Strathmore
UNIVERSITY

STRATHMORE INSTITUTE OF MATHEMATICAL SCIENCES

AND ILAB AFRICA

MASTER OF SCIENCE IN DATA SCIENCE & ANALYTICS

END OF SEMESTER EXAMINATION

DSA 8202 TIME SERIES ANALYSIS AND FORECASTING

DATE: 28th February 2024 Time: 3 Hours

Instructions

1. This examination consists of **FIVE** questions.
2. Answer **Question ONE (COMPULSORY)** and you may choose any other **TWO** additional questions (from Question 2 to 6). Each question is worth 20 marks (This exam has a total of 60 points).
3. Except for Question ONE multiple choice, you must show all of your working for full credit in answering the questions. Correct answers without any working will not earn you any points.

++

Question ONE (COMPULSORY)

1-A. Multiple Choice Questions (10 points total – no need to explain)

1) The concept of "spurious regression" refers to a situation where
a. two or more non-stationary series are regressively analyzed, leading to misleadingly high R-squared values.
b. the regression model mistakenly includes too many lagged terms. *
c. the residuals of the model are perfectly autocorrelated.
d. the model parameters are incorrectly identified as significant due to random chance.

2) Which of the following is not a characteristic of a cointegrated time series?
a. The series move together over time and have a long-run equilibrium relationship.
b. The series can be represented by an error correction model. short
c. Each series is stationary at their level.
d. Deviations from the long-run equilibrium are temporary.

3) In the context of time series analysis, Granger causality tests are used to determine if

$X_t = \alpha + \hat{\beta} Y_t + e_t$

NS NS NS

a. one variable can predict another variable.
b. two series have a unit root.
c. the series is stationary.
d. residuals are normally distributed.

4) The presence of seasonality in a time series implies that
a. there is a deterministic trend in the data.
b. the series exhibits regular patterns at specific periodic intervals.
c. the time series is non-stationary.
d. autocorrelation within the series is always positive.

5) When a time series model includes both autoregressive and moving average terms, it is specifically referred to as a(n)
a. AR model.
b. MA model.
c. ARMA model.
d. VAR model.

6) Impulse Response Functions (IRFs) in the context of Vector Autoregression (VAR) models are used to
a. predict the future values of a time series based on past shocks.
b. determine the causality direction between two time series.
c. measure the response of one variable to a one-time shock to another variable while holding everything else constant.
d. estimate the long-run equilibrium relationship between variables in a cointegrated system.

7) Which of the following is not a method for testing the stationarity of a time series?
a. Phillips-Perron test.
b. Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test.
c. Engle-Granger test.
d. Augmented Dickey-Fuller test.

8) In vector autoregression (VAR) models, the optimal number of lags is often determined using
a. the Cochrane-Orcutt procedure.
b. the Durbin-Watson statistic.
c. information criteria like AIC or BIC.
d. the Phillips-Perron test.

9) Which of the following is true about forecasting with ARIMA models?
a. The differencing order must be determined after estimating the model. x
b. The model can only forecast stationary series. x
c. The model parameters include the number of AR terms, differencing, and MA terms.
d. Forecast accuracy decreases as the forecast horizon increases.

10) Non-stationary time series data can be transformed into stationary data by
a. applying exponential smoothing.
b. taking the natural logarithm of the series.
c. differencing the series.
d. increasing the sample size.

gretl: ADF test

```
Augmented Dickey-Fuller test for YNS
including one lag of (1-L)YNS
sample size 298
unit-root null hypothesis: a = 1

   test without constant
   model: (1-L)y = (a-1)*y(-1) + ... + e
   1st-order autocorrelation coeff. for e: -0.207
   estimated value of (a - 1): -0.00608704
   test statistic: tau_nc(1) = -1.43452
   asymptotic p-value 0.1415

Augmented Dickey-Fuller regression
OLS, using observations 3-300 (T = 298)
Dependent variable: d_YNS

              coefficient   std. error   t-ratio    p-value
  ----------------------------------------------------------
  YNS_1       -0.00608704   0.00424327   -1.435     0.1415
  d_YNS_1      0.506102     0.0503010    10.06      1.12e-020 ***

  AIC: -447.818   BIC: -440.424   HQC: -444.858
```

1-B. (Each sub-question is worth 2 points x 5 = 10 points total)

a) Compare and contrast a white noise process with a stationary process.

