Booth School of Business, University of Chicago
Business 41910, Time Series Analysis, Autumn Quarter 2008, Mr. Ruey S. Tsay

Solutions to Final Exam

Problem A: (45 pts) Answer briefly the following questions. Each question has three points.

1. Define a temporary change in the mean for an ARMA(p, q) process at the time index $t = h$.

   Answer: Let $I_t^{(h)}$ be the indicator for the time index $t = h$. The model is

   $$Z_t = \frac{\omega}{1 - \delta B} I_t^{(h)} + X_t,$$

   where Z_t is the observed time series and X_t is the outlier-free ARMA(p, q) process, ω is the initial change in the level, and $0 < \delta < 1$ is the discounting rate of the intervention impact.

2. Consider the model

   $$z_t = \mu + \frac{\omega}{1 - B} I_t^{(h)} + X_t, \quad t = 1, \ldots, T,$$

   where μ is a constant, ω is a parameter, and $I_t^{(h)}$ is the indicator for the time index h. Suppose that X_t is a Gaussian white noise series with mean zero and variance σ^2. Derive the maximum likelihood estimate of ω? What is the distribution of the estimate?

   Answer: Under normaility, the MLE of μ is $\bar{z}_1 = \sum_{t=1}^{h-1} z_t/(h-1)$ and the MLE of $\mu + \omega$ is $\bar{z}_2 = \sum_{t=h}^{T} z_t/(T - h + 1)$. Using the properties of MLE, the MLE estimate of ω is $\hat{\omega} = \bar{z}_2 - \bar{z}_1$. Using properties of sample mean, $\bar{z}_1 \sim N(\mu, \sigma_1^2)$ and $\bar{z}_2 \sim N(\mu + \omega, \sigma_2^2)$, where σ_1^2 and σ_2^2 are

   $$\begin{aligned} \sigma_1^2 &= \sigma^2/(h-1) \\ \sigma_2^2 &= \sigma^2/(T - h + 1). \end{aligned}$$

   Consequently, $\hat{\omega} \sim N(\omega, \sigma^2[1/(h-1) + 1/(T - h + 1)])$, where we have used the independence of $\bar{z}_1$ and $\bar{z}_2$.

3. Consider the model of the prior question. Suppose that X_t follows the MA(1) model $X_t = a_t - \theta a_{t-1}$, where a_t is a Gaussian white noise series with mean zero and variance σ_a^2. A natural estimate of ω is $\hat{\omega} = \bar{z}_2 - \bar{z}_1$, where $\bar{z}_1 = \sum_{t=1}^{h-1} z_t/(h-1)$ and $\bar{z}_2 = \sum_{t=h}^{T} z_t/(T - h + 1)$. What is the distribution of this estimate?

   Answer: The key idea here is to work out the variances of $(h-1)\bar{z}_1$ and $(T - h + 1)\bar{z}_2$ under the MA(1) assumption. The distribution of $\hat{\omega}$ will be normal under the normality assumption. It turns out that

$$\begin{aligned}\text{Var}[\bar{z}_1] &= \frac{(1+\theta^2)\sigma_a^2 + (1-\theta)^2(h-2)\sigma_a^2}{(h-1)^2}\\ \text{Var}[\bar{z}_2] &= \frac{(1+\theta^2)\sigma_a^2 + (1-\theta)^2(T-h)\sigma_a^2}{(T-h+1)^2}.\end{aligned}$$

In addition, the covariance between $\bar{z}_1$ and $\bar{z}_2$ is $-\theta\sigma_a^2/[(h-1)(T-h+1)]$. Therefore, $\hat{\omega}$ is given by

$$\frac{(1+\theta^2)\sigma_a^2 + (1-\theta)^2(h-2)\sigma_a^2}{(h-1)^2} + \frac{(1+\theta^2)\sigma_a^2 + (1-\theta)^2(T-h)\sigma_a^2}{(T-h+1)^2} + \frac{2\theta\sigma_a^2}{(h-1)(T-h+1)}.$$

4. Consider the estimate $\hat{\omega}$ of the prior problem. Let $h = [T/2]$, which is the integer part of $T/2$. Derive the limiting distribution of $\sqrt{T}\hat{\omega}$ as $T \to \infty$ when (a) $\theta = 0.9$ and (b) $\theta = 0$. Discuss the impact of θ on the limiting distribution.

   Answer: From the prior solution, the asymptotic variance of $\sqrt{T}\hat{\omega}$ is

$$(1-\theta)^2\sigma_a^2 T\left(\frac{h-2}{(h-1)^2} + \frac{T-h}{(T-h+1)^2}\right) \to 4(1-\theta)^2\sigma_a^2.$$

   Therefore, for case (a): $\theta = 0.9$, $\sqrt{T}\hat{\omega} \sim N(\omega, 4(1-0.9)^2\sigma_a^2)$. For case (b): $\theta = 0$, $\sqrt{T}\hat{\omega} \sim N(\omega, 4\sigma_a^2)$. From the results, θ plays an important role in determining the limiting distribution of $\sqrt{T}\hat{\omega}$.

