BINDURA UNIVERSITY OF SCIENCE EDUCATION

FACULTY OF COMMERCE

DEPARTMENT OF BANKING AND FINANCE

FINANCIAL ECONOMETRICS (BS450)

DURATION: THREE HOURS

MAR 2024

INSTRUCTIONS TO CANDIDATES:

1) Answer **Question ONE** and any other **THREE** questions

2) Candidates will need non-programmable calculators.

3) Each question carries equal marks.

4) Electronic data saving devices are not allowed into the examination halls.

ADDITIONAL MATERIAL

Standard Normal Distribution Tables

Question One (Compulsory)

(a) In the last decade, stockbrokers have drastically changed the way they do business. Internet trading has become quite common, and online trades can cost as little as $7. It is now easier and cheaper to invest in the stock market than ever before. What are the effects of these changes? To help answer this question, a financial analyst randomly sampled 366 American households and asked each to report the age category of the head of the household and the proportion of its financial assets that are invested in the stock market.

The age categories are:

Young (less than 35)

Early middle age (35 to 49)

Late middle age (50 to 65)

Senior (older than 65)

The analyst was particularly interested in determining whether the ownership of stocks varied by age. Some of the data are listed in the table below.

Young	Early Middle Age	Late Middle Age	Senior
24	28	82	67
35	7	0	77
68	61	61	32
42	53	0	74

Required:

Use Analysis of Variance (ANOVA) to determine if there are differences in stock ownership between the four age groups and comment on your results. (20)

(b) The table below shows output of an Analysis of Variance (ANOVA).

Regression Analysis

The regression equation is
FOOD = 954 + 10.9 INCOME + 748 FSIZE + 565 STUDENT

Predictor	Coef	Stdev	t-ratio	p
Constant	954	1581	0.60	0.563
INCOME	10.92	31.53	0.35	0.738
FSIZE	748.4	303.0	2.47	0.039
STUDENT	564.5	495.1	1.14	0.287

s = 572.7 R-sq = 80.4% R-sq(adj) = 73.1%

Analysis of Variance

SOURCE	DF	SS	MS	F	p
Regression	3	10762903	3587634	10.94	0.003
Error	8	2623764	327970		
Total	11	13386667			

SOURCE	DF	SEQ SS
INCOME	1	4607817
FSIZE	1	5728753
STUDENT	1	426332

Required:

Analyse whether there is statistically significant difference between group means. (5)

[25 Marks]

Question Two

Assume Mr Robert observes the selling price and sales volume of milk for 10 randomly selected weeks. The data he has collected are presented in the Table below:

Week	Weekly Sales (1000s of gallons)	Selling Price ($)
1	10	1.30
2	6	2.00
3	5	1.70
4	12	1.50
5	10	1.60
6	15	1.20
7	5	1.60
8	12	1.40
9	17	1.00
10	20	1.10

Required:

Calculate the following

i. β and interpret the result (4)

ii. α and interpret the result (2)

iii. SST (5)

iv. SSR (2)

v. SSE (5)

vi. Correlation Coefficient (3)

vii. Coefficient of Determination and comment on the goodness of fit. (4)

[25 Marks]

Question Three

a) Data collected from a random sample of 5 General Motors salespersons including statistics from the data are as follows:

Number of Years of Schooling (X1)	Motivation as measured by Higgins Motivation Scale (X2)	Annual Sales in Dollars (Y)
12	32	350 000
14	35	399 765
15	45	429 000
16	50	435 000
18	65	433 000

	Mean	Standard Deviation
Number of Years of Schooling	15	2.236
Motivation	45.4	13.164
Annual Sales	$409 353	$36 116.693

Correlation between Number of Years of Schooling and Motivation ($r_{x1,x2}$) = 0.968

Correlation between Number of Years of Schooling and Annual Sales ($x_{x1,y}$) = 0.880

Correlation between Motivation and Annual Sales ($r_{x2,y}$) = 0.772

Required

Using the figures from the tables above, calculate the following multiple regression parameters and comment on your results.

