Exam II: Information and Practice Problems

Statistics 509 – Winter 2022

April 13, 2022

Information: The Second Exam will be Friday, April 22 from 1:30-3:30 in room **SEB1202** (in the School of Education Building). The exam will cover the material in Lectures 7-12 covered in class and corresponding relevant material in the book. The exam is open textbook (only the textbook) and open notes. Below is a set of practice Exam Problems – solutions for the problems will be posted on Monday, April 18. Additional practice problems are the exercise problems from the lectures and exercises in the book.

Practice Exam Problems

1. Problem 8 on page 514 in Ruppert/Matteson, but change part (d) to be a 4% expected return, and add the following part **(e)**.

   (e) Determine the β's of asset C and asset D relative to the tangency portfolio.

2. Problem 7 on page 513 of Ruppert/Matteson.

3. Suppose have two stationary AR time series

$$X_n = .5X_{n-1} + \epsilon_n$$

$$Y_n = \epsilon'_n + .6\epsilon'_{n-1}$$

   where ϵ and ϵ' are both white-noise processes with mean 0 and standard deviation of 0.5, and which are independent of each other.

   (a) Suppose $X'_n = X_n + X_{n+1}$. Show that $X' = \{X'_n\}$ is a weakly stationary process and derive it's auto-covariance function.

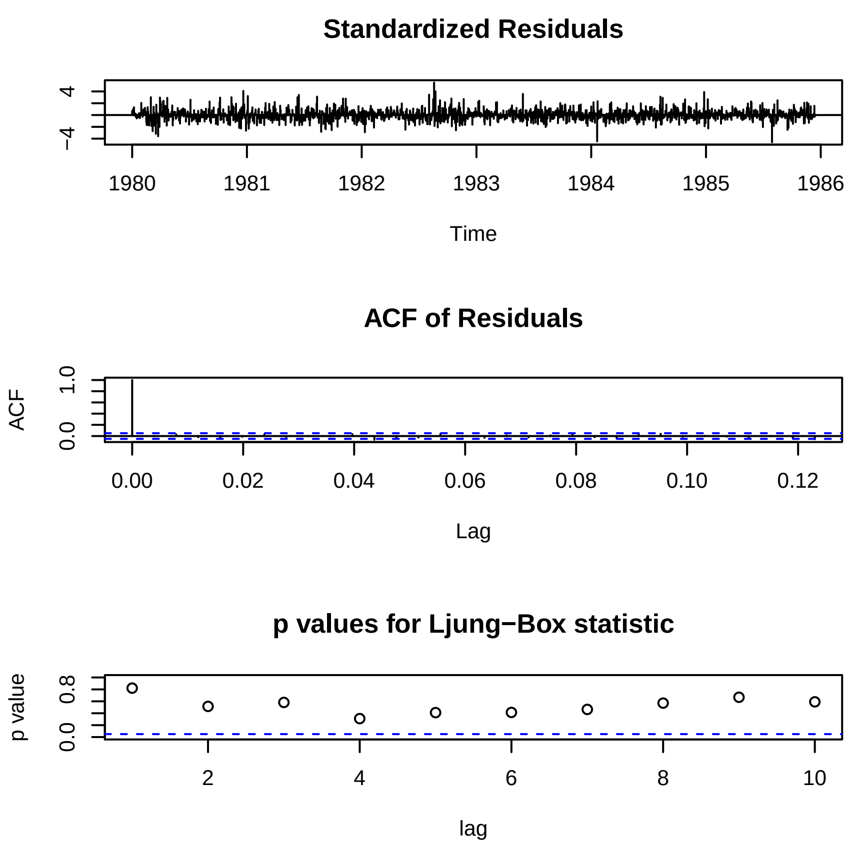
   (b) Suppose generate a new time series of $Z_n = X_n + Y_n$ – verify that this is a stationary process, and derive its mean and autocovariance function.