5. Consider the data $\{r_1, r_2, \ldots, r_n\}$ from the model $r_t = \mu + e_t$, where $e_t = \sigma_t a_t$ with a_t being a Gaussian white noise with variance σ_a^2 and $\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2$. Write down the conditional likelihood function of the data given r_1.

   Answer: Let F_{t-1} be the σ-field generated by $r_{t-1}, r_{t-2}, \ldots, r_1$. The conditional likelihood function is

$$\begin{aligned}f(r_1,\ldots,r_n|r_1) &= \prod_{t=2}^{T} f(r_t|F_{t-1})\\ &= \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi}\sigma_t}\exp[-\frac{(r_t-\mu)^2}{2\sigma_t^2}].\end{aligned}$$

6. Consider two time series

$$\begin{aligned}X_t &= 0.4X_{t-1} + a_t,\\ Y_t &= 1.2Y_{t-1} - 0.32Y_{t-2} + b_t\end{aligned}$$

   where $\{a_t\}$ and $\{b_t\}$ are two independent white noises with unit variance. What is the model ARMA model of $Z_t = X_t + Y_t$?

Answer: Applying $(1-0.8B)(1-0.4B)$ to Z_t, we have

$$\begin{aligned}(1-0.8B)(1-0.4B)Z_t &= (1-0.8B)a_t + b_t \\ &= (1-\theta B)e_t,\end{aligned}$$

where θ and σ_e^2, the variance of e_t, are determined by $1.64\sigma_a^2 + \sigma_b^2 = (1+\theta^2)\sigma_e^2, \quad -0.8\sigma_a^2 = -\theta\sigma_e^2$.

7. Consider an AR(1) model $z_t = \phi z_{t-1} + a_t$, where a_t is a white noise with variance σ_a^2. Suppose that among the 200 observations $\{z_1, \ldots, z_{200}\}$, z_{100} and z_{101} are missing. Derive the conditional distribution of (z_{100}, z_{101}) given the model and the data.

   Answer: From the AR(1) model, the two missing values are related to $z_{99}, z_{100}, z_{101}, z_{102}$. More specifically, we can obtain

$$\begin{bmatrix} \phi z_{99} \\ 0 \\ z_{102} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \phi & -1 \\ 0 & \phi \end{bmatrix} \begin{bmatrix} z_{100} \\ z_{101} \end{bmatrix} + \begin{bmatrix} -a_{100} \\ a_{101} \\ a_{102} \end{bmatrix}.$$

   This is a multiple linear regression with 3 bservations and two unknown parameter. Express the regression as $\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{E}$, where $\boldsymbol{\beta} = (z_{100}, z_{101})'$. The estimate of $\boldsymbol{\beta}$ is $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{X}'\boldsymbol{Y})$, which is distributed as $N(\boldsymbol{\beta}, \sigma_a^2(\boldsymbol{X}'\boldsymbol{X})^{-1})$.

8. Consider an ARIMA(p,d,q) process z_t satisfying $\phi(B)(1-B)^d z_t = \theta(B)a_t$, where a_t is a Gaussian white noise series with mean zero and variance σ_a^2. Suppose that the observed data is $y_t = \frac{\theta(B)}{\phi(B)(1-B)^d}(a_t + \omega I_t^{(h)})$, where $I_t^{(h)}$ is the indicator for time index h. Given the model, derive an estimate of ω? What is the distribution of the estimate?

   Answer: ω is the size of an innovational outlier. It is estimated by the residual at time $t = h$. Thus, letting $e_t = [(1-B)^d\phi(B)/\theta(B)]y_t$, $\hat{\omega} = e_h$. Under the null hypothesis of no oultier, $\hat{\omega}$ is normal with mean zero and variance σ_a^2.

9. Consider the ARMA(1,1) model $X_t = 0.8X_{t-1} + a_t - 0.4a_{t-1}$, where a_t is a Gaussian white noise with mean zero and variance 1. Let $Z_t = X_{2t} + X_{2t-1}$ for $t \geq 1$. What is the order of the ARMA model for Z_t?

   Answer: Z_t is a 2-aggregate process. Applying $(1-0.8^2B)Z_t$, we see that Z_t follows an ARMA(1,1) model.

10. Let r_t be the daily log return of an asset with mean zero. Describe a method to test the conditional heteroscedasticity in r_t. What is the test statistic? What is the asymptotic distribution of the test?

