

(One cansimply notes that the LEV(1) parameter estimate is significant.)

6. (6 points) To better understand the leverage effect, use the fitted EGARCH(1,1) model to calculate the ratio $\frac{\sigma_t^2(\epsilon=-2)}{\sigma_t^2(\epsilon=2)}$, where ϵ_t denotes the *iid* sequence of the innovations defined in class.

Answer: Let us start by parts. Recall that $h_t = \log(\sigma_t^2)$. Then,

$$h_t(\varepsilon = -2) = -0.4108 + 0.9583h_{t-1} + 0.2037(2 - 0.243(-2)).$$

On the other hand,

$$h_t(\varepsilon = 2) = -0.4108 + 0.9583h_{t-1} + 0.2037(2 - 0.243(2)).$$

Hence,

$$\frac{\sigma_t^2(\varepsilon = -2)}{\sigma_t^2(\varepsilon = 2)} = \frac{e^{0.2037(2)(1+0.243)}}{e^{0.2037(2)(1-0.243)}} = e^{0.1980} = 1.2190.$$

In other words, the volatility of a negative shock (with magnitude 2) is 21.9% higher than the volatility that with a positive shock of the same magnitude.