b) Sketch possible ACF and PACF for an ARMA (1,2) process.

c) Below are results from a unit root test in GRETL.

```
gretl: ADF test

Augmented Dickey-Fuller test for YNS
including one lag of (1-L)YNS
sample size 298
unit-root null hypothesis: a = 1

   test without constant
   model: (1-L)y = (a-1)*y(-1) + ... + e
   1st-order autocorrelation coeff. for e: -0.207
   estimated value of (a - 1): -0.00608704
   test statistic: tau_nc(1) = -1.43452
   asymptotic p-value 0.1415

Augmented Dickey-Fuller regression
OLS, using observations 3-300 (T = 298)
Dependent variable: d_YNS

              coefficient    std. error    t-ratio    p-value
  -------------------------------------------------------------
  YNS_1      -0.00608704    0.00424327    -1.435     0.1415
  d_YNS_1     0.506102      0.0503010     10.06      1.12e-020 ***

  AIC: -447.818   BIC: -440.424   HQC: -444.858
```

```
Critical Values:
        1%: -3.432
        5%: -2.862
       10%: -2.567
```

Note:

State the null and alternatively clearly for the test. What do you conclude?

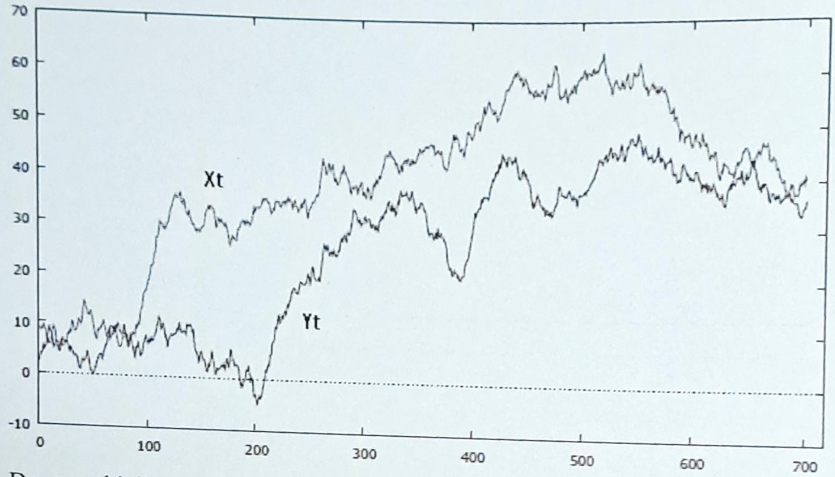
d) Can we improve the Augmented Dickey-Fuller test above by adding more lags? Explain your answer.

e) Explain 3 kinds of non-stationarity we may come across in time series analysis (use figures and equations for your answer).

++

Question TWO (Each sub-question is 5 points x 4 = 20 points total)

Suppose we have two time series X_t and Y_t observed over a period of t=700. A time-series plot is provided below:



(a) Do you think any of the above series are stationary? Are any non-stationary? Which one(s)?

(b) Do you think the two series could be cointegrated? What do you mean by cointegration?

(c) Assuming that Y_t and X_t are I(1) processes, you decide to regress Y_t on X_t by OLS giving:

$$\widehat{Yt} = \underset{(0.87052)}{-7.33993} + \underset{(0.020501)}{0.837184}\, Xt$$

$$T = 700 \quad \bar{R}^2 = 0.7045 \quad F(1, 698) = 1667.6 \quad \hat{\sigma} = 8.5326$$

(standard errors in parentheses)

Would you accept the above results? What would you be worried about? How would you check that your regression is valid?

(d) Alternative specifications to understanding the relationship between Y_t and X_t are:

Dynamic model: $$\widehat{d\_Yt} = \underset{(0.037586)}{0.0469393} + \underset{(0.038184)}{0.0279278}\, d\_Xt$$

Error Correction Model (ECM): $$\widehat{d\_Yt} = \underset{(0.037460)}{0.0468043} + \underset{(0.038056)}{0.0277219}\, d\_Xt - \underset{(0.0043928)}{0.0104865}\, uhat\_1$$

where d_Y_t denotes the first difference of Y_t, d_X_t denotes the first difference of X_t, and uhat_1 is the residual from the model in (c) above.

Of the two models (dynamic and ECM), which one would you accept and why?