i. Multiple Correlation coefficient (R) (4)

ii. b_1 (4)

iii. b_2 (4)

iv. α (2)

v. You interviewed a potential sales person and she had 13 years of schooling and she scored 49 on the Higgins Motivation scale. Determine the amount of money the sales person is likely to raise annually. (3)

b) The manager of a departmental store is thinking about establishing a new billing system for the store's credit customers. After a thorough financial analysis, she determines that the new system will be cost-effective only if the mean monthly account is more than \$170. A random sample of 400 monthly accounts is drawn, for which the sample mean is \$178. The manager knows that the accounts are approximately normally distributed with a standard deviation of \$65. $\alpha = 10\%$.

Required:

Using Critical Values Approach, assess if the manager can conclude from this, that the new system will be cost-effective. (8)

[25 Marks]

Question Four

a) Consider the earnings model below: $Wage_i=\beta_0+\beta_1 Exper_i+\beta_2 Educ_i+u_i$, where $Wage$ is measured in dollars per hour, $Exper$ is work experience in years, and $Educ$ is the number of years of schooling.

```
  Source |       SS       df       MS              Number of obs =     100
---------+------------------------------           F(  2,    97) =   16.47
   Model |  2057.5037     2  1028.75185            Prob > F      =  0.0000
Residual |  6059.71269    97  62.4712648           R-squared     =  0.2535
---------+------------------------------           Adj R-squared =  0.2381
   Total |  8117.21639    99  81.9920847           Root MSE      =  7.9039

------------------------------------------------------------------------------
        wage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        Educ |   1.435782   .321546     4.47   0.000     .7976026    2.073962
       Exper |    .328525   .0658247    4.99   0.000     .1978813    .4591687
       _cons |  -11.91922   4.750254   -2.51   0.014    -21.34716   -2.491275
------------------------------------------------------------------------------
```

Required:

Using the regression results shown in the table, summarise the overall goodness of fit of the model.

(5)

b) The table below shows results of the regression assumptions tested using EViews 10.

Null Hypothesis: GDP has a unit root
Exogenous: None
Lag Length: 0 (Automatic - based on SIC, maxlag=9)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.339168	0.0001
Test critical values:	1% level	-2.621185	
	5% level	-1.948886	
	10% level	-1.611932	

Null Hypothesis: D(RMT) has a unit root
Exogenous: None
Lag Length: 1 (Automatic - based on SIC, maxlag=9)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-7.154405	0.0000
Test critical values:	1% level	-2.624057	
	5% level	-1.949319	
	10% level	-1.611711	

F-Bounds Test — Null Hypothesis: No levels relationship

Test Statistic	Value	Signif.	I(0)	I(1)
			Asymptotic: n=1000	
F-statistic	6.617348	10%	2.63	3.35
K	2	5%	3.1	3.87
		2.5%	3.55	4.38
		1%	4.13	5

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.754185	Prob. F(2,29)	0.4741
Obs*R-squared	1.878766	Prob. Chi-Square(2)	0.3429

Heteroskedasticity Test: Breusch-Pagan-Godfrey

F-statistic	0.235907	Prob. F(6,31)	0.0124

Obs*R-squared	1.659295	Prob. Chi-Square(6)	0.0152
Scaled explained SS	4.029729	Prob. Chi-Square(6)	0.0147

Ramsey RESET Test
Equation: UNTITLED
Specification: DINV DINV(-1) DINV(-2) DINV(-3) DINV(-4) DRMT GDP C
Omitted Variables: Squares of fitted values

	Value	df	Probability
t-statistic	0.346914	30	0.7111
F-statistic	0.120349	(1, 30)	0.7111

Required:

Comment on the implications of these assumption results at 5% level of significance. (20)

[25 Marks]

Question Five

Paul Scott, Vice President of Dulibadzimu Water Power, is worried about the possibility of a takeover attempt and the fact that the number of common shareholders have been decreasing since 2001. He instructed you to study the number of common shareholders since 1980 and come up with a forecast for 2023. You decided to investigate the most potential predictor variables namely earnings per share, dividends per share and pay-out ratio.

