Queen Mary
University of London

Main Examination period 2020 – May/June – Semester B
Online Alternative Assessments

MTH6139 / MTH6139P: Time Series

You should attempt ALL questions. Marks available are shown next to the questions.

In completing this assessment, you may use books, notes, and the internet. You may use calculators and computers, but you should show your working for any calculations you do. You must not seek or obtain help from anyone else.

At the start of your work, please **copy out and sign** the following declaration:

> I declare that my submission is entirely my own, and I have not sought or obtained help from anyone else.

All work should be **handwritten**.

You have **24 hours** in which to complete and submit this assessment. When you have finished your work, you should upload photographs or scans of your work using the upload tool on the QMplus page for the module. You should also email a copy of your work to maths@qmul.ac.uk with your student number and the module code in the subject line.

One of the photographs you upload should feature the first page of your work together with either yourself or your student ID card.

You are not expected to spend a long time working on this assessment. We expect you to spend about **2 hours** to complete the assessment, plus the time taken to scan and upload your work. Please try to upload your work well before the end of the assessment period, in case you experience computer problems.

IFoA exemptions
This module counts towards IFoA actuarial exemptions. For your submission to be eligible for IFoA exemptions, you must submit within the first **3 hours** of the assessment period. You may then submit a second version later in the assessment period if you wish, which will count only towards your degree.

Examiners: W. W. Yoo, D. S. Coad

Question 1 [23 marks].
Let $\{Y_t\}$ be a zero-mean weakly stationary process with autocovariance function $\gamma(h)$. Denote by s_t the seasonal component with period 4 such that $s_t = s_{t-4}$, and let a and b be some constants such that $a \neq 0, b \neq 0$.

(a) Suppose our time series is $X_t = a + bt + s_t + Y_t$.

(i) Define the operator ∇_4. **[2]**

(ii) Explain what is the main use of ∇_4. **[2]**

(iii) Show that $\nabla_4 X_t$ is a weakly stationary process and express its autocovariance function in terms of $\gamma(h)$. **[9]**

(b) Suppose now $X_t = (a + bt)s_t + Y_t$.
Here you can use the fact that addition or subtraction of weakly stationary processes is still weakly stationary.

(i) Show that $\nabla_4 X_t$ is not weakly stationary in this case. **[4]**

(ii) Find a differencing operator that will reduce X_t to a weakly stationary process. **[6]**

Question 2 [17 marks].

(a) Suppose we are given the linear filter

$$a_{-2} = -\frac{3}{35}, a_{-1} = \frac{12}{35}, a_0 = \frac{17}{35}, a_1 = \frac{12}{35}, a_2 = -\frac{3}{35}$$

with $a_j = 0$ for $j > 2$ and $j < -2$. Would the cubic trend $m_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3$ pass through this linear filter without distortion? Justify your answer. **[9]**

(b) Exponential smoothing is defined by the following recursion:

$$\widehat{m}_1 = X_1,$$
$$\widehat{m}_t = \alpha X_t + (1 - \alpha)\widehat{m}_{t-1}, \quad t = 2, \ldots, n.$$

Suppose we have time series data $\{8, 5, 2, 4, 9, 12, 10\}$ with $X_1 = 8, X_2 = 5$ and so on until $X_7 = 10$. Assume that $\alpha = 0.7$. Calculate $\widehat{m}_5$ using exponential smoothing. **[8]**

Question 3 [12 marks]. In the following, let $\{Z_t\}$ be a white noise process with mean 0 and variance 1. Describe the behavior of the autocorrelation function (ACF) and partial autocorrelation function (PACF) for the following processes.

(a) $X_t = Z_t + \sum_{j=1}^{3} (0.2)^j Z_{t-j}$ **[3]**

(b) $X_t = -0.3X_{t-1} + 0.1X_{t-2} + Z_t$ **[3]**

(c) $X_t = 0.5X_{t-1} + Z_t + 0.3Z_{t-1}$ **[3]**

(d) $X_t = 0.4X_{t-1} - 0.4Z_{t-1} + Z_t$. **[3]**

Question 4 [28 marks].

