NAMIBIA UNIVERSITY
OF SCIENCE AND TECHNOLOGY

FACULTY OF HEALTH, NATURAL RESOURCES AND APPLIED SCIENCES

DEPARTMENT OF AGRICULTURE AND NATURAL RESOURCES SCIENCES

QUALIFICATION: BACHELOR OF SCIENCE IN AGRICULTURE (AGRIBUSINESS MANAGEMENT)	
QUALIFICATION CODE: 07BAGA	**LEVEL:** 7
COURSE CODE: BEA611S	**COURSE NAME:** BASIC ECONOMETRICS FOR AGRICULTURE
DATE: JANUARY 2023	
DURATION: 3 HOURS	**MARKS: 100**

SECOND OPPORTUNITY/SUPPLEMENTARY EXAMINATION QUESTION PAPER	
EXAMINER(S)	PROF DAVID UCHEZUBA
MODERATOR:	MR MWALA LUBINDA

INSTRUCTIONS
1. Answer ALL the questions. 2. Write clearly and neatly. 3. Number the answers clearly.

PERMISSIBLE MATERIALS

1. Examination question paper
2. Answering book

THIS QUESTION PAPER CONSISTS OF 9 PAGES (Excluding this front page)

Section 1 Multiple choice

Question 1

Consider the following models

$E(y/x_i) = \beta_1 + \beta_2 x_i^2 + \mu_i$ .... (A)

$E(y/x_i) = \beta_1 + \beta_2^2 x_i + \mu_i$ .... (B)

Which of the following statements about equations (A) and (B) is **incorrect**?

A) Equation (A) is linear in parameter and (B) is non-linear in parameter
B) Equation (A) is linear in variable and (B) is non-linear in variable
C) Equation (A) is non-linear in variable and (B) is linear in variable
D) Equation (A) is linear in parameter and (B) linear in variable

Question 2

The farmer's consumption function is fitted as

$y_i = \beta_1 + \beta_2 x_i + \mu_i$

Which of the following is **INCORRECT** about why μ_i was included in the model?

A) We do not know other variables affecting consumption expenditure (y)
B) Even if we know, we may not have information (data) about all factors affecting (y)
C) There may be measurement errors in the way data was collected
D) We include μ_i because it is a non-random and systematic component of the model

Question 3

According to the Gauss-Markov theorem, which of the following statements is **NOT CORRECT**?
An estimator says the ordinary least square (OLS) estimator $\hat{\beta}_2$, is said to be the best linear unbiased estimator of β_2, if the following conditions hold.

A) $\hat{\beta}_2$, must be a linear function of the dependent variable (y)
B) $\hat{\beta}_2$, must be unbiased, i.e, its average or expected value $E(\hat{\beta}_2)$, must be equal to β_2
C) $\hat{\beta}_2$, must have minimum variance
D) $\hat{\beta}_2$, must have a mean of zero

Question 4

An unbiased estimator such as $\hat{\beta}_2$, with the least (minimum) variance is said to be
A) An inefficient estimator
B) An efficient estimator
C) A random noise
D) An asymptote

Question 5

Consider the following regression model estimated using the OLS method

$$Y = \underset{(91.2835)}{463.5136} - \underset{(0.1213)}{0.3901x_1} + \underset{(0.0477)}{0.17925x_2}$$

(Standard errors are in parenthesis)

Using equation (12.1), calculate the t-statistic for the x_1 and x_2 variables

A) 3.2159 and 3.710
B) 3.2801 and 3.7578
C) 3.2159 and 3.7578
D) 3.2009 and 3.7011

Question 6

Which one of the following **incorrectly** defines the coefficient of correlation between variables?
A. Its value is between -1 and +1.
B. It can be positive or negative
C. It is a measure of association
D. It is independent of the origin and scale
E It is the same as R^2

Question 7

The statistical significance of a parameter in a regression model refers to:
a) The conclusion of testing the null hypothesis that the parameter is equal to zero, against the alternative that it is non-zero.
b) The probability that the OLS estimate of this parameter is equal to zero.
c) The interpretation of the sign (positive or negative) of this parameter.
d) All of the above

Question 8

All of the following are possible effects of multicollinearity EXCEPT:
a) the variances of regression coefficients estimators may be larger than expected
b) the signs of the regression coefficients may be opposite of what is expected
0) a significant F ratio may result even though the t ratios are not significant
d) removal of one data point may cause large changes in the coefficient estimates
6) the VIP is zero

Question 9

Suppose that you estimate the model Y =50 +B$_{I}$X + u. You calculate residuals and find that the explained sum of squares is 400 and the total sum of squares is 1200. The R-squared is

a) 0.25
b) 0.33
c) 0.5
d) 0.67

Question 10

In linear regression, the assumption of homoscedasticity is needed for
I. unbiasedness
II simple calculation of variance and standard errors of coefficient estimates.
III. the claim that the OLS estimator is BLUE.

a) I only.
b) II only.
c) III only.
d) II and III only.
e) I, II, and III.