Required:

a) Critically analyse the steps you would take to formulate the best model to investigate the most potential predictor. (19)

b) Explain in detail any **three** (3) sources of errors in a regression model. (6)

[25 Marks]

QUESTION SIX

Explain the concepts of autocorrelation and heteroscedasticity in financial econometrics, clearly articulating how they affect the estimation of financial models and also highlighting the methods used to address them.

[25 Marks]

END OF EXAMINATION

Formula Sheet (2024)

Hypothesis Test Statistics and Confidence Intervals

1 - α Confidence Interval **Point Estimate $\pm$ Maximum Error E**		**Hypothesis Test Value (Statistic)** NULL Hypothesis: Use the statement containing the condition of equality either directly or implied, as the Null Hypothesis H_0.
(TI-84)	**Single Population**	(TI-84)
One Sample for mean μ (σ is known)		
(ZInterval) $\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$	Use the *Normal* z -Table for the critical value z	(Z-Test) $z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$
One Sample for mean μ (σ is unknown)		
(TInterval) $\bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}$	$df = n - 1$ Use the t-distribution Table for the critical value t	(T-Test) $t = \frac{\bar{x}-\mu}{s/\sqrt{n}}$
One Sample for Proportion p		
(1-PropZInt) $\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$	Use the *Normal* z -Table for the critical value z	(1-PropZTest) $z = \frac{\hat{p}-p}{\sqrt{pq/n}}$

Linear Regression Formulas

$$b = \frac{n\sum XY - \sum X \sum Y}{n\sum X^2 - (\sum X)^2}$$

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)\cdot\left(\sum y^2 - \frac{(\sum y)^2}{n}\right)}}$$

Multiple Regression Formulas

$$\hat{b}_1 = \frac{(\sum x_1 y)(\sum x_2{}^2) - (\sum x_2 y)(\sum x_1 x_2)}{(\sum x_1{}^2)(\sum x_2{}^2) - (\sum x_1 x_2)^2}$$

$$\hat{b}_2 = \frac{(\sum x_2 y)(\sum x_1{}^2) - (\sum x_1 y)(\sum x_1 x_2)}{(\sum x_1{}^2)(\sum x_2{}^2) - (\sum x_1 x_2)^2}$$

$$\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X}_1 - \hat{b}_2 \bar{X}_2$$

The Formula for R

$$R = \sqrt{\frac{\left[\left(r_{y,x1}\right)^2 + \left(r_{y,x2}\right)^2\right] - \left(2 r_{y,x1} r_{y,x2} r_{x1,x2}\right)}{1 - \left(r_{x1,x2}\right)^2}}$$

ANOVA (Analysis of Variance) Formula – F Statistic

Sum of squares Due to Error: $SSE = \sum (y_i - \hat{y}_i)^2$
Total sum of squares: $SST = \sum (y_i - \bar{y})^2$
Sum of Squares Due to Regression: $SSR = \sum (\hat{y}_i - \bar{y})^2$
Relationship Among SST,SSR,and SSE: $SST = SSR + SSE$
Coefficient of determination: $r^2 = \frac{SSR}{SST}$

Source of Variation	Sums of Squares (SS)	Degrees of Freedom (df)	Mean Squares (MS)	F
Between Treatments	$\text{SSB} = \Sigma n_j \left(\bar{X}_j - \bar{X}\right)^2$	k-1	$\text{MSB} = \frac{SSB}{k-1}$	$F = \frac{\text{MSB}}{\text{MSE}}$
Error (or Residual)	$\text{SSE} = \Sigma\Sigma \left(X - \bar{X}_j\right)^2$	N-k	$\text{MSE} = \frac{\text{MSE}}{N-k}$	
Total	$\text{SST} = \Sigma\Sigma \left(X - \bar{X}\right)^2$	N-1		

Autocorrelation Coefficient (r_k)

$$r_k = \frac{\sum_{t=k+1}^{n}(Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2}$$