KENYATTA UNIVERSITY

UNIVERSITY EXAMINATIONS 2010/2011

INSTITUTIONAL BASED PROGRAMME (IBP) – AUGUST SESSION

EXAMINATION FOR THE DEGREE OF MASTER OF SCIENCE IN AGRIBUSINESS MANAGEMENT

KBT 503: ECONOMETRICS

DATE: Tuesday, 28th December, 2010 TIME: 9.00 a.m. – 12.00 p.m.

INSTRUCTIONS:

1. Be sure your **student number and other details specified** are printed clearly on the front of all exam booklet used.
2. **Please label clearly** each of your answers in the exam booklets with the appropriate number and letter.
3. **Please write legibly**.

Answer question **ONE** and any other **THREE**.

1. A researcher is using data for a sample of 25 business schools that offer Agribusiness Masters degrees to investigate the relationship between the annual salary gain of graduates Y_i (measured in thousands of dollars per year) and annual tuition fees X_i (measured in thousands of dollars per year). Preliminary analysis of the sample data produces the following sample information:

$$N = 25 \qquad \sum_{i=1}^{N} Y_i = 1{,}034.97 \qquad \sum_{i=1}^{N} X_i = 528.599 \qquad \sum_{i=1}^{N} Y_i^2 = 45{,}237.19$$

$$\sum_{i=1}^{N} X_i^2 = 11{,}432.92 \qquad \sum_{i=1}^{N} X_i Y_i = 22{,}250.54 \qquad \sum_{i=1}^{N} x_i y_i = 367.179$$

$$\sum_{i=1}^{N} y_i^2 = 2{,}390.67 \qquad \sum_{i=1}^{N} x_i^2 = 256.241 \qquad \sum_{i=1}^{N} \hat{y}_i^2 = 526.147$$

where $x_i \equiv X_i - \overline{X}$, $y_i \equiv Y_i - \overline{Y}$ and $\hat{y}_i \equiv \hat{Y}_i - \overline{Y}$ for i = 1, ..., N. Use the above sample information to answer all the following questions. **Show explicitly all formulas and calculations.**

a) Use the above information to compute OLS estimates of the intercept coefficient β_1 and the slope coefficient β_2.

b) Interpret the slope coefficient estimate you calculated in part (a) – i.e., explain in words what the numeric value you calculated for β_2 means.

c) Calculate an estimate of σ^2, the error variance

d) Compute the value of R^2, the coefficient of determination for the estimated OLS sample regression equation. Briefly explain what the calculated value of R^2 means.

e) Compute the estimated variance of β_2 and the estimated standard error of β_2.

2. You are conducting an econometric investigation into the list prices of food processing machines sold in Sudan in 2010. Your particular interest is in comparing the determinants of list prices for foreign and domestic machines. The sample data consist of 74 observations on the following variables:-

P_i = the list price of machine i, measured in *hundreds* of dollars;

WGT_i = the weight of machine i, measured in *hundreds* of kgs;

F_i = 1 if machine i is a foreign machine, = 0 if machine i is a domestic machine.

The regression model of car prices you propose to use is the pooled regression model:

$P_i = \beta_1 + \beta_2 WGT_i + \beta_3 WGT_i^2 + \beta_4 F_i + \beta_5 F_i WGT_i + \beta_6 F_i WGT_i^2 + u_i$

where the β_j (j = 1, 2,.., 6) are regression coefficients, and u_i is a random error term.

Using the sample data described above, your research assistant computes OLS estimates of regression equation (1) and OLS estimates of two restricted versions of equation (1). For each of the three sample regression equations estimated on the sample of N = 74 observations, the following table contains the OLS coefficient estimates (with estimated standard errors in parentheses below the coefficient estimates) and the summary statistics RSS (residual sum-of-squares), TSS (total sum-of-squares). and N (number of sample observations).

Regressors		(1)	(2)	(3)
Intercept	β_1	159.34 (41.295)	55.267 (37.995)	134.19 (39.978)
WGT_i	β_2	−9.8414 (2.5845)	−3.7146 (2.4334)	−7.2731 (2.6917)
WGT_i^2	β_3	0.19851 (0.039583)	0.11211 (0.038311)	0.15142 (0.043373)
F_i	β_4	−75.918 (126.55)	32.470 (6.4941)	----
F_iWGT_i	β_5	3.4648 (9.9967)	----	----
$F_iWGT_i^2$	β_6	0.032625 (0.19341)	----	----
	RSS =	21496.33	28352.19	38477.99
	TSS =	63506.54	63506.54	63506.54
	N =	74	74	74

Note: The symbol "----" means that the corresponding regressor was omitted from the estimated sample regression equation. The figures in parentheses below the coefficient estimates are the estimated standard errors. RSS is the Redual Sum-of-Squares, TSS is the Total Sum-of-Squares, and N is the number of sample observations.