++

Question THREE (Each sub-question is 5 points x 4 = 20 points total)

(a) Consider an AR(1) model, $x_t = \varphi x_{t-1} + u_t$. What are the conditions on φ that make x_t stationary? How about non-stationary?

(b) The result below is for some variable x_t.

	Coefficient	Std. Error	t-Statistic	Prob.
C	0.868177	0.334573	2.594884	0.0102
AR(1)	0.975461	0.019471	50.09854	0.0000
MA(1)	-0.909851	0.039596	-22.97840	0.0000

Write down the ARMA(p,q) model in full.

(c) What are the conditions for the ARMA model in (b) to be valid for forecasting?

(d) Given the above results, forecast the one-period-ahead value for x_{t+1}, given $x_t = 100$ and $e_t = 5$

++

Question FOUR (Each sub-question is 5 points x 4 = 20 points total)

a) Explain how the concepts of overfitting and underfitting are intrinsically related forecasting performance?

b) Explain the difference between static and dynamic forecasting.

c) Here are actual and forecasted values for t = 10.

T	1	2	3	4	5	6	7	8	9	10
Actual	250	110	500	200	330	490	670	210	435	375
Forecast	265	140	480	215	290	515	750	210	420	285

Find the Mean Absolute Error (MAE), Mean Square Error (MSE) and the Root Mean Square Forecasting Error (RMSFE).

d) Compare and comment on the values of the RMSE and MAE above.

++

Question FIVE (Each sub-question is 5 points x 4 = 20 points total)

a) Derive the mean and variance of an AR(1) process. What are the conditions for the AR(1) process to be stationary?

b) In time series, what is the difference between autocorrelation and partial autocorrelation?

c) Given a time series, explain how you could check for the presence of an ARCH effect.

d) Clearly explain the ARCH and GARCH models. What are the advantages of GARCH over ARCH model?

++

Question SIX (Each sub-question is 5 points x 4 = 20 points total)

Suppose that industrial production (IP), money supply (M) and trade balance (tb) are jointly determined by a VAR(2,2) model (with a constant be the only exogenous variable)

(a) Write out the VAR(2,2) model in equation form

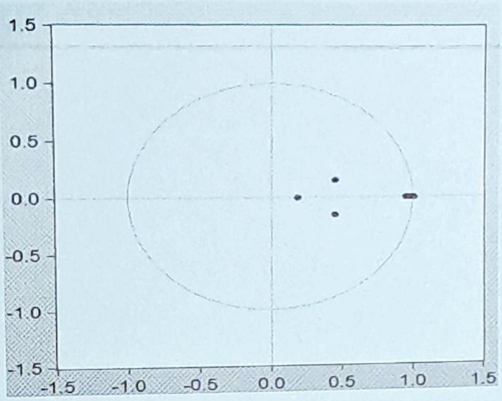
(b) Suppose you ask GRETL to calculate some information criteria as follows:

VAR Lag Order Selection Criteria
Endogenous variables: IP M1 TB3
Exogenous variables: C
Date: 06/02/15 Time: 13:07
Sample: 1959M02 1995M04
Included observations: 428

Lag	LogL	LR	FPE	AIC	SC	HQ
0	NA	NA	22944378	25.46221	25.49067	25.47345
1	NA	7306.121	0.786018	8.272854	8.386661	8.317801
2	NA	264.4260	0.437446	7.686819	7.885982*	7.765478
3	NA	50.97121	0.403867	7.606935	7.891453	7.719304
4	NA	38.17713	0.384206	7.556998	7.926871	7.703078*
5	NA	8.899167	0.392165	7.577454	8.032683	7.757244
6	NA	27.72597	0.382229	7.551720	8.092305	7.765221
7	NA	39.14914*	0.362037*	7.497350*	8.123290	7.744562
8	NA	11.50713	0.367005	7.510852	8.222147	7.791775

