UNIVERSITY OF KELANIYA – SRI LANKA

CENTRE FOR DISTANCE AND CONTINUING EDUCATION

Bachelor of Science (General) Degree (External) Examination – May 2025

ACADEMIC YEAR 2019 – Semester I

STATISTICS

STAT 36513 – Time Series Analysis (R)

No. of questions: Four (04) **No. of pages: Six (06)** **Duration: Two and Half (2 1/2) hrs.**

Answer all questions

01.

*(a)*Define what is a Time Series.

*(b)*What is meant by a non-stationary time series?

*(c)*Explain one statistical test that can be used to assess the stationarity of a time series.

*(d)*State four main components of a time series.

*(e)*Identify component/s of below two time series plots with suitable justification.

Figure 1: Quarterly Price

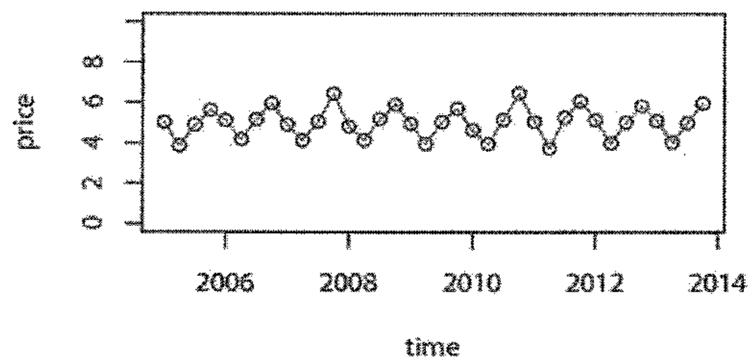
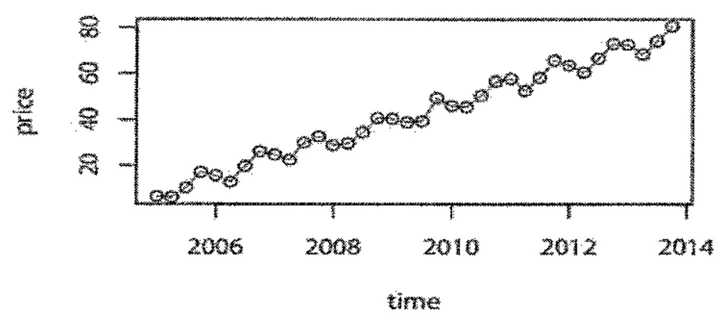


Figure 2: Quarterly Price



02.

(a) Briefly explain the following techniques:

(i) Differencing

(ii) Log Transformation

(b) Sunshine Electronics, a medium-sized retailer, has been monitoring its monthly sales revenue (in LKR '000) for one full year. The management is keen to analyze the sales trend and identify any underlying patterns or seasonality to improve inventory planning and budget forecasts for the upcoming year.

The recorded sales revenue for each month in 2024 is shown below:

Sales (LKR "Thousands")

Month	Sales
January	120
February	130
March	145
April	155
May	160
June	170
July	190
August	200
September	220
October	230
November	240
December	260

As a data analyst for the company, you are asked to analyze this time series and provide insights into whether the data is stationary, and if not, what steps are needed to make it so.

(i). Plot the time series and briefly describe it.

(ii). Apply the first-order differencing to the series. Present the differenced data and comment on whether the differenced series appears stationary.

(iii). If the first difference is still non-stationary. Explain other transformations might be appropriate before differencing again? Explain the reasoning.

03.

(a) List three key assumptions of time series models that are commonly used as measures of model adequacy. For each assumption, mention one statistical test or diagnostic method used to assess it.

(b) Define a general ARMA (p, q) process.

(c) A time series $\{Y_t\}$ be defined by:

$$Y_t = 0.5y_{t-1} + \varepsilon_t + 0.4\varepsilon_{t-1}$$

where $\{\varepsilon_t\}$ is a white noise with mean zero and variance σ^2.

(i) Identify the time series model given by the above equation.

(ii) Is the above process stationary? Justify your answer.

*(d)*Consider the following process:

$$X_t = \varepsilon_t - 0.6\varepsilon_{t-1} + 0.2\varepsilon_{t-2}$$

where $\{\varepsilon_t\}$ is a white noise with mean zero and variance σ^2.