Question 11
Which of the following is/are consequences of over-specifying a model (including irrelevant variables on the right-hand side)?
I. The variance of the estimators may increase.
II. The variance of the estimators may stay the same.
III. Bias of the estimators may increase.
a) I only.
b) II only.
c) III only.
d) I and II only.
e) I, II, and III.

Question 12
Heteroscedasticity means that
a) Homogeneity cannot be assumed automatically for the model.
b) the observed units have different preferences.
c) the variance of the error term is not constant.
d) agents are not all rational.

Question 13
By including another variable in the regression, you will
a) look at the t-statistic of the coefficient of that variable and include the variable only if the coefficient is statistically significant at the 1% level.
b) eliminate the possibility of omitted variable bias from excluding that variable.
c) decrease the regression R2 if that variable is important.
d) decrease the variance of the estimator of the coefficients of interest.

Question 14
Which of the following statements is TRUE concerning OLS estimation?
a) OLS minimises the sum of the vertical distances from the points to the line
b) OLS minimises the sum of the squares of the vertical distances from the points to the line
c) OLS minimises the sum of the horizontal distances from the points to the line
(d) OLS minimises the sum of the squares of the horizontal distances from the points to the line.

Question 15

The residual from a standard regression model is defined as

a) The difference between the actual value, y, and the mean, $\overline{y}$

b) The difference between the fitted value, $\hat{y}$, and the mean, $\overline{y}$

c) The difference between the actual value, y, and the fitted value, $\hat{y}$

d) The square of the difference between the fitted value, $\hat{y}$, and the mean, $\overline{y}$

Section 2 True or False

Question 1

If the null hypothesis is not rejected, it is true. True or False.

Question 2

The higher the value of the σ^2, the larger the variance of $\hat{\beta}_2$ given in question 4. True or False

Question 3

The conditional and unconditional means of a random variable are the same thing. True or False.

Question 4

In the two-variable population regression function (PRF), if the slope coefficient β_2 is zero, the intercept β_1 is estimated by the sample mean $\overline{Y}$.

Question 5

The conditional variance, var $(Y_i \mid X_i) = \sigma^2$, and the unconditional variance of Y, var $(Y) = \sigma^2_Y$, will be the same if X had no influence on Y.

Section 3 – General

Question 1

Question 1.1. What is the meaning of the following econometrics terms

i).	Intercept (constant)	(1 mark)
ii).	Cross-section data	(1 mark)
iii).	Response variable	(1 mark)

iv).	Linear regression line	(1 mark)
v).	Predictor	(1 mark)
vi).	Linear model	(1 mark)
vii).	Multivariate model	(1 mark)
viii).	Regression analysis	(1 mark)
ix).	Residual	(1 mark)
x).	Slope coefficient	(1 mark)

Question 1.2.

Using hypothetical data, the relationship between child nutrition and stunting was estimated as follows.

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$$

Where, Y = Average height of pupils aged 5 (measured in metres) and X = Household Dietary Diversity Score (a measure of the diversity of food intake).

The estimated coefficients are

$\hat{\beta}_1 = 0.088$ (0.0412), $\hat{\beta}_2 = 0.7165$ (0.2547), $R^2 = 0.91$.

(Figures in parenthesis are standard errors).

1.2.1. Interpret the slope coefficient	(2 marks)
1.2.2. Calculate the T-statistic for the slope coefficient.	(2 marks)
1.2.3. Calculate the T-statistic for the intercept coefficient	(2 marks)
1.2.4. Interpret the the R^2 value	(2 marks)
1.2.5. Give two properties of the coefficient of correlation between Y and X	(2 marks)

Question 2

In a model $y_i = \alpha + \beta x_i + u_i, i = 1,\ldots, N$, the following sample moments have been calculated from 10 observations.

