MACHAKOS UNIVERSITY

University Examinations for 2020/2021 Academic Year

SCHOOL OF BUSINESS AND ECONOMICS

DEPARTMENT OF ECONOMICS

THIRD YEAR FIRST SEMESTER EXAMINATION FOR

BACHELOR OF ECONOMICS AND FINANCE

EES 303: SAMPLE SURVEY

DATE: 16/8/2021 **TIME: 11.00-1.00 PM**

INSTRUCTIONS:

(i) Answer question one (Compulsory) and any other two questions

(ii) Do not write on the question paper

(iii) Show your working clearly

QUESTION ONE (Compulsory) (30 MARKS)

a) Distinguish clearly between observational and non- observational errors. (2 marks)

b) Explain the main sources of observational errors (10 marks)

c) Given two populations of size 95 and 240, use the provided table of random numbers to draw a simple random sample of size 5 and 10 respectively from the populations. Explain the steps followed in each case. (10 marks)

d) Distinguish between the following terms as used in sampling survey;

i. Target population and Accessible population (2 marks)

ii. Judgemental sampling and Quota sampling (2 marks)

iii. Questionnaire and an Interview (2 marks)

iv. Systematic bias and sampling error (2 marks)

QUESTION TWO (20 MARKS)

a) Under what circumstances would a stratified random sampling considered appropriate? Using an illustration, explain how you can select such a sample. (6 marks)

b) A simple random sample of 100 water meters within a community is monitored to estimate the average daily water consumption per household over a specified dry spell. The sample mean and sample variance are found to be $\bar{y}$ = 12.5 and s^2 = 1252. If we assume that there are N = 10,000 households within the community;

i) Calculate the interval estimate of the true mean daily consumption, and place a bound on the error of estimation. (3 marks)

ii) Estimate the total number of gallons of water used daily during the dry spell and use it to calculate the interval estimate of the population total. (3 marks)

c) Explain the importance of sampling in solving business problems? (8 marks)

QUESTION THREE (20 MARKS)

a) An advertising is interested in estimating the population mean of household in the country that view show X. The country contains two towns, A and B and a rural area. There are 155 households in town A, 62 in town B and 93 in the rural area. The advertising firm finds that obtaining an observation from a rural household costs more than obtaining a response in town A or B. The increase is due to the costs of traveling from one rural household to another. The cost per observation in each town is estimated to be \$9 (that is, $c_1 = c_2 = 9$, and the costs per observation in the rural area to be \$16 (that is, $c_3 = 16$). The stratum standard deviations (approximated by the strata sample variances from a prior survey) are $s_1 = 5, s_2 = 15, s_3 = 10$ 10. Find the overall sample size *n* and the stratum sample sizesn_1 , $n_2$and n_3 that allow the firm to estimate, at minimum cost, the average television-viewing time with a bound on the error of estimation equal to 2 hours. (6 marks)

b) Nature and the target audience define the type of research report in survey sampling. Explain four types of research report that you know. (8 marks)

c) Under what circumstances would you consider cluster sampling (6 marks)

QUESTION FOUR (20 MARKS)

a) Identify and briefly explain any four non-random sampling procedures (8 marks)

b) State and explain four main methods of Data collection a researcher can use in a survey (8 marks)

c) A certain type of circuit board manufactured for installation in computers has 12 microchips per board. During the quality control inspection of ten of these boards, the numbers of defective microchips on each of the ten boards were as follows:

2, 0, 1, 3, 2, 0, 0, 1, 3, 4

Estimate the proportion of defective microchips in the population from which this sample was drawn. (4 marks)

QUESTION FIVE (20 MARKS)

a) A government park charges admission according to car load rather than by person and a park official wants to estimate the number of persons per car for a particular summer holiday. The officer knows from past experience that there should be 400 cars entering the park and wants to sample 80 cars. To obtain an estimate of the variance, he uses repeated systematic sampling with 10 samples of 8 cars each4. Using the data given below, estimate the average number of persons per car and place a bound on the error of estimation. (7 marks)

Data on the number of persons per car (responses are in parentheses)

Random starting point	Second element	Third element	Fourth element	Fifth element	Sixth element	Seventh element	Eighth element
2(3)	52(4)	102(5)	152(3)	202(6)	252(1)	302(4)	352(4)
5(5)	55(3)	105(4)	155(2)	205(4)	255(2)	305(3)	355(4)
7(2)	57(4)	107(6)	157(2)	207(3)	257(2)	307(1)	357(3)
13(6)	63(4)	113(6)	163(7)	213(2)	263(3)	313(2)	363(7)
26(4)	76(5)	126(7)	176(4)	226(2)	276(6)	326(2)	376(6)
31(7)	81(6)	131(4)	181(4)	231(3)	281(6)	331(7)	381(5)
35(3)	85(3)	135(2)	185(3)	235(6)	285(5)	335(6)	385(8)
40(2)	90(6)	140(2)	190(5)	240(5)	290(4)	340(4)	390(5)
45(2)	95(6)	145(3)	195(6)	245(4)	295(4)	345(5)	395(4)
46(6)	96(5)	146(4)	196(6)	246(3)	296(3)	346(5)	396(3)