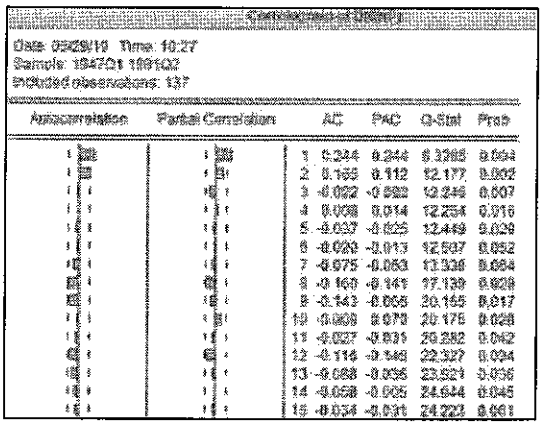
(i) State name of the process.

(ii) Compute the mean and the autocovariance function γ_h for $h = 0,1,2,3$.

04.

(a) ACF and PACF plots of different time series are given below. Write down all candidate time series models for each series using given plots.

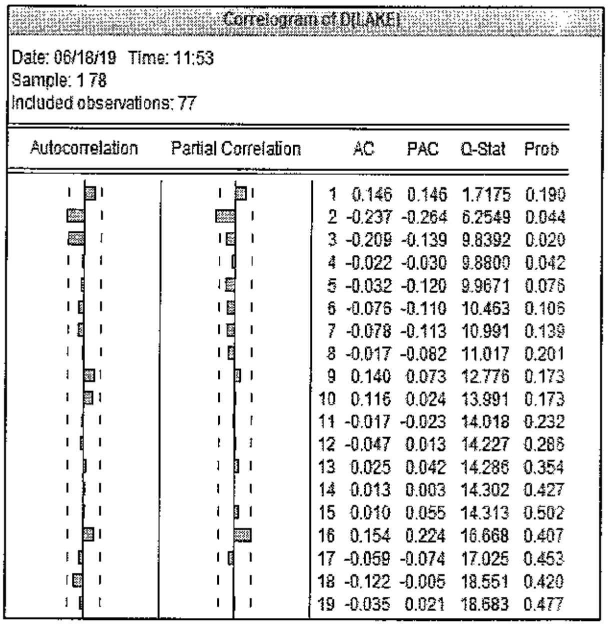
(i) Correlogram of the first differenced series of Quarterly GDP



Date: 02/28/19 Time: 10:27
Sample: 1947Q1 1981Q2
Included observations: 137

	AC	PAC	Q-Stat	Prob
1	0.244	0.244	8.3365	0.004
2	0.165	0.112	12.177	0.002
3	-0.022	-0.092	12.246	0.007
4	0.008	0.014	12.254	0.016
5	-0.037	-0.025	12.449	0.029
6	-0.020	-0.013	12.507	0.052
7	-0.075	-0.053	13.336	0.064
8	-0.160	-0.141	17.139	0.029
9	-0.143	-0.065	20.155	0.017
10	-0.008	0.073	20.175	0.028
11	-0.027	-0.031	20.282	0.042
12	-0.114	-0.146	22.327	0.034
13	-0.088	-0.036	23.521	0.036
14	-0.058	-0.005	24.044	0.045
15	-0.034	-0.031	24.223	0.061

(ii) Correlogram of first differenced series of Monthly Water Pollution Index



Correlogram of D(LAKE)

Date: 06/18/19 Time: 11:53
Sample: 1 78
Included observations: 77

	AC	PAC	Q-Stat	Prob
1	0.146	0.146	1.7175	0.190
2	-0.237	-0.264	6.2549	0.044
3	-0.209	-0.139	9.8392	0.020
4	-0.022	-0.030	9.8800	0.042
5	-0.032	-0.120	9.9671	0.076
6	-0.076	-0.110	10.463	0.106
7	-0.078	-0.113	10.991	0.139
8	-0.017	-0.082	11.017	0.201
9	0.140	0.073	12.776	0.173
10	0.116	0.024	13.991	0.173
11	-0.017	-0.023	14.018	0.232
12	-0.047	0.013	14.227	0.286
13	0.025	0.042	14.286	0.354
14	0.013	0.003	14.302	0.427
15	0.010	0.055	14.313	0.502
16	0.154	0.224	16.668	0.407
17	-0.059	-0.074	17.025	0.453
18	-0.122	-0.005	18.551	0.420
19	-0.035	0.021	18.683	0.477

(*b*) A statistician has identified the following models based on the analysis of a company's datasets. Review the given information and write down the model equations using appropriate notation. Briefly explain your choice of parameters.