(a) Consider the series

$$X_t = \sin(2\pi U t),$$

$t = 1, 2, \ldots$, where U has a uniform distribution on the interval $(0, 1)$. Prove that X_t is weakly stationary.
(Hint: Use the trigonometric identities $\cos(A - B) = \cos(A)\cos(B) + \sin(A)\sin(B)$ and $\cos(A + B) = \cos(A)\cos(B) - \sin(A)\sin(B)$ with the facts $\cos(2\pi k) = 1$, $\sin(2\pi k) = 0$ for $k = 0, 1, 2, \ldots$.) **[9]**

(b) Consider the AR(2) process

$$X_t = 0.45X_{t-1} - 0.05X_{t-2} + Z_t,$$

where $\{Z_t\}$ is a white noise process with mean 0 and variance 1.

(i) Check whether this process is causal and invertible. Explain your answers. **[5]**

(ii) Its autocorrelation function obeys the linear difference equation

$$\rho(h) - 0.45\rho(h-1) + 0.05\rho(h-2) = 0, \quad h \geq 2.$$

Find the initial values $\rho(0)$ and $\rho(1)$ by solving

$$\rho(h) - 0.45\rho(h-1) + 0.05\rho(h-2) = 0, \quad h = 1.$$ **[5]**

(iii) The general solution of the linear difference equation in (ii) is of the form

$$\rho(h) = k_1 z_1^{-h} + k_2 z_2^{-h},$$

where z_1 and z_2 are the roots of the characteristic polynomial from the AR(2) process above, and k_1 and k_2 are constants. Find $\rho(h)$ by computing the values for these quantities. **[9]**

Question 5 [20 marks]. Figure 1 below shows the atmospheric CO_2 concentration in parts per million (ppm) at Mauna Loa. The data are monthly from 1959 to 1997.

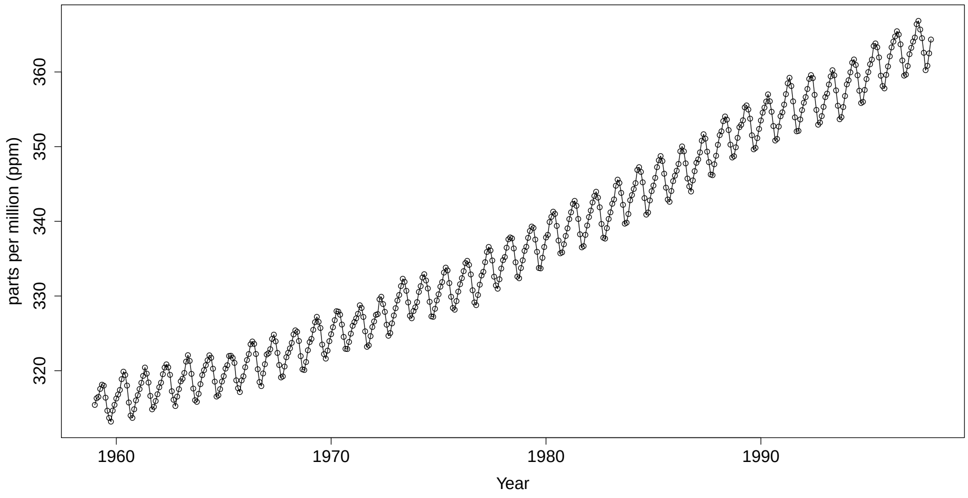

Figure 1: CO_2 concentration in ppm.

(a) Describe the main features of the time series shown by the plot in Figure 1. **[4]**

First-order differencing and then lag-12 differencing were performed on the CO_2 data. Figure 2 shows the sample ACF and PACF plots after applying both differencing operators, i.e, $\nabla_{12}\nabla$ on the CO_2 data.