a) Use the estimation results for regression equation (1) in column (1) to compute the OLS estimates of the intercept coefficient, the slope coefficient of WGTi and the slope coefficient of WGT_i^2 for *domestic* machines. Use the estimation results for regression equation (1) in column (1) to compute the OLS estimates of the intercept coefficient, the slope coefficient of WGT_i and the slope coefficient of WGT_i^2 for *foreign* machines.

a) Compare the goodness-of-fit to the sample data of the three sample regression equations (1), (2) and (3) in the table. Calculate the value of an appropriate goodness-of-fit measure for each of the sample regression equations (1), (2) and (3) in the table. Which of the three sample regression equations provides the best fit to the sample data? Which of the three sample regression equations provides the worst fit to the sample data?

3. The sales manager of a large retail chain has been given the task of determining why certain stores have been performing more poorly than expected. He believes that part of the problem is that some stores appear to be victims of an inefficient distribution system. Other possible factors include the size of the population in the surrounding area and the amount of advertising used. He decides to take a sample of 50 stores, and for each he determines the following variables:

y = total sales last year (in \$ 1,000s)

x_1 = population (in 1,000s)

x_2 = advertising expenditure (in \$1,000s)

x_3 = distribution-efficiency index (ranging from 1 = poor distribution to 10 = excellent distribution; this is an ordinal scale, but it will be treated as an interval-scaled variable

The data were run on SPSS software package, with the following results:

MULTIPLE R	0.7625
R SQUARE	0.5814
ADJUSTED R SQUARE	0.5541
STANDARD ERROR	6.9051

ANALYSIS OF VARIANCE	DF	SUM OF SQUARES	MEAN SQUARE	F
REGRESSION	3	3046.1646	1015.3882	21.2957
RESIDUAL	46	2193.2984	47.6804	
		5239.463		

VARIABLES IN EQUATION

VARIABLE	B	STD ERROR B	F
x_1	34.0791	4.1138	68.6262
x_2	0.1259	0.1063	1.4028
x_3	-0.9932	0.6551	2.2986
Constant	13.6225	6.7198	4.1096

a) Can we conclude at the 5% significance level that this model is useful in analyzing annual sales?

b) Do these data provide sufficient evidence at the 5% significance level to enable us to conclude that the distribution-efficiency index is a factor in determining total annual sales?

c) Does the amount of advertising influence sales? (Test with $\alpha = 0.05$).

4. A meat packer was interested in getting a better understanding of the variables that influence the price of beef in Rwanda. After some thought, she produced the following list of 4 variables that she believes affect the price of beef (in hundred weights, on the hoof). These are:

- Average beef consumption
- Price of chicken
- Price of pork
- Price of lamb

Using data from the last 120 months and a statistical package she produced the data shown in the print out below:-

Predictor	**Coef**	**Stdev**	**t-ratio**
Constant	5.260	7.982	0.659
BEEF	1.259	0.397	3.174
CHICKEN	1.214	0.601	2.020
PORK	2.386	2.708	0.881
LAMB	1.366	0.544	2.513

a) Are the signs of β_1, β_2, β_3 and β_4 in the expected direction? Explain

b) Which of the 4 independent variables are linearly related to the dependent variable? (Test with $\alpha = 0.05$)

5. You have been commissioned to investigate the relationship between annual Research and Development (R & D) expenditures (Y) and total annual sales

revenues (X) for agribusiness firms. You have assembled data for a sample of 32 agribusiness firms, where Y_i is annual R & D expenditures of the i-th firm (measured in millions of dollars per year) and X_i is total annual sales revenues of i-th firm (measured in millions of dollars per year). Your research assistant has used the sample data to estimate the following OLS sample regression equation, where the figures in parentheses below the coefficient estimates are the estimated standard errors of the coefficient estimates:

$$Y_i = \underset{(20.515)}{-0.5772} + \underset{(0.0024487)}{0.04063}X_i + u_i \qquad (i = 1,\ldots,N) \quad N = 32$$

a) Compute the two-sided 95% confidence interval for the slope coefficient β_2.

b) Perform a test of the null hypothesis $H_0: \beta_2 = 0$ against the alternative hypothesis $H_1: \beta_2 \neq 0$ at the 1% significance level (i.e., for significance level $\alpha = 0.01$). Show how you calculated the test statistic. State the decision rule you use, and the inference you would draw from the test. Briefly explain what the test outcome means.

6. Given 8 observations of x and y, the following summations were computed:

$\sum x = 40 \qquad \sum y = 100 \qquad \sum xy = 600 \qquad \sum x^2 = 250 \qquad \sum y^2 = 2{,}000$

a) Determine a 99% confidence interval for the expected value of y when x =

b) Determine a 95% prediction interval for the value of y when x = 3.