

M.A. Examination, 2019
Semester-II
Economics
Course- C9 [New] /Course – VI [Old]
(Computer Application)

Time: Three Hours

Full Marks: 40

Questions are of value as indicated in the margin

Answer *any four* questions

1. (a) What do you mean by ‘unit of analysis’ in the context of survey data? What should be the units of analysis for the following research questions?

- (i) Factors affecting monthly expenditures of students staying away from home for their study
- (ii) Empowerment of women through their involvements in Self-Help Groups
- (iii) Financial performance of Self Help Groups
- (iv) Productivity of agricultural plots
- (v) Factors affecting per-capita consumption expenditure of households

(b) What do you mean by ‘open-ended’ and ‘close-ended’ questions in a primary survey? Given a suitable example of an open-ended question and then modify it to a close-ended one.

(c) Suppose the following question is put to a household respondent in a primary survey:

*“Put tick on the following sources of earning for your household over last one year:
1 = Own agriculture; 2 = Daily labour; 3 = migrant labour; 4 = Own business; 5 = Salaried employment; 6 = Any other”*

How many columns will you use to enter the data in a spreadsheet for this response? Explain your answer.

5+3+2=10

2. (a) Explain the concept of ‘p-value’ as reported against a regression result (assume least square regression) by a statistical software. Show the position of estimated t-value for Regressor3 in Question #4 (below) in a diagram representing the appropriate t-distribution (you can use up to two decimal points for reported t-value and p-value for simplicity).

(b) What do you mean by ‘degrees of freedom’? What will be the degrees of freedom for the following statistics?

- (i) $\sum_{i=1}^n (x_i - 10)^5$, where x_i ’s are sample values randomly selected from a large population.
- (ii) $\sum_{i=1}^7 e_i^2$ where e_i ’s are estimated residual terms of a 5-variable linear regression estimated by OLS

6+4=10

3. Suppose you need to estimate the parameters of a production function

$$Q = A.L^\alpha.K^\beta.u \quad (\text{where } u \text{ is an error term})$$

Raw data on production (Q_i), labour input (L_i) and capital employed (K_i) is provided in an excel sheet. Describe in details how you would proceed to estimate the parameters.10

P.T.O.

(2)

4. Consider the following as part of an Excel spreadsheet showing the result of an OLS regression:

	A	B	C	D	E	F
1	Multiple R	0.242421				
2	R Square	0.058768				
3	Adjusted R Square	0.046260				
4	Standard Error	6.923368				
5	Observations	?				
6	ANOVA					
7		<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
8	Regression	4	900.8371	?	4.698416	0.001083
9	Residual	301	?	47.93303		
10	Total	?	15328.68			
11						
12		<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	
13	Intercept	?	8.636340	-1.275470	0.203127	
14	Regressor 1	-0.000054	0.000020	-2.750610	0.006309	
15	Regressor2	0.045694	0.021137	2.161791	0.031422	
16	Regressor3	0.016212	?	1.831304	0.068044	
17	Regressor4	-0.064860	0.034223	-1.895300	0.059011	

(a) What formulae would you use to fill up the cells indicated by interrogation mark ('?')? [Use cell references in your formulae, not the values].

(b) Point out the significance levels against each regressor following standard demarcations (10%, 5% and 1% levels)

(c) Comment on the overall significance of the regression result.

(d) What do you mean by the value of 'standard error' given in cell B4? $3+4+1+2=10$

5. Consider an individual level data involving three variables X, Y and D in a data set such that $Y = f(X, D)$. Dis a binary variable representing gender of the individual and is coded as 1 or 2 (1=Male; 2=Female). It is known that the functional relationship between Y and X is linear. You want to check whether the gender affects the slope, intercept or both slope and intercept in the relationship between Y and X.

