

 W1-2-60-1-6

**JOMO KENYATTA UNIVERSITY OF AGRICULTURE AND TECHNOLOGY**

**UNIVERSITY EXAMINATIONS 2013/2014**

**YEAR I SEMESTER I EXAMINATION FOR THE DEGREE OF MASTERS IN APPLIED STATISTICS**

**STA 3106: TIME SERIES AND ANALYSIS**

**DATE: DECEMBER 2013 TIME: 3 HOURS**

**INSTRUCTIONS:**

**-ANSWER QUESTION ONE AND ANY OTHER TWO QUESTIONS**

**- PLEASE START EACH QUESTION ON A NEW PAGE CLEARLY INDICATING**

 **THE QUESTION NUMBER**

**- THE TOTAL MARKS IS 60 MARKS**

**QUESTION ONE: COMPULSORY (20 MARKS)**

(i) What is the main difference between time series and multiple regression (analysis)

 model? [2 marks]

(ii) Mention THREE methods of replacing missing values when they are present in your time

 series date. [3 marks]

(iii) List and discuss THREE methods of determining trend in time series. [6 marks]

(iv) What are the main tool to obtain stationality in mean and variance for given time series,

 and which method of transformation should be done first? [3 marks]

(v) Define briefly the FOUR components in time series. [6 marks]

**QUESTION TWO (20 MARKS)**

Provide answer for the following questions:

(i) For non-stationary time series data, write step by step the procedures which lead you to

 fitting adequate Box-Jenkins time series model. [8 marks]

(ii) Mention the methods step by step for checking the fitted if Box-Jenkins time series

 model is adequate or not. [6 marks]

(iii) Justify briefly the difference between ARIMA and SARIMA models. [3 marks]

(iv) Describe some known non-linear type of time series model with their characteristics. [3 marks]

**QUESTION THREE (20 MARKS)**

The following table summarizes the number of tuberculosis cases registered through year;

|  |  |
| --- | --- |
| **Year**  | **Tuberculosis Cases** |
| 2003 | 2 |
| 2004 | 4 |
| 2005 | 2 |
| 2006 | 4 |
| 2007 | 3 |
| 2008 | 5 |
| 2009 | 3 |
| 2010 | 5 |
| 2011 | 4 |
| 2012 | 4 |

For this data, from the preliminary analysis the following output are obtained:

1. The data is approximately stationed with mean and variance in time-series sense
2. The autocorrelation coefficient at lay 2 is highly significant compared to others, with value 0.662 and Box-Lying p-value = 0/014
3. None of partial autocorrelation coefficients is out of the confidence limits.

Based on the above given information:

1. Select appropriate Box-Jenkins Time series model and justify your answer. [5 marks]
2. Write the fitted equation for the model using ordinary least square method, show each step clearly that you used to fit this model. [15 marks]

**QUESTION FOUR (20 MARKS)**

The following are SPSS analysis output related to correlogram for annual yield of grain time series data (1852 – 1925). From these output, (See next pate)

1. How many yearly observations were included in the analysis? [2 marks]
2. Write the general Null and Alternative hypothesis to be tested by the Box-Lying statistics. [2 marks]
3. Write the Null and alternative hypothesis for lay 16 related to the population autocorrelation coefficient and make your conclusion and decision on it at 5% level of significance. [6 marks]
4. According to the illustrated correlogram, is the transformed time series stationary in mean and variance? Why? [2 marks]
5. Interpret the value of autocorrelation coefficient at log 1 [4 marks]
6. By considering only the most significance autocorrelation coefficient and partial autocorrelation coefficient, what type of Box-Jenkins Time series model is appropriate for this transformed data. [2 marks]
7. Write the general equation of your chosen model for the above questions. (vi) [2 marks]