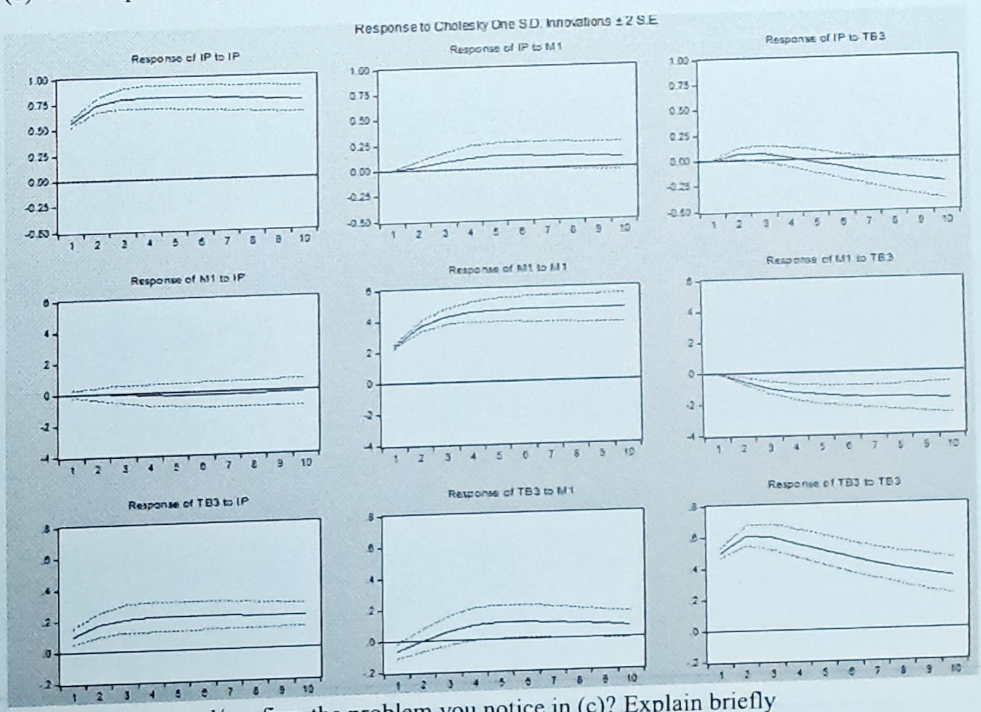
What do you conclude?

(c) However, the inverse roots of the AR characteristic polynomial as follows



What seems to be the problem?

(d) Your impulse response functions are as follows:



Do these correspond/confirm the problem you notice in (c)? Explain briefly

(e) What would you do to solve the problem? What kind of model would you consider running? Briefly explain.

Strathmore
UNIVERSITY

STRATHMORE INSTITUTE OF MATHEMATICAL SCIENCES
AND ILAB AFRICA
MASTER OF SCIENCE IN DATA SCIENCE & ANALYTICS
END OF SEMESTER EXAMINATION
DSA 8202: TIME SERIES ANALYSIS AND FORECASTING

DATE: 1ST March 2023 Time: 3 Hours

Instructions

1. This examination consists of **FIVE** questions.
2. Answer **Question ONE (COMPULSORY)** and any other **TWO** questions (from Question 2 to 6). Each question is worth 20 marks (This exam has total 60 points).
3. Except for Question ONE multiple choice, you must show all of your working for full credit in answering the questions. Correct answers without any working will not earn you any points.

+++

Question ONE (COMPULSORY)

1-A. Multiple Choice Questions (10 points – no need to explain)

1) In time series econometrics, pseudo out-of-sample forecasting can be used for the following reasons with the exception of
   a. giving the forecaster a sense of how well the model forecasts at the end of the sample.
   b. estimating the RMSFE (root mean square forecasting error).
   c. analyzing whether or not a time series contains a unit root.
   d. evaluating the relative forecasting performance of two or more forecasting models.

2) Autoregressive models include
   a. current and lagged values of the error term.
   b. lags of the dependent variable, and lagged values of additional predictor or explanatory variables.
   c. current and lagged values of the residuals.
   d. lags and leads of the dependent variable.

3) Negative autocorrelation in the change of a variable implies that
a. the variable contains only negative values.
b. the series is not stable.
c. an increase in the variable in one period is, on average, associated with a decrease in the next.
d. the data are negatively trended.

4) Stationarity means that the
a. error terms are not correlated.
b. probability distribution of the time series variable does not change over time.
c. time series has a unit root.
d. forecasts remain within 1.96 standard deviation outside the sample period.

5) To choose the number of lags in either an autoregression or a time series regression model with multiple predictors, you can use any of the following test statistics with the exception of the
a. *F*-statistic.
b. Akaike information criterion.
c. Bayes information criterion.
d. augmented Dickey-Fuller test.