(*i*) MA (4) model

Dependent Variable: D(RAINFALL_DATA)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 06/10/19 Time: 16:56
Sample: 2 80
Included observations: 79
Convergence achieved after 49 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.028077	0.030124	-0.932061	0.3544
MA(1)	-1.157072	0.176645	-6.542868	0.0000
MA(2)	0.032284	0.172782	0.182569	0.8479
MA(3)	-0.079231	0.178314	-0.444331	0.6581
MA(4)	0.409509	0.181359	2.258420	0.0269
SIGMASQ	1.609119	0.421259	3.819689	0.0003

R-squared	0.632051	Mean dependent var	-0.022250
Adjusted R-squared	0.606849	S.D. dependent var	2.103984
S.E. of regression	1.319234	Akaike info criterion	3.641974
Sum squared resid	127.8477	Schwarz criterion	3.721632
Log likelihood	-132.8981	Hannan-Quinn criter.	3.613771
F-statistic	25.07944	Durbin-Watson stat	2.103949
Prob(F-statistic)	0.000000		

Inverted MA Roots	.94-.30i	.94+.30i	-.36+.54i	-.36-.54i

(*ii*) AR (2) Model

Dependent Variable: D(LAKE)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 06/13/19 Time: 09:14
Sample: 2 78
Included observations: 77
Convergence achieved after 17 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000234	0.073727	0.003178	0.9975
AR(1)	0.203212	0.108669	1.870013	0.0655
AR(2)	-0.308431	0.097079	-3.177114	0.0022
SIGMASQ	0.472973	0.079262	5.967240	0.0000

R-squared	0.105381	Mean dependent var	0.006104
Adjusted R-squared	0.068616	S.D. dependent var	0.731876
S.E. of regression	0.706321	Akaike info criterion	2.195969
Sum squared resid	36.41889	Schwarz criterion	2.317725
Log likelihood	-80.54480	Hannan-Quinn criter.	2.244670
F-statistic	2.866330	Durbin-Watson stat	2.053251
Prob(F-statistic)	0.042359		

Inverted AR Roots	.10+.55i	.10-.55i

(*iii*) ARMA (2,2) Model

Dependent Variable: EXPORTS
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 08/28/19 Time: 21:21
Sample: 1995Q1 2017Q1
Included observations: 89
Convergence not achieved after 500 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	286468.1	119759.0	2.392037	0.0190
AR(1)	1.542940	0.067048	23.01296	0.0000
AR(2)	-0.546842	0.063580	-8.600848	0.0000
MA(1)	0.023909	0.084524	0.282861	0.7780
MA(2)	0.996919	0.745112	1.337940	0.1846
SIGMASQ	15559450	11415437	1.362031	0.1769

R-squared	0.998120	Mean dependent var	297671.2
Adjusted R-squared	0.998006	S.D. dependent var	91477.45
S.E. of regression	4084.502	Akaike info criterion	19.70451
Sum squared resid	1.38E+09	Schwarz criterion	19.87228
Log likelihood	-870.8566	Hannan-Quinn criter.	19.77213
F-statistic	8811.406	Durbin-Watson stat	1.710093
Prob(F-statistic)	0.000000		

Inverted AR Roots	.99	.55
Inverted MA Roots	-.01-1.00i	-.01+1.00i

*(c)*Following two outputs are related to two fitted models for a time series. Discuss the adequacy of the fitted models and recommend the most suitable model with justifications.

Model 1

Dependent Variable: X
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 08/29/24 Time: 12:50
Sample: 1 80
Included observations: 80
Convergence achieved after 4 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.301910	0.284496	-1.061243	0.2919
AR(1)	0.479206	0.099113	4.834929	0.0000
SIGMASQ	1.418132	0.262289	5.406756	0.0000

R-squared	0.234230	Mean dependent var	-0.290840
Adjusted R-squared	0.214340	S.D. dependent var	1.369433
S.E. of regression	1.213830	Akaike info criterion	3.265479
Sum squared resid	113.4505	Schwarz criterion	3.354805
Log likelihood	-127.6191	Hannan-Quinn criter.	3.301292
F-statistic	11.77622	Durbin-Watson stat	1.530790
Prob(F-statistic)	0.000034		
Inverted AR Roots	.48		