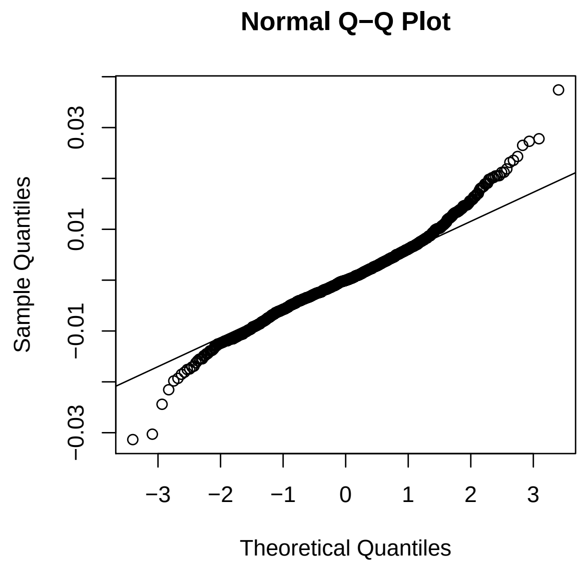
   (c) With Z as in **(b)**, find the optimal linear predictor of X_{25} based on Z_{24} and determine the mean-squared prediction error of this linear predictor.

4. Suppose had the following results a time series analysis in R applied to log of the daily price data from Dow Jones Utility Average Index.

```
> X = read.csv("Data//DJUtilityAv-daily.csv",header=TRUE)
> X_logprice = rev(log(X$Adj.Close))
> X_logprice.ts = ts(data=X_logprice,start=c(1980,1),frequency=252,
+ names=c("DJUtilityAv_DailyLogPrice"))
> auto.arima(X_logprice.ts,max.p=4,max.q=4,ic="bic")
Series: X_logprice.ts
ARIMA(1,1,0)
```

```
Coefficients:
         ar1
      0.2009
s.e.  0.0253

sigma^2 estimated as 4.633e-05:  log likelihood=5352.84
AIC=-10701.67   AICc=-10701.67   BIC=-10691.05
> DJUtility_ARIMA = arima(X_logprice.ts, order = c(1,1,0), method = "ML")
> tsdiag(DJUtility_ARIMA)
> qqnorm(DJUtility_ARIMA$residuals)
> qqline(DJUtility_ARIMA$residuals)
```





(a) What model was selected by auto.arima, and what criteria was used for this selection?

(b) From these results, what are the estimated coefficients and the estimated standard deviation for the white noise process?

(c) Based on the diagnostic plots, what can you say about the fit of the this ARIMA model to the daily log-price data, i.e., what is good and what is not?

(d) What alternative models/approaches might you consider based on these results?

5. Suppose have time series $X = \{X_n\}_{n=0}^{\infty}$ following a first-order AR model

$$X_n = \alpha X_{n-1} + \epsilon_n \qquad\qquad n = 1, 2, \ldots$$

where assume that $|\alpha| \leq 1$ and $\{\epsilon_n\}$ is a white noise process with mean zero and variance of σ^2. The least-squares estimates of α is given by

$$\hat{\alpha} = \frac{\sum_{i=1}^{n} X_i X_{i-1}}{\sum_{i=1}^{n} X_i^2}.$$

(a) Assuming $|\alpha| < 1$, ergodicity, and $\{X_n\}$ is a stationary process, derive the limits of the normalized numerator and denominator of $\hat{\alpha}$, i.e.,

$$\frac{1}{n}\sum_{i=1}^{n} X_i X_{i-1} \quad \text{and} \quad \frac{1}{n}\sum_{i=1}^{n} X_i^2.$$

Justify your answer.

(b) Assuming $X_0 = 0$ and $\alpha = 1$, derive expressions for the expected value of the normalized numerator and denominator of $\hat{\alpha}$. *Hint:* Can use that $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$.

(c) Based on results in **(a)** and **(b)**, would you say that $\hat{\alpha}$ is a reasonable estimate of α for all cases. Explain your answer.

(d) Based on results in **(a)** and **(b)** how might one test the hypothesis of $H_o : \alpha = 1$ vs. $H_1 : \alpha < 1$, i.e., what would be the general form of the test based on the test statistic $\hat{\alpha}$.

6. **(a)** Problem 12 on page 357 in Ruppert/Matteson.