b) The management of a large utility company is interested in the average amount of time delinquent bills are overdue. A systematic sample will be drawn from an alphabetical list of N= 2500 overdue customer accounts. In a similar survey conducted the previous year the sample variance was found to s^2= 100 days. Determine the sample size required to estimate μ, the average amount of time utility bills are overdue with a bound on the error of estimation of B = 2 days. (3 marks)

c) A systematic bias results from error in the sampling procedures and it cannot be reduced or eliminated by increasing the sample size. What can you say are the causes responsible for these errors? Explain briefly. (10 marks)

Table of Random Numbers

2952	6641	3992	9792	7979	5911	3170	5624
4167	9524	1545	1396	7203	5356	1300	2693
2370	7483	3408	2762	3563	1089	6913	7691
0560	5246	1112	6107	6008	8126	4233	8776
2754	9143	1405	9025	7002	6111	8816	6416

Statistical formulas

Simple random sampling

1. Bound on the error of estimation $B = 2\delta\theta$
2. Sample Variance $S^2 = \frac{1}{n-1} \Sigma_{i=1}^{n} [y_i - \bar{y}]^2$
3. Standard error of the population proportion:

$$S_{\bar{p}} = \sqrt{\left[\frac{N-n}{N}\right]\left[\frac{\bar{p}(1-\bar{p})}{n-1}\right]}$$

4. Sample size for the population mean

$$n = \frac{Ns^2}{N\left(\frac{B^2}{4}\right) + s^2}$$

5. Sample size for the population total

$$n = \frac{Ns^2}{\frac{B^2}{4N} + s^2}$$

6. Sample size for estimating a population proportion

$$n = \frac{N\bar{p}(1-\bar{p})}{N\left(\frac{B^2}{4}\right) + \bar{p}(1-\bar{p})}$$

Stratified Sampling

7. Estimate of the standard error of the mean which is expressed as follows;

$$s_{\bar{y}_{st}} = \sqrt{\frac{1}{N^2}\sum_{h=1}^{H} N_h(N_h - n_h)\frac{s_h^2}{n_h}}$$

8. Point Estimator of the Population Proportion:

$$\bar{p}_{st} = \sum_{h=1}^{H}\left(\frac{N_h}{N}\right)\bar{p}_h$$

9. **S**ample size for population mean

$$n = \frac{(\sum_{h=1}^{H} N_h S_h)^2}{N^2\left(\frac{B^2}{4}\right) + \sum_{h=1}^{H} N_h S_h^2}$$

10. Sample size when estimating a population proportion:

$$n = \frac{\left(\sum_{h=1}^{H} N_h \sqrt{p_h(1-p_h)}\right)^2}{N^2\left(\frac{B^2}{4}\right) + \sum_{h=1}^{H} N_h \left(\sqrt{p_h(1-p_h)}\right)^2}$$

11. $n = \frac{\left(\sum_{h=1}^{H} N_h S_h\right)^2}{N^2\frac{B^2}{4} + \sum_{h=1}^{H} N_h S_h^2}$

12. $n = \frac{\left[\sum_{h=1}^{H} \frac{N_h S_h}{\sqrt{ch}}\right]\left[\sum_{h=1}^{H} N_h S_h \sqrt{ch}\right]}{N^2\frac{B^2}{4} + \sum_{h=1}^{H} N_h S_h^2}$

Systematic sampling

13. Sample size:

$$n = \frac{Ns^2}{(N-1)\left(\frac{B^2}{4}\right) + s^2}$$

14. Estimated variance of the mean $\hat{V}(\hat{\mu}) = \frac{N-n}{N} \cdot \frac{S_{\bar{y}}^2}{n_s}$

15. $S_{\bar{y}}^2 = \frac{\sum_{i=1}^{n_s}(\bar{y}_i - \hat{\mu})^2}{n_s - 1}$

16. Variance of the paired selection model

$$Var(\bar{y}_{sy}) = \frac{1-f}{n^2} \sum_{n=1}^{\frac{n}{2}} (y_{ha} - y_{hb})^2$$

17. Variance of successive difference model

$$Var(\bar{y}_{sy}) = \frac{1-f}{2n(n-1)} \sum_{g=1}^{n-1} (y_g - y_{g+1})^2$$

Cluster sampling

18. Estimated Variance of the Mean

$$\hat{V}(\bar{y}_{cl}) = \left(\frac{N-n}{Nn\bar{M}^2}\right) S_{\bar{y}_{cl}}^2$$

19. $S_{\bar{y}_{cl}}^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y}_{cl} m_i)^2}{n-1}$