$\sum Y = 8$, $\sum X = 40$, $\sum (Y - \bar{Y})^2 = 26$, $\sum (X - \bar{X})^2 = 200$, and $\sum (X - \bar{X})(Y - \bar{Y}) = 20$

2.1. Estimate the slope parameter (4 marks)

2.2. Estimate the intercept parameter (4 marks)

2.2. Determine the function $\hat{y}$ (3 marks)

2.3. Calculate the value $\hat{y}$ for $x = 10$ (3 marks)

2.4. Obtain the 95% confidence interval for the calculated $\hat{y}$ (6 marks)

Question 3

Consider the following regression model, $y_t = \beta_1 + \beta_2 x_t + \varepsilon_t$. Where, y_t = consumption expenditure, x = income, β_1 = Constant, β_2 = Slope, ε = Error term. Which of the above

3.1. Has fixed values in repeated sampling. (2 mark)

3.2 Is a stochastic variable. (2 mark)

3.3 Is a non-stochastic variable (2 mark)

3.4 Has zero mean in a classical linear regression. (2 mark)

3.5 Is a parameter. (2 mark)

3.6. The analysis of the variance of a regression model is given below.

	df	SS	MS	F	Significance F
Regression	1	-	0.2701	745.9286	0.0000
Residual	-	0.0040	-		
Total	12	0.2741			

i). Complete the table (6 marks)

ii) What is the null hypothesis of this test? (2 marks)

iii) Do you reject or fail to reject this null? Why? (2 marks)

Question 4

A post-regression Breusch-Pagan-Godfrey test was conducted to test for a violation of a classical linear regression assumption. The result results of the test are shown below.

Description	Statistics
F-statistics	3.1405
P-value (F1, 11)	0.1040

4.1	What is the name of this test?	(2 marks)
4.2.	What is the null hypothesis for this test	(2 marks)
4.3.	Do you reject or fail to reject this null? Why?	(2 marks)
4.4.	What are the implications of violating this assumption	(4 marks)
4.5.	How do you detect this assumption is violated?	(6 marks)
4.6.	What measure would you adopt to remedy this problem?	(4 marks)

Statistical formula

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

$$\hat{\beta}_2 = \frac{\sum(Y_i - \bar{Y})(X_i - \bar{X})}{(X_i - \bar{X})}$$

$$\hat{\sigma}^2 = \frac{\sum\hat{\mu}_i^2}{n-2}$$

$$\text{var}(\hat{\beta}_2) = \frac{\hat{\sigma}^2}{\sum(X_i - \bar{X})^2}$$

$$\text{var}(\hat{\beta}_1) = \frac{\sum(X_i - \bar{X})^2}{n\sum(X_i - \bar{X})^2}$$

$$R^2 = 1 - \frac{\sum\mu_i^2}{\sum(Y_i - \bar{Y})^2}$$

$$\hat{\beta}_2 = \frac{\sum(Y_i - \bar{Y})(X_i - \bar{X})}{\sum(X_i - \bar{X})^2}$$

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

$$se(\hat{\beta}_1) = \sqrt{\frac{\sum X_i^2}{n\sum(X_i - \bar{X})^2}}$$

$$se(\hat{\beta}_2) = \frac{\sigma}{\sqrt{\sum(X_i - X)^2}}$$

$$se(\hat{\beta}_1) = \sqrt{\frac{\sum X_i^2}{n\sum(X_i - \bar{X})^2}\sigma}$$

$$se(\hat{\beta}_2) = \frac{\sigma}{\sqrt{\sum n(X_i - X)^2}}$$

$$JB = n\left[\frac{S^2}{6} + \frac{(K-3)^2}{24}\right]$$

$$JB = n\left[\frac{S^2}{6} - \frac{(K-3)^2}{22}\right]$$

See the attached Durbin Watson table

	$k' = 1$		$k' = 2$		$k' = 3$		$k' = 4$	
n	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
6	0.610	1.400	—	—	—	—	—	—
7	0.700	1.356	0.467	1.896	—	—	—	—
8	0.763	1.332	0.559	1.777	0.368	2.287	—	—
9	0.824	1.320	0.629	1.699	0.455	2.128	0.296	2.588
10	0.879	1.320	0.697	1.641	0.525	2.016	0.376	2.414
11	0.927	1.324	0.658	1.604	0.595	1.928	0.444	2.283
12	0.971	1.331	0.812	1.579	0.658	1.864	0.512	2.177
13	1.010	1.340	0.861	1.562	0.715	1.816	0.574	2.094
14	1.045	1.350	0.905	1.551	0.767	1.779	0.632	2.030
15	1.077	1.361	0.946	1.543	0.814	1.750	0.685	1.977
16	1.106	1.371	0.982	1.539	0.857	1.728	0.734	1.935
17	1.133	1.381	1.015	1.536	0.897	1.710	0.779	1.900
18	1.158	1.391	1.046	1.535	0.933	1.696	0.820	1.872

END