   (b) Problem 16 on page 359 in Ruppert/Matteson, and add another part which is to compute the mean-squared prediction error of the 3 predictions, assuming perfectly accurate estimates on the coefficients.

7. Suppose that X_n is time series which corresponds to an ARCH(2) model of the form

$$X_n = \sigma_n \epsilon_n$$

where

$$\sigma_n^2 = 1 + .4 \cdot X_{n-1}^2 + .2 \cdot X_{n-2}^2$$

(a) Derive the conditional distribution of X_3 given $X_1 = 1, X_2 = 2$ assuming $\{\epsilon_n\}$ are iid $\mathcal{N}(0, 1)$.

(b) Derive the conditional distribution of X_3 given $X_1 = 1, X_2 = 2$ assuming $\{\epsilon_n\}$ are iid DExp$(0, 1)$.

(c) Answer **(a)** and **(b)** if the equations are

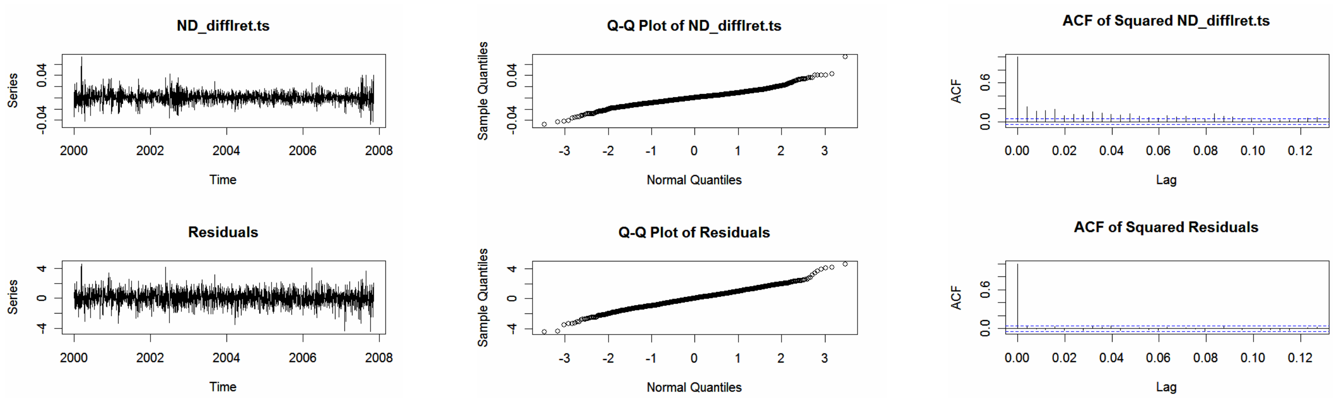
$$X_n = 0.5 + \sigma_n \epsilon_n$$

where

$$\sigma_n^2 = 1 + .4 \cdot (X_{n-1} - 0.5)^2 + .2 \cdot (X_{n-2} - 0.5)^2$$

8. Suppose have the following results for a GARCH estimation on differenced log-returns of the Daily Adjusted Closing Prices of the NASDAQ from Jan-2000 to Nov-2007 – note we are assuming a mean of 0 for the process.

```
> summary(ND_garchest11)

Call:
garch(x = ND_difflret.ts, order = c(1, 1), itmax = 200, grad = c("analytic"))

Model:
GARCH(1,1)

Coefficient(s):
    Estimate  Std. Error  t value Pr(>|t|)
a0 2.743e-06   5.323e-07    5.153 2.57e-07 ***
a1 8.888e-02   9.438e-03    9.418  < 2e-16 ***
b1 8.848e-01   1.206e-02   73.375  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Diagnostic Tests:
        Jarque Bera Test

data:  Residuals
X-squared = 75.8134, df = 2, p-value < 2.2e-16

        Box-Ljung test

data:  Squared.Residuals
X-squared = 2.1571, df = 1, p-value = 0.1419
```



(a) Based only on the plot of the differenced log-returns, does it appear that GARCH(1,1) model is better than a stationary white-noise for modeling this time series of the differenced log-returns. Explain your answer.