6) The random walk model is an example of a
a. deterministic trend model.
b. binomial model.
c. stochastic trend model.
d. stationary model.

7) Heteroskedasticity- and autocorrelation-consistent standard errors

a. result in the OLS estimator being BLUE.
b. should be used when errors are autocorrelated.
c. are calculated when using the Cochrane-Orcutt iterative procedure.
d. have the same formula as the heteroskedasticity robust standard errors in cross-sections.

8) The following is not a consequence of X_t and Y_t being cointegrated:
a. X_t and Y_t are both I(1), then for some θ, $Y_t - \theta X_t$ is I(0).
b. X_t and Y_t have the same stochastic trend.
c. in the expression $Y_t - \theta X_t$, θ is called the cointegrating coefficient.
d. integrating one of the variables gives you the same result as integrating the other.

9) The order of integration
a. can never be zero.
b. is the number of times that the series needs to be differenced for it to be stationary.
c. is the value of ϕ_1 in the quasi difference ($\Delta Y_t - \phi_1 Y_{t-1}$).
d. depends on the number of lags in the VAR specification.

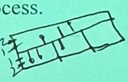
10) To test the null hypothesis of a unit root, the ADF test
   a. has higher power than the so-called DF-GLS test.
   b. uses complicated iterative techniques.
   c. cannot be calculated if the variable is integrated of order two or higher.
   d. uses a test-statistic and a special critical value.

1-B. (10 points)

a) What is a white noise process? What is a stationary process? What is the difference between a white noise process and a stationary process?

white noise just + noise with no pattern = variance mean no autocov

b) Sketch possible ACF and PACF for an ARMA(2,1) process.



c) Below are results from a unit root test.

Null Hypothesis: TBILL has a unit root
Exogenous: Constant
Bandwidth: 3.82 (Andrews using Bartlett kernel)

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-1.519035	0.5223
Test critical values:	1% level	-3.459898	
	5% level	-2.874435	
	10% level	-2.573719	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	0.141569
HAC corrected variance (Bartlett kernel)	0.107615

p-f = high
critical val
less
Non-stat

State the null for the test. What do you conclude?

d) What does it mean to de-trend a series? Explain why de-trending is useful in time series analysis. remove long term isolate short term fluctu

e) Briefly explain what use Theil (U2) statistics is? And what is it used for?

++

Question TWO

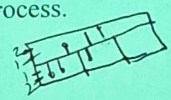
Suppose we have two time series X_t and Y_t observed over a period of t=700. In fact the specific equations (data generating process - DGP) are $X_t = X_{t-1} + e_t$ and $Y_t = Y_{t-1} + e_t$, where $e_t \sim N(0,1)$.

10) To test the null hypothesis of a unit root, the ADF test
   a. has higher power than the so-called DF-GLS test.
   b. uses complicated iterative techniques.
   c. cannot be calculated if the variable is integrated of order two or higher.
   d. uses a test-statistic and a special critical value. ✓

1-B. (10 points)

a) What is a white noise process? What is a stationary process? What is the difference between a white noise process and a stationary process? white noise just ts noise with no pattern ≡ variance mean no autocorr

b) Sketch possible ACF and PACF for an ARMA(2,1) process.



c) Below are results from a unit root test.

Null Hypothesis: TBILL has a unit root
Exogenous: Constant
Bandwidth: 3.82 (Andrews using Bartlett kernel)

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-1.519035	0.5223
Test critical values:	1% level	-3.459898	
	5% level	-2.874435	
	10% level	-2.573719	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	0.141569
HAC corrected variance (Bartlett kernel)	0.107615

p-f = high critical value less Non-stationary

State the null for the test. What do you conclude?

d) What does it mean to de-trend a series? Explain why de-trending is useful in time series analysis. remove long term isolute short term fluctu',

