Sample Exam for Analysis of Financial Data

BRIEF SKETCHES OF ANSWERS GIVEN IN BOLD BETWEEN THE QUESTIONS

This exam contains three questions. Please do all questions.

1. *(worth 30% of grade)* Define the following concepts and discuss their importance for financial data analysis.

a) Autoregressive conditional heteroskedasticity
b) The spurious regression problem
c) Variance decomposition

The answer to these questions would just be to briefly summarize textbook material from the relevant sections of the textbook. I will not provide detailed answers. Note that these answers can be mostly written in words, but in some cases it would be good to have some mathematical formulae. For instance, to define ARCH you should write out the ARCH equation (before discussing in words how this can be used to estimate financial volatility and why volatility is important to the finance researcher). For variance decomposition, you might give an example like in the textbook (see pages 207-209 of textbook).

But I do stress that well-written, brief answers that get straight to the point will receive higher grades than long, poorly written answers.

2. *(worth 35% of grade)*

Suppose you had data on a stock price in two different countries. You are interested in causality issues and want to find out whether movements in stock prices in one country effect stock prices in another. Describe all the steps you would go through (i.e. what models you would estimate and what testing procedures you would use) to investigate your question of interest.

You should begin by testing for a unit root in each variable using a Dickey Fuller test. For stock price data you typically find that they have unit roots (but you might want to briefly explain how you would proceed if you did not find unit roots). If both variables have unit roots then you would test for unit root using the Engle-Granger or the Johansen test (provide a few details on one or the other test). If cointegration is present you would then test for Granger causality (provide a few details on what this is) as described on pages 189-190 of the textbook. If cointegration is not present you would difference the data (you might relate this to the idea of building a model for stock returns) and then do Granger causality testing with differenced variables (similar to example on pages 186-187 of textbook). An ambitious answer could bring in VARs and VECMs.

3. (35%)

I have collected data on two financial time series variables, a long term interest rate (X_t) and a short term interest rate (Y_t) and run various regressions using this data. Excel outputs containing results for these regressions are below and labelled as “OUTPUT 1”, “OUTPUT 2”, “OUTPUT 3” and “OUTPUT 4”. To be specific:

- OUTPUT 1 contains results from a regression of ΔY on one lag of Y. That is,
$$\Delta Y_t = \alpha + \beta \times Y_{t-1} + e_t .$$
- OUTPUT 2 contains results from a regression of ΔX on one lag of X.
- OUTPUT 3 contains results from the simple regression of Y on X.
- OUTPUT 4 takes the residuals, e, from the regression of Y on X (i.e. the one in OUTPUT 3) and regresses Δe on one lag of e.

NOTE: For this mock exam, I have just copied Excel outputs into the exam below. In the real exam, I will make up proper tables of results (e.g. in the format of Tables 9.4 or 10.2 in the textbook)

i) Define and describe the Dickey-Fuller test. Can this test be done using any of the OUPTPUTS above? If yes, what does the Dickey-Fuller test tell you about the properties of Y and Y? You may assume that the 5% critical value for the Dickey-Fuller test is -2.89.

THE DICKEY FULLER TEST IS DESCRIBED ON PAGES 154-156 OF THE TEXTBOOK. OUTPUTS 1 AND 2 DO CONTAIN RELEVANT REGRESSIONS. SINCE THE T-STATS ARE SMALL (SMALLER THAN THE DICKEY FULLER CRITICAL VALUE) IN BOTH CASES WE CAN CONCLUDE THAT UNIT ROOTS ARE PRESENT IN BOTH X AND Y.

ii) Define and describe the Engle-Granger test for cointegration. Does cointegration seem to be present in this data set? You may assume that the 5% critical value for the Engle-Granger test is -2.89.

COINTEGRATION TESTING IS DICUSSED ON PAGES 170-173 OF THE TEXTBOOK. OUTPUT 4 CAN BE USED TO DO THE ENGLE GRANGER TEST IS. COMPARING -11.7749 TO THE ENGLE-GRANGER CRITICAL VALUE OF -2.89 WE CAN REJECT THE HYPOTHESIS THAT THE ERRORS HAVE A UNIT ROOT. THUS COINTEGRATION IS PRESENT

iii) Can you obtain an estimate of the long run multiplier from any of these OUTPUTS? If yes, what is the estimate of the long multiplier?

SINCE X AND Y ARE COINTEGRATED, OUTPUT 3 CAN BE USED TO GIVE US A MULTIPLIER OF 1.93891. NOTE, HOWEVER, THAT IF X AND Y WERE NOT COINTEGRATED, THEN OUTPUT 3 WOULD HAVE BEEN A SPURIOUS REGRESSION AND WE WOULD NOT HAVE BEEN ABLE TO USE IT TO CALCULATE THE MULTIPLIER.

OUTPUT 1

Regression Statistics	
Multiple R	0.100336
R Square	0.010067
Adjusted R Square	0.003957
Standard Error	0.149963
Observations	164

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.03705	0.03705	1.647497	0.201133
Residual	162	3.643202	0.022489		
Total	163	3.680253			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.209069	0.148738	1.405625	0.16175	-0.08465	0.502784	-0.08465	0.502784
Y-lagged	-0.01519	0.011833	-1.28355	0.201133	-0.03856	0.008179	-0.03856	0.008179

OUTPUT 2

Regression Statistics	
Multiple R	0.071587
R Square	0.005125
Adjusted R Square	-0.00102
Standard Error	0.010183
Observations	164

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	8.65E-05	8.65E-05	0.834485	0.362336
Residual	162	0.016798	0.000104		
Total	163	0.016885			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.011895	0.002202	5.400638	2.33E-07	0.007545	0.016244	0.007545	0.016244
X-lagged	-0.00148	0.00162	-0.9135	0.362336	-0.00468	0.001719	-0.00468	0.001719

OUTPUT 3

Regression Statistics	
Multiple R	0.993897
R Square	0.987831
Adjusted R Square	0.987755
Standard Error	0.109426
Observations	164

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	157.4576	157.4576	13149.98	5.1E-157
Residual	162	1.939785	0.011974		
Total	163	159.3974			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	9.99487	0.023857	418.9481	8.8E-248	9.947759	10.04198	9.947759	10.04198
X	1.93891	0.017434	114.6734	5.1E-157	1.964787	2.033641	1.964787	2.033641

OUTPUT 4

Regression Statistics	
Multiple R	0.680224
R Square	0.462705
Adjusted R Square	0.459368
Standard Error	0.10957
Observations	163

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	1.664557	1.664557	138.6493	1.75E-23
Residual	161	1.932888	0.012006		
Total	162	3.597445			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-0.00013	0.008583	-0.01468	0.988308	-0.01708	0.016824	-0.01708	0.016824
Resid(-1)	-0.9397	0.079805	-11.7749	1.75E-23	-1.0973	-0.7821	-1.0973	-0.7821