(b) Interpret the results of the GARCH analysis and say whether the estimates are appropriate/valid. Justify your answer and explain what are the deficiencies in this analysis.

(c) This data goes through the end of November 2007 – based on the estimated model being valid and assuming that the difference in log returns on the last trading day in November was $x_n = .0140$ and the standard deviation was $\hat{\sigma}_n = 0.0214$, estimate the standard deviation of the difference of log-returns on the first trading day in December (σ_{n+1}) and also estimate the square of differenced log-return (X^2_{n+1}).

9. Suppose have a weakly stationary process

$$X_n = \epsilon_n + .5\epsilon_{n-1} + .5\epsilon_{n-2} \qquad\qquad n = 1, 2, \ldots$$

where $\{\epsilon_n\}$ is a white noise process with variance σ^2.

(a) Derive the mean and autocovariance of X_n - you may derive it from first principles or rigorously quote results from class.

(b) Suppose $Y_n = X_n + X_{n-1}$. Is this time series $\{Y_n\}$ ARMA(p, q)? Justify your answer and specify parameters p and q if your answer is yes.

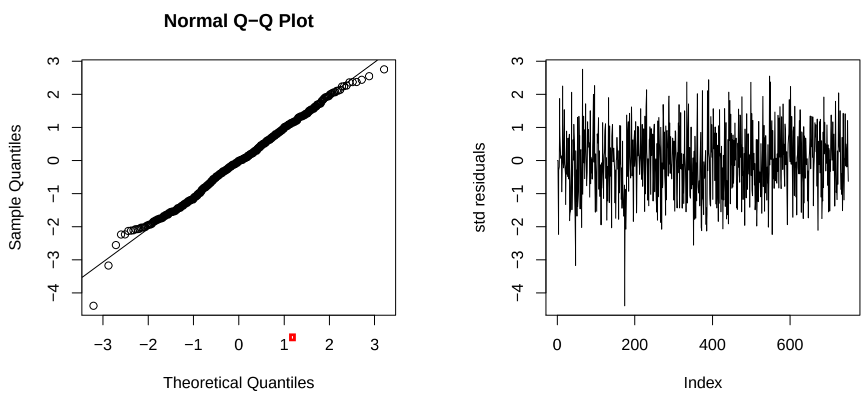
(c) What is your best linear prediction of X_{n+2} based on knowing only X_n? What is the the mean-squared prediction error of this linear predictor?

10. Suppose have daily price data for SP Midcap 400 index (for 2001-2003) and the following is the R-output of an analysis (commands and plots).

```
> X = read.csv("Data//SPMidcap400.csv",header=TRUE)
> X_logret = diff(log(X$AdjClose))
> X_logret.ts = ts(data=X_logret,start=c(2001,1),frequency=252,names=c("SPMidcap400"))
```

```
>
> SP400_garch <- garchFit(~arma(2,0)+garch(1,1), data=X_logret.ts[1:750], cond.dist=c("snorm
> summary(SP400_garch)

Title:
 GARCH Modelling

Call:
 garchFit(formula = ~arma(2, 0) + garch(1, 1), data = X_logret.ts[1:750],
    cond.dist = c("snorm"), include.mean = T, include.skew = F,
    trace = F)

Mean and Variance Equation:
 data ~ arma(2, 0) + garch(1, 1)
<environment: 0x0000000004a0c358>
 [data = X_logret.ts[1:750]]

Conditional Distribution:  snorm

Std. Errors:  based on Hessian

Error Analysis:
         Estimate  Std. Error  t value Pr(>|t|)
mu      6.879e-04   4.299e-04    1.600   0.1096
ar1     2.861e-02   3.788e-02    0.755   0.4501
ar2    -8.086e-02   3.757e-02   -2.152   0.0314 *
omega   5.324e-06   2.651e-06    2.008   0.0446 *
alpha1  8.996e-02   2.206e-02    4.078 4.53e-05 ***
beta1   8.792e-01   3.024e-02   29.071  < 2e-16 ***
---

Standardised Residuals Tests:
                                Statistic p-Value
 Jarque-Bera Test   R    Chi^2  0.6848627 0.7100419
 Shapiro-Wilk Test  R    W      0.9953828 0.02412903
 Ljung-Box Test     R    Q(10)  10.54462  0.3940778
 Ljung-Box Test     R    Q(15)  17.11203  0.3122116
 Ljung-Box Test     R    Q(20)  20.2074   0.445025
 Ljung-Box Test     R^2  Q(10)  13.88344  0.1783752
 Ljung-Box Test     R^2  Q(15)  14.37448  0.4973381
 Ljung-Box Test     R^2  Q(20)  18.14707  0.5777204
 LM Arch Test       R    TR^2   16.49336  0.1696692

Information Criterion Statistics:
      AIC       BIC       SIC      HQIC
-5.880602 -5.843641 -5.880729 -5.866360

> qqnorm(SP400_garch@residuals/SP400_garch@sigma.t)
> qqline(SP400_garch@residuals/SP400_garch@sigma.t)
```

```
> plot(SP400_garch@residuals/SP400_garch@sigma.t,type='l',ylab='std residuals')
```