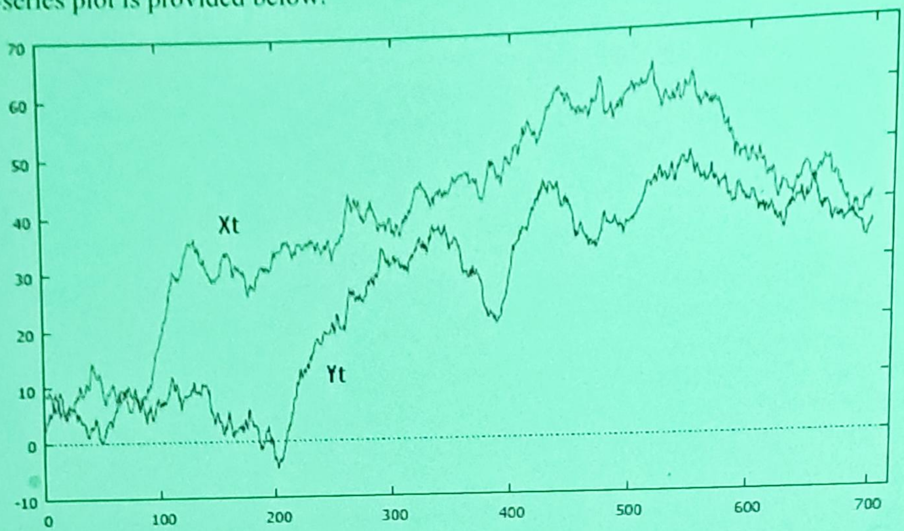
e) Briefly explain what use Theil (U2) statistics is? And what is it used for?

++

Question TWO

Suppose we have two time series X_t and Y_t observed over a period of t=700. In fact the specific equations (data generating process - DGP) are $X_t = X_{t-1} + e_t$ and $Y_t = Y_{t-1} + e_t$, where $e_t \sim N(0,1)$.

A time-series plot is provided below:



(a) Do you think the series are stationary or non-stationary? If they are non-stationary, what kind of non-stationarity would describe the above series best?

(b) To test for stationary, you decide to run Y_t on its lag, Y_{t-1}. Estimation by OLS gives the following:

$$\widehat{Yt} = \underset{(0.071910)}{0.131553} + \underset{(0.0023905)}{0.996752}\,\text{Yt\_1}$$

$$T = 699 \quad \bar{R}^2 = 0.9960 \quad F(1,697) = 1.7386\text{e+}005 \quad \hat{\sigma} = 0.99174$$

(standard errors in parentheses)

Does it make sense to test the null hypothesis H_o: $\beta = 1$ to establish a unit root or non-stationary? Explain why or why not. no - assume

(c) You try a different specification, running the first difference of Y_t or d_Y_t on the lag of Y_t. Estimation by OLS gives:

$$\widehat{\text{d\_Yt}} = \underset{(0.071910)}{0.131553} - \underset{(0.0023905)}{0.00324768}\,\text{Yt\_1}$$

$$T = 699 \quad \bar{R}^2 = 0.0012 \quad F(1,697) = 1.8458 \quad \hat{\sigma} = 0.99174$$

(standard errors in parentheses)

9.45997
7.76357

0.99606
0.99174
0.00426

Based on the above results, is Y_t stationary or non-stationary? State clearly the null and alternative hypothesis. (Use critical Dickey-Fuller statistic of -2.87 for t > 500 and 5% level of significance.)

(d) Assume Y_t and X_t are non-stationary or more specifically I(1). To understand the relationship between Y_t and X_t we regress one on the other by OLS giving:

$$\widehat{Yt} = \underset{(0.87052)}{-7.33993} + \underset{(0.020501)}{0.837184}\,Xt$$

$$T = 700 \quad \bar{R}^2 = 0.7045 \quad F(1, 698) = 1667.6 \quad \hat{\sigma} = 8.5326$$

(standard errors in parentheses)

Would you accept the above results? What in particular would you be worried about? How would you check that your regression is valid?

(e) Alternative specifications to understanding the relationship between Y_t and X_t are:

Dynamic model:
$$\widehat{d\_Yt} = \underset{(0.037586)}{0.0469393} + \underset{(0.038184)}{0.0279278}\,d\_Xt$$

Error Correction Model (ECM):
$$\widehat{d\_Yt} = \underset{(0.037460)}{0.0468043} + \underset{(0.038056)}{0.0277219}\,d\_Xt - \underset{(0.0043928)}{0.0104865}\,uhat\_1$$

where d_Y_t denotes the first difference of Y_t, d_X_t denotes the first difference of X_t, and uhat_1 is the residual from the model in (d) above.

What does the dynamic model tell you about the relationship between Y_t and X_t ? How about the ECM model? Of the two models (dynamic and ECM), which one would you accept and why?