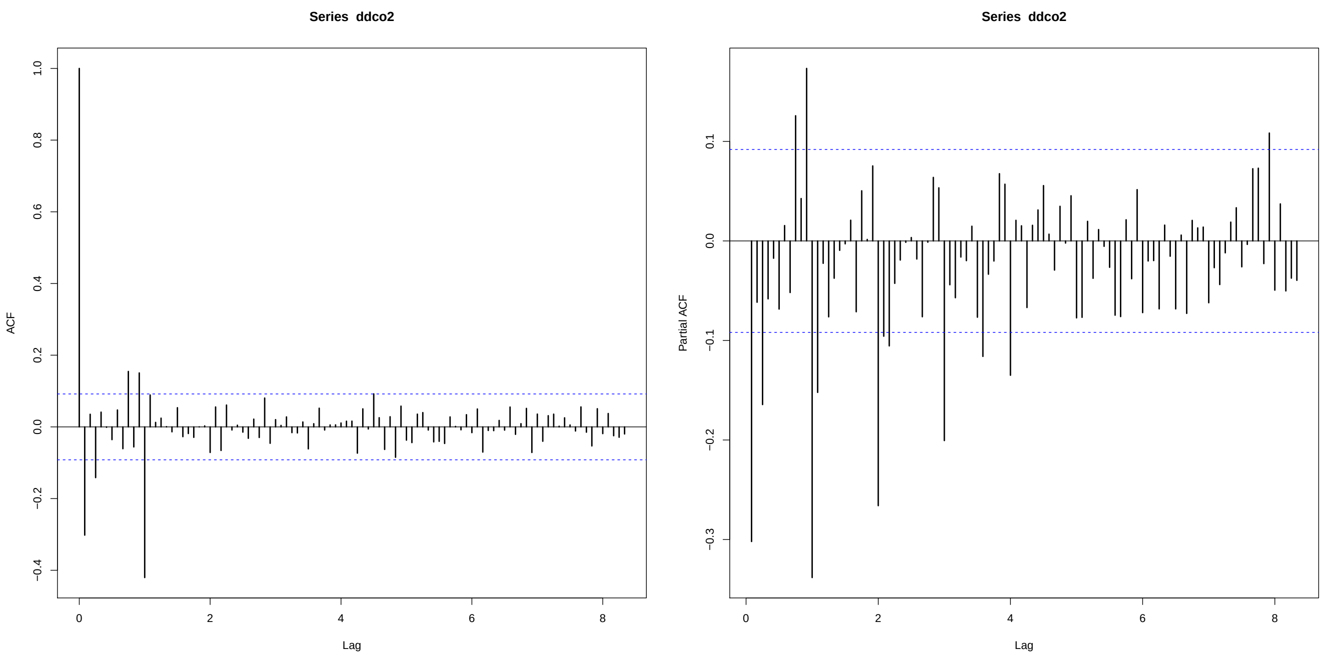

Figure 2: Left: ACF of $\nabla_{12}\nabla$ on CO_2 data, Right: PACF of $\nabla_{12}\nabla$ on CO_2 data.

The R commands to produce these plots are given below.

```
#name of data is co2
dco2 = diff(co2) #1st order differencing
ddco2 = diff(dco2, 12) #lag-12 differencing

par(mfrow=c(1,2))
acf(ddco2, lag.max=100) #plot sample acf
pacf(ddco2, lag.max=100) #plot sample pacf
```

(b) State two possible SARIMA models indicated by the sample ACF and PACF plots in Figure 2. Explain how you arrived at this conclusion. **[8]**

(c) Suppose that an ARIMA$(0, 1, 0) \times (0, 1, 2)_{12}$ model is fitted to the CO_2 data. Figure 3 below shows the resulting model diagnostics of the residuals and running the aggregated Box-Ljung test statistics on these residuals yielded

```
        Box-Ljung test

data:  resid(m1$fit)
X-squared = 49.46, df = 18, p-value = 9.104e-05
```

Would you recommend this model to scientists working to understand the time dependence of these data? If no, give suggestions on how to improve this model. Explain your reasoning. **[8]**