    Answer: Consider the multiple linear regression $r_t^2 = \alpha_0 + \sum_{i=1}^k \alpha_i r_{t-i}^2 + \epsilon_t$. A test can be constructed by testing $H_o : \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$ versus $H_a : \alpha_i \neq 0$ for some i. The usual F-statistic can be used. The asymptotic distribution is chi-square with k degrees of freedom.

11. Consider the model

$$\begin{aligned} r_t &= 0.01 + e_t, \quad e_t = \sigma_t a_t, \quad a_t \sim_{iid} N(0,1) \\ \sigma_t^2 &= 0.01 + 0.85\sigma_{t-1}^2 + 0.1 e_{t-1}^2. \end{aligned}$$

Suppose that $r_{100} = 0.1$ and $\sigma_{100}^2 = 0.09$. Compute the 1-step ahead forecast of r_t at the forecast origin $t = 100$? What is the associated volatility forecast?

Answer: $r_{100}(1) = 0.01$. Also, $e_{100} = 0.1 - 0.01 = 0.09$. Thus, $\sigma_{101}^2 = 0.01 + 0.85(0.09) + 0.1(.09)^2 = 0.0873$. The volatility forecast is $\sqrt{0.0873} = 0.295$.

12. What is the impact when the serial correlations in the residuals of a linear regression model are overlooked?

Answer: Incorrect estimation of the variances of coefficient estimates. In other words, the t-ratios are not reliable.

13. Suppose that X_t follows the AR(1) model $X_t = 0.9X_{t-1} + a_t$, where a_t is a Gaussian white noise series with mean zero and variance σ_a^2. What is the model for the overdifferenced series $Z_t = (1-B)X_t$?

Answer: $Z_t = (1-B)X_t = (1-B)\frac{1}{1-0.9B}a_t$ so that $(1-0.9B)Z_t = (1-B)a_t$, which is a non-invertible ARMA(1,1) model.

14. Describe two sources by which a GARCH model can introduce excess kurtosis.

Answer: (a) The excess kurtosis of the innovation a_t and (b) the dynamic of the GARCH structure.

15. Let X_t be the daily price range, (High minus low), of a particular stock. Assume that $X_t/\psi_t = \epsilon_t$, where ϵ_t is an exponential distribution with mean 1 and $\psi_t = 0.1 + 0.9\psi_{t-1} + 0.05X_{t-1}$. What is the E($X_t$)? What is Var($X_t$)?

Answer: Since $E(X_t) = E(\psi_t \epsilon_t) = E(\psi_t)$ and using stationarity, we have $E(\psi_t) = 0.1 + 0.9E(\psi_{t-1}) + 0.05E(X_{t-1})$. Therefore, $E(X_t) = 0.1/(1-0.9-0.05) = 2$. Using the result of Lec15-08, we have $\text{Var}(X_t) = \frac{2^2(1-.9^2-2(.05)(.9))}{1-2(.05)^2-.9^2-2(.05)(.9)} = 4.21$.

Problem B. (20 pts) Suppose that the univeriate time series z_t follows the model

$$z_t = z_{t-1} + y_t, \quad y_t = (1-\theta_1 B - \theta_2 B^2)a_t, \quad t = 1, \ldots, T,$$

where a_t is a Gaussian white noise with mean 0 and variance σ_a^2 and the MA model is invertible. Derive the limiting distributions of the following statistics as $T \to \infty$:

1. (5 points) $T^{-2}\sum_{t=1}^{T} z_{t-1}^2$

Answer: Based on the Theorem of Lec11, the key quantity to consider is $\sigma^2 = \lim_{T\to\infty} E(T^{-1}S_T^2)$. Since $y_t = (1-\theta_1 B - \theta_2 B^2)a_t$, it is easy to calculate that (a) $\sigma_y^2 = (1+\theta_1^2+\theta_2^2)\sigma_a^2$ and (b) $\sigma^2 = \sigma_y^2(1+2\rho_1+2\rho_2) = \sigma(1-\theta_1-\theta_2)^2$. Therefore,

$$T^{-2}\sum_{t=1}^{T} z_{t-1}^2 \Rightarrow \sigma_a^2(1-\theta_1-\theta_2)^2 \int_0^1 W(r)^2 dr,$$

where $W(.)$ denotes the standard Brownion motion.

2. (5 points) $T^{-1}\sum_{t=1}^{T} z_{t-1}y_t$

   Answer: $T^{-1}\sum_{t=1}^{T} z_{t-1}y_t \Rightarrow \frac{\sigma_a^2(1-\theta_1-\theta_2)^2}{2}\left(W(1)^2 - \frac{1+\theta_1^2+\theta_2^2}{(1-\theta_1-\theta_2)^2}\right)$.