++

Question THREE

The following least squares residuals come from a sample of T=10.

T	1	2	3	4	5	6	7	8	9	10
$\hat{e}_t$	0.28	-0.31	-0.09	0.03	-0.37	-0.17	-0.39	-0.03	0.03	1.02

(a) Calculate the Durbin-Watson statistic, $DW = \dfrac{\sum_{t=2}^{T}(\hat{e}_t - \hat{e}_{t-1})^2}{\sum_{t=1}^{T}\hat{e}_t^2}$. Interpret your results.

(b) Consider an AR(1) model, $x_t = \varphi x_{t-1} + u_t$: What do you think are the conditions on φ that make x_t stationary? How about non-stationary?

$X_t = 0.868177 + 0.975461 X_{t-1}$

(c) The result below is for some variable x_t.

$-0.909851\, e_{t-1}$

$+ e_t$

	Coefficient	Std. Error	t-Statistic	Prob.
C	0.868177	0.334573	2.594884	0.0102
AR(1)	0.975461	0.019471	50.09854	0.0000
MA(1)	-0.909851	0.039596	-22.97840	0.0000

Write down the ARMA(p,q) model in full. $X_t = \phi x_{t-1} + \theta e_{t-1} + e_t$

$X_t = 0.975461$

(d) What are the conditions for the ARMA model in (c) to be valid for forecasting?

(e) Given the above results, forecast the one-period-ahead value for x_{t+1}, given $x_t = 100$ and $e_t = 5$

++

Question FOUR

休 a) Explain the difference between static and dynamic forecasting.

Day 13:

The series below is on Thailand's actual monthly inflation and forecasted inflation by AR(2) by dynamic forecasting method for the year 2014.

date	Actual inflation (π) actual test	Forecast 1 (f1)	predicted today
2013 - Dec	3.63		
2014 - Jan	3.39	3.60	3.63
2014 - Feb	3.23	3.58	3.63
2014 - Mar	2.69	3.56	3.63
2014 - Apr	2.42	3.54	3.63
2014 - May	2.27	3.52	3.63
2014 - Jun	2.25	3.50	3.63
2014 - Jul	2.00	3.48	3.63
2014 - Aug	1.59	3.47	3.63
2014 - Sep	1.42	3.45	3.63
2014 - Oct	1.46	3.44	3.63
2014 - Nov	1.92	3.42	3.63
2014 - Dec	1.67	3.41	3.63
	2994	4.97	

$Var = \frac{1}{f_i} \Sigma f_i$

Evaluate the forecasting performance of forecast 1 (generated by an AR(2) model) with a naïve forecast model (Random Walk) using

$MSE = Bias^2 + Var^2$

MAE =

b) Root mean square forecasting error (RMSFE) and Mean absolute error (MAE).

c) Bias, Variance and Covariance Proportion

$Cov = \frac{1}{f} \Sigma FI$

Var =

d) What do you conclude from the above? Which is better, AR(2) or Random Walk for forecasting?

++

Question FIVE

a) Derive the mean and variance of an AR(1) process. What are the conditions for the AR(1) process to be stationary?

b) In time series, what is the difference between autocorrelation and partial autocorrelation?

c) Given a time series, explain how you could check for the presence of an ARCH effect.

d) Clearly explain the ARCH and GARCH models. What are the advantage of GARCH over ARCH model?

++

Question SIX

Suppose that industrial production (IP), money supply (M) and trade balance (tb) are jointly determined by a VAR(2) model (with a constant be the only exogenous variable)

(a) Write out the model in equation form

(b) What is the main advantage of VAR models over simpler models like the ARMA?

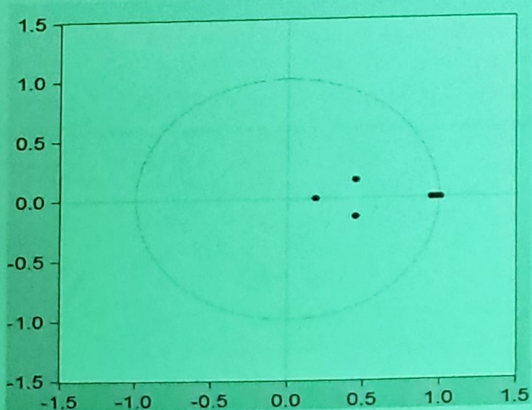
(c) Suppose you ask GRETL to calculate some information criteria as follows:

VAR Lag Order Selection Criteria
Endogenous variables: IP M1 TB3
Exogenous variables: C
Date: 06/02/15 Time: 13:07
Sample: 1959M02 1995M04
Included observations: 428

Lag	LogL	LR	FPE	AIC	SC	HQ
0	NA	NA	22944378	25.46221	25.49067	25.47345
1	NA	7306.121	0.786018	8.272854	8.386661	8.317801
2	NA	264.4260	0.437446	7.686819	7.885982*	7.765478
3	NA	50.97121	0.403867	7.606935	7.891453	7.719304
4	NA	38.17713	0.384206	7.556998	7.926871	7.703078*
5	NA	8.899167	0.392165	7.577454	8.032683	7.757244
6	NA	27.72597	0.382229	7.551720	8.092305	7.765221
7	NA	39.14914*	0.362037*	7.497350*	8.123290	7.744562
8	NA	11.50713	0.367005	7.510852	8.222147	7.791775

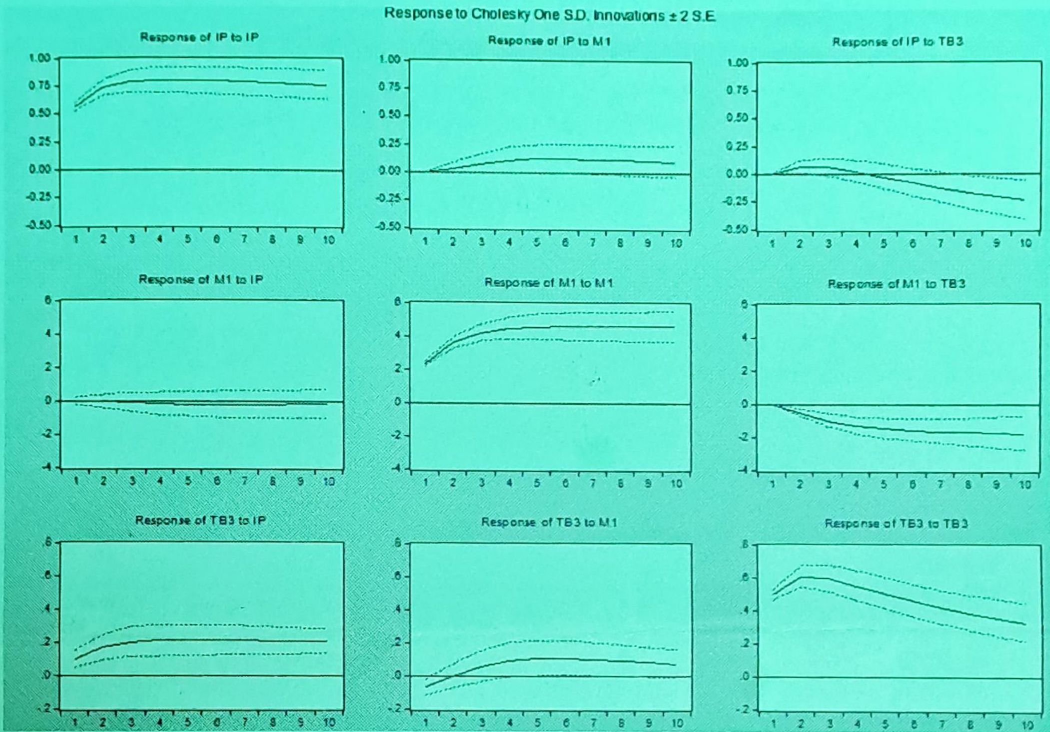
What do you conclude?

(d) However, the inverse roots of the AR characteristic polynomial in EViews is as follows



What seems to be the problem?

(e) Your impulse response functions are as follows:



Do these correspond/confirm the problem you notice in (c) ? Explain briefly

(f) What would you do to solve the problem? What kind of model would you consider running? Briefly explain.

(g) Mention other things you would check to see whether your model fits the data well.

(h) Explain briefly the intuition of a reduced-VAR and a structural VAR? When would you run either a reduced or structural VAR?

(i) Now if you were consider running a VECM model here, what would you have to consider?

(j) Write down a possible VECM model for industrial production (IP), money supply (M) and trade balance (tb) based on all you have done above.