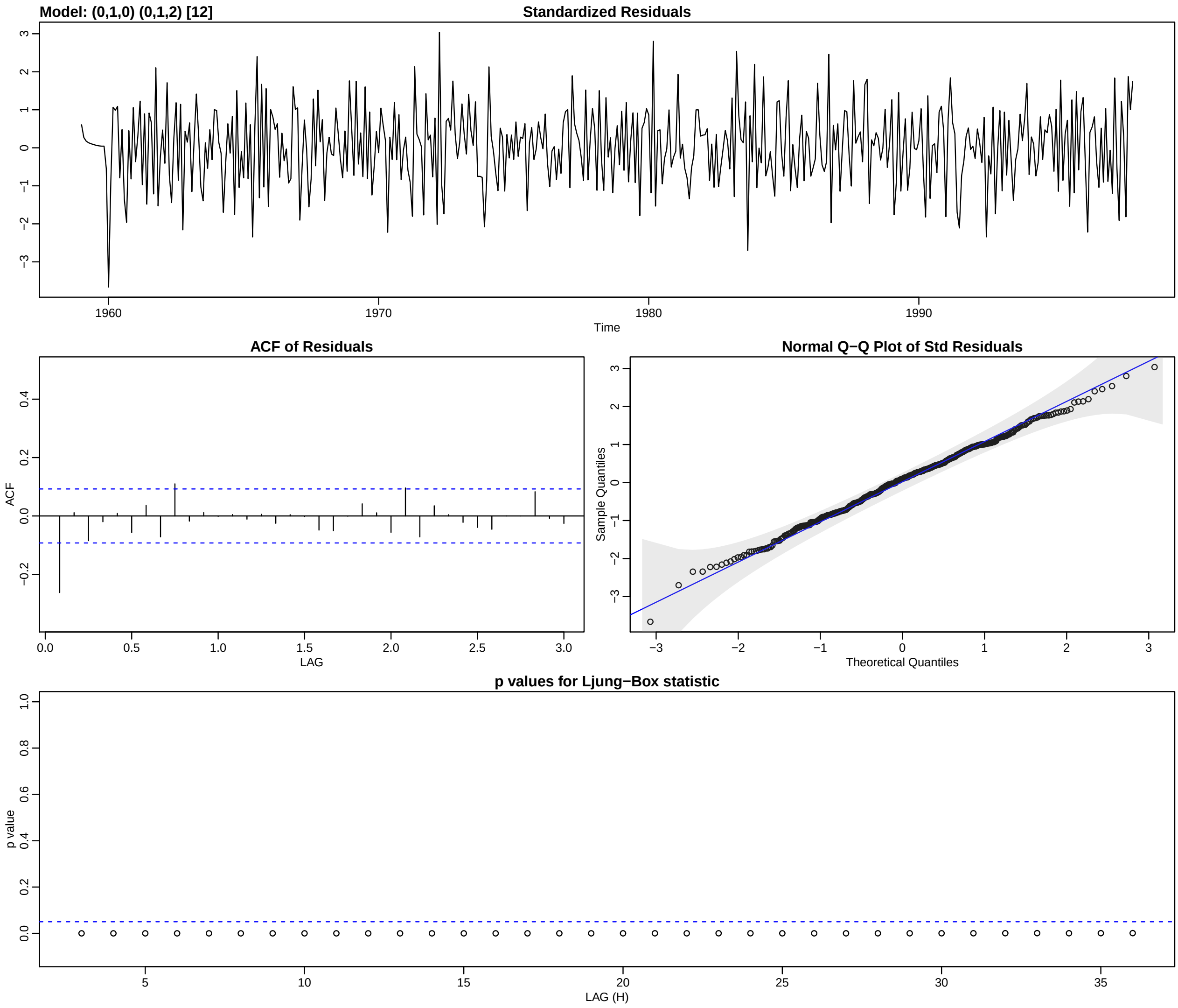

Figure 3: Model diagnostics for fitting ARIMA$(0, 1, 0) \times (0, 1, 2)_{12}$ to the CO_2 data.

End of Paper.