

CONFIDENTIAL



UNIVERSITI TUN HUSSEIN ONN MALAYSIA

**FINAL EXAMINATION
SEMESTER I
SESSION 2015/2016**

COURSE NAME : SURVEY AND SAMPLING
METHODS
COURSE CODE : BWB 21103
PROGRAMME : 2 BWQ
EXAMINATION DATE : DECEMBER 2015 / JANUARY
2016
DURATION : 3 HOURS
INSTRUCTION : ANSWER ALL QUESTIONS

THIS QUESTION PAPER CONSISTS OF SEVEN (7) PAGES

CONFIDENTIAL

- Q1**
- (a) Define the concept of census study and give ONE example. (2 marks)
 - (b) Differentiate the concepts of cross-sectional survey and cohort study. Give ONE example of each studies. (4 marks)
 - (c) Differentiate the concepts of accuracy and reliability. (4 marks)
 - (d) Suppose that a survey is being planned for purposes of estimating the average number of hours spent exercising daily by adults (18 years of age or older) living in a certain area at Batu Pahat. Suppose this list consists of five households and the information of hours spent for exercising are shown in **Table Q1(d)**.

Table Q1(d) : Number Of Hours Spent For Exercising By Adults

Household	Number of Hours Spent for Exercising by All Adults
1	2
2	9
3	14
4	40
5	12

- (i) Compute and list down all possible samples if sample of two households are selected. (1 mark)
- (ii) For the first sampling plan, select all possible sample as listed in **Q1(d)(i)** and calculate the expected value of mean and variance of the estimated mean number of hours spent for exercising. (4 marks)
- (iii) For the second sampling plan which consist of household (1 and 2), (1 and 3), (1 and 4), (1 and 5), and (2 and 3), calculate the expected value of mean and variance of the estimated mean number of hours spent for exercising. (4 marks)
- (iv) Which sampling plan is bias? Explain your answer. (1 mark)

- Q2** (a) Define a simple random sampling with replacement and without replacement. (2 marks)
- (b) Give two methods on how to take a simple random sample. (2 marks)
- (c) A community in the San Francisco Bay area consists of approximately 50000 persons, of whom approximately 40% are Caucasians, 25% are African American, 20% are Hispanic, and 15% are Asian. It is desired to estimate in this community, the proportion of persons who are not covered by some form of health insurance. One would like to be 95% certain that this estimate is within 15% of true proportion, which is believed to lie somewhere between 10% and 20% of the total population. Assuming simple random sampling, how large a sample is needed for 10% proportion. (4 marks)
- (d) A community within a city contains 3000 households and 10,