3. (5 points) $T(\hat{\pi}-1)$, where $\hat{\pi}$ is the least squares estimate of the AR(1) model $z_t = \pi z_{t-1} + e_t$.

   Answer: $T(\hat{\pi}-1) \Rightarrow \frac{(1/2)(W(1)^2-(1+\theta_1^2+\theta_2^2)/(1-\theta_1-\theta_2)^2)}{\int_0^1 W(r)^2 dr}$.

4. (5 points) The t-ratio for testing $H_o : \pi = 1$ vs $H_a : \pi < 1$.

   Answer: $t_\pi \Rightarrow \frac{\frac{|1-\theta_1-\theta_2|}{2\sqrt{1+\theta_1^2+\theta_2^2}}\left[W(1)^2 - \frac{1+\theta_1^2+\theta_2^2}{(1-\theta_1-\theta_2)^2}\right]}{\left[\int_0^1 W(r)^2 dr\right]^{1/2}}$.

Problem C. (15 pts) Consider a time series $z_t = T_t + S_t + e_t$, where the components follow the models

$$(1-B)^2 T_t = \epsilon_{1t}, \quad (1+B+B^2+B^3)S_t = \epsilon_{2t}, \quad (1-0.5B)e_t = \epsilon_{3t},$$

where ϵ_{it} are independent white noise series with mean zero and variance σ_i^2, $i = 1, 2, 3$. Answer the following questions:

1. (3 points) Write down a state-space model for T_t.

   Answer: The obsevration eq. is $T_t = [1, 0]S_t$, whereas the state-transition eq. is

$$\begin{bmatrix} T_t \\ T_{t-1} \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} T_{t-1} \\ T_{t-2} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \epsilon_{1t}.$$

2. (3 points) Write down a state-space model for S_t.

   Answer: The obsevration eq. is $S_t = [1, 0]X_t$, whereas the state-transition eq. is

$$\begin{bmatrix} S_t \\ S_{t-1} \\ S_{t-2} \end{bmatrix} = \begin{bmatrix} -1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} S_{t-1} \\ S_{t-2} \\ S_{t-3} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \epsilon_{2t}.$$

3. (3 points) Write down a state-space model for e_t.

   Answer: The observatiob eq. is $e_t = e_t$ and the state-transition eq. is $e_t = 0.5e_{t-1} + \epsilon_{3t}$.

4. (6 points) Write down a state-space model for z_t.

Answer: Let $\boldsymbol{S}_t = (T_t, T_{t-1}, S_t, S_{t-1}, S_{t-2}, e_t)'$ and $\boldsymbol{\epsilon}_t = (\epsilon_{1t}, \epsilon_{2t}, \epsilon_{3t})'$. We have the observation eq. $z_t = (1,0,1,0,0,1)\boldsymbol{S}_t$ and the state-transition eq. is $\boldsymbol{S}_t = \boldsymbol{F}\boldsymbol{S}_{t-1} + \boldsymbol{G}\boldsymbol{\epsilon}_t$, where

$$\boldsymbol{F} = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{G} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Problem D. (15 pts) Consider the daily log returns, in percentages, of the stock of Adobe Systems, Inc. from 1997 to 2006. Analysis of the series, including GARCH modelling, is given in the attached output. Use the output to answer the following questions.

1. (3 pts) Ignoring the conditional heteroscedasticity, we saw a significant Ljung-Box statistic for the log return series. Write down the fitted AR(1) model.

   Answer: $(1 + 0.051B)(r_t - 0.083) = a_t$, where $\sigma_a^2 = 12.04$.

2. (3 pts) Is there any ARCH effect in the residuals of the AR(1) model? Why?

   Answer: Yes, the Ljung-Box Q statistic of the squared residuals shows $Q(10) = 142.52$ with p-value close to 0.

3. (3 pts) Write down the fitted Gaussian GARCH(1,1) model.

   Answer: Mean equation is $r_t = 0.115 + e_t$, $e_t = \sigma_t a_t$ with a_t being iid N(0,1). The volatility equation is $\sigma_t^2 = 0.034 + 0.035e_{t-1}^2 + 0.963\sigma_{t-1}^2$.

4. (3 pts) Write down the fitted GARCH(1,1) model with Student-t innovations.

   Answer: Mean equation is $r_t = 0.041 + e_t$, $e_t = \sigma_t a_t$ with a_t being a standardized Student-t distirbution with 4.91 degrees of freedom. The volaitlity equation is $\sigma_t^2 = 0.008 + 0.018e_{t-1}^2 + 0.981\sigma_{t-1}^2$.

5. (3 pts) Compare the two fitted GARCH(1,1) model. Which one is preferred? Why?

   Answer: The GARCH(1,1) model with Student-t innovation is preferred because it has a high log likelihood value and allows for heavy tails.