(a) What model is being fit to this time series, and state clearly the final mathematical form of the model as estimated?

(b) Based on these estimates, what is the model for the conditional variance of the the white noise process?

(c) Based on diagnostics as shown here, how satisfied are you with this model? What alternative models might you investigate to get an even better fit? How would you select between this model and your proposed alternatives?

11. **(a)** Suppose X and σ^2 are two random variables and that X conditional on σ^2 is $\mathcal{N}(0, \sigma^2)$, i.e. that conditioned on the value of σ^2, X is mean-zero normal with variance σ^2. Show that

$$\frac{E(X^4)}{[\text{Var}(X)]^2} = 3\left[1 + \frac{\text{Var}(\sigma^2)}{[E(\sigma^2)]^2}\right].$$

What does this say about the excess kurtosis of X relative to that of a normal random variable?

Hint: For the above problem, recall that from your basic probability/statistics class the results that

$$E(X) = E(E(X|\sigma)),\ E(X^2) = E(E(X^2|\sigma)),\ E(X^4) = E(E(X^4|\sigma)).$$

(b) Suppose that X_n is a GARCH(1,1) process with parameters of α_o, α_1 and β_1, and the innovations $\{\epsilon_n\}$ are iid $\mathcal{N}(0,1)$. Show that the kurtosis of X_n (assuming stationarity) is given by

$$\kappa = \frac{E(X_n^4)}{[\text{Var}(X_n)]^2} = \frac{3\left[1 - (\alpha_1 + \beta_1)^2\right]}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} > 3$$

when $(\alpha_1 + \beta_1)^2 + 2\alpha_1^2 < 1$.

(c) What are the implications for the GARCH(1,1) process if $(\alpha_1 + \beta_1)^2 + 2\alpha_1^2 \geq 1$?

12. Exercise 6 on page 448 in Ruppert/Matteson.

13. Exercise 5 on page 448 in Ruppert/Matteson.

14. Suppose $Y_{1,t}, Y_{2,t}$ are log-prices satisfying the following model for log-returns,

$$\begin{aligned} \triangle Y_{1,t} \equiv Y_{1,t} - Y_{1,t-1} &= \phi_1(Y_{1,t-1} - \lambda Y_{2,t-1}) + \epsilon_{1,t} \\ \triangle Y_{2,t} \equiv Y_{2,t} - Y_{2,t-1} &= \phi_2(Y_{1,t-1} - \lambda Y_{2,t-1}) + \epsilon_{2,t} \end{aligned}$$

where $\epsilon_{1,t}, \epsilon_{2,t}$ are white noise processes.

(a) Show that the above implies that $Y_{1,t} - \lambda Y_{2,t}$ satisfies an AR(1) model and identify the coefficients (subtract λ times second equation from the first equation).

(b) Specify sufficient conditions on ϕ_1, ϕ_2, λ implying that the AR(1) process is stationary.

(c) Problem 2 on page 463 in Ruppert/Matteson – basically adding μ_1 and μ_2 to the righthand side of the two equations above.