Model 2

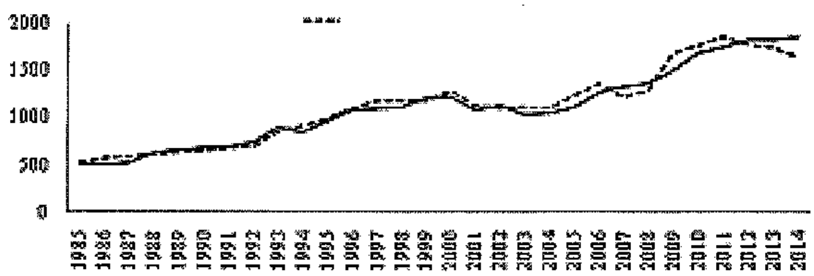
Dependent Variable: X
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 08/29/24 Time: 12:51
Sample: 1 80
Included observations: 80
Convergence achieved after 28 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.283518	0.211109	-1.342996	0.1832
MA(1)	0.751945	0.077668	9.681491	0.0000
SIGMASQ	1.083381	0.217372	4.984001	0.0000

R-squared	0.414991	Mean dependent var	-0.290840
Adjusted R-squared	0.399796	S.D. dependent var	1.369433
S.E. of regression	1.060938	Akaike info criterion	3.003381
Sum squared resid	86.67044	Schwarz criterion	3.092707
Log likelihood	-117.1352	Hannan-Quinn criter.	3.039194
F-statistic	27.31096	Durbin-Watson stat	1.817383
Prob(F-statistic)	0.000000		
Inverted MA Roots	-.75		

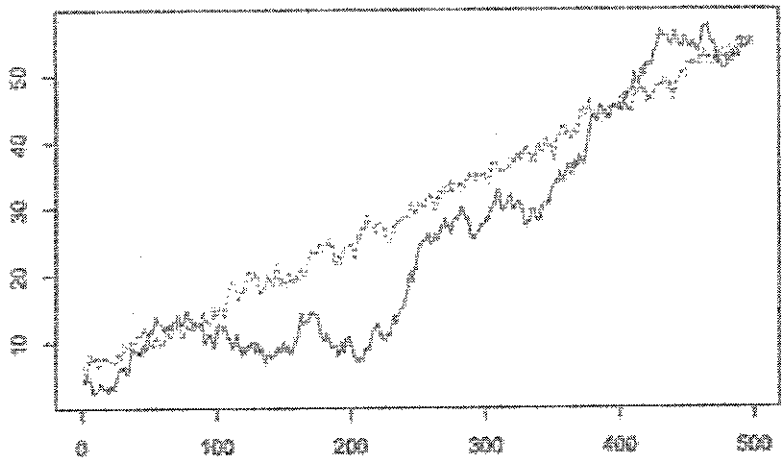
*(d)*A researcher has drawn following graphs to visualize actual verses fitted values for the test set of three time series. Interpret each graph separately.

*(i)*Graph 1: Yearly Mango Production



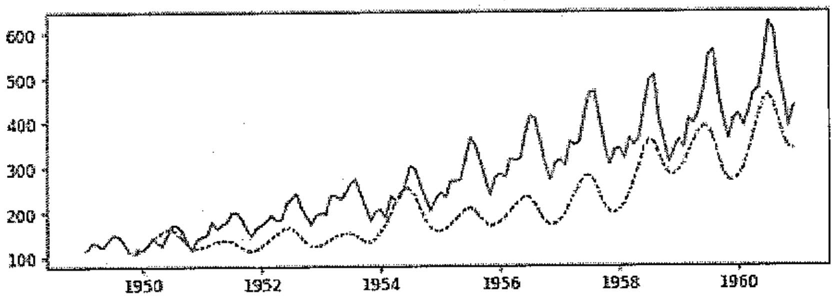
------- Actual ______ Fitted

*(ii)*Graph 2: Daily Sales



____ Actual ------ Fitted

*(iii)*Graph 3: Monthly Air Passengers



____ Actual ------ Fitted

.... End of Paper ....

Helpful Information:

Autocovariance Function of MA(q) Process:

$$\gamma_h = \begin{cases} \sigma_y^2(\theta_0\theta_h + \theta_1\theta_{h+1} + \cdots + \theta_{q-h}\theta_q) \ , & h = 0,1,2,\ldots q \\ 0 \ , & h > q \end{cases}$$

where $\theta_0 = 1$.