- Write down the model you would like to estimate.
- Write down the STATA commands sequentially that will give you the required regression result (assuming the data file is already open in STATA)
- Explain your conclusion relating to the different possible outcomes (significance/insignificance) of the parameter estimates. $2+4+4=10$

(3)

[Question #6, #7 and #8 are based on the data described at the end of this question paper]

6. Given the description of the data at the end of this question paper, sequentially write the command lines for a single STATA DO-file that will execute the following tasks:

- (i) Open the data file, generate a new variable 'fsize' that calculates the number of household members
- (ii) Generate a dummy variable named 'migrant' such that its value is '1' if the corresponding household has any migrant labour, and '0' otherwise.
- (iii) Calculate the total agricultural landholding in katha and store the value in a new variable named 'totkatha'.
- (iv) Put a command that results in a table showing mean value of agricultural land across the 'migrant' status of the households
- (v) Save the data file under the name 'answer_6' at the same location with an option to overwrite any existing file with the same name. 10

7. Given the description of the data at the end of this question paper, sequentially write the command lines for a single STATA DO-file that will execute the following tasks:

- (i) Open the data file, delete all the variables except household serial number, age and sex information of the household members
- (ii) Rearrange the data at individual level (long format)
- (iii) Create a new variable called 'ageclass' which is defined as follows:
ageclass=1 means the age is less than 30
ageclass=2 means the age is greater than or equal to 30 and less than 60
ageclass=3 means the age is greater than or equal to 60
- (iv) Label the 'ageclass' variable values as 'young', 'middle age' and 'old' respectively
- (v) Put a command that generates frequency distribution of 'ageclass'
- (vi) Put a command that generates frequency distribution of 'ageclass' for male members only
- (vii) Put a command that results in a table showing mean value of age across male and female
- (viii) Save the data file under the name 'answer_7' at the same location with an option to overwrite any existing file with the same name. 10

8. Given the description of the data at the end of this question paper, sequentially write the command lines for a single STATA DO-file that will execute the following tasks:

- (i) Open the data file, generate a new variable 'fsize' that calculates the number of household members
- (ii) Generate a new variable named 'mpcc' which represents per capita monthly consumption expenditure of the households
- (iii) You want to run a linear regression where 'mpcc' is to be explained by age of the head of the household, number of family members in the household and total agricultural landholding. Generate required variables for this purpose (if not already existing) and arrange them in that order in your dataset. Delete all other variables.
- (iv) Put a command that executes the regression and shows the result.
- (v) Save the data file under the name 'answer_8' at the same location with an option to overwrite any existing file with the same name. 10

P.T.O.

Description of data for question #6, #7 and #8:

Location of data file on computer: "C:\Users\Desktop\STATA\Test.dta"

Number of observation: **100** [household level survey data]

Variable name	Variable description
hhs1	household serial number (1, 2,3.....100)
age1	Age of the head of the household (first member) in completed years
sex1	Sex of the head of the household (first member): coded as 1 = Male; 2 = Female
occu1	Occupation code of the head (first member) of the household (code details are provided below)
age2*	Age of the second household member in completed years
sex2*	Sex of the second household member: coded as before
occu2*	Occupation code of the second household member
age3*	Age of the third household member in completed years
sex3*	Sex of the third household member: coded as before
occu3*	Occupation code of the third household member
age4*	Age of the fourth household member in completed years
sex4*	Sex of the fourth household member: coded as before
occu4*	Occupation code of the fourth household member
age5*	Age of the fifth household member in completed years
sex5*	Sex of the fifth household member: coded as before
occu5*	Occupation code of the fifth household member
religion	Religion code of the household: 1 = Hindu; 2 = Muslim; 3 = Others
caste	Caste code of the household: 1 = SC; 2 = ST; 3 = OBC; 4 = General
landb*	Agricultural landholding of the household (in bigha)
landk*	Agricultural landholding of the household (in katha)
mcon	monthly average total consumption expenditure of the household (Rs.)

** Implies some of the values of these variable might be missing (i.e, appears as 'dots')*

Codes used for occupation 1 = Own agriculture; 2 = Daily labour; 3 = migrant labour; 4 = Own business; 5 = Salaried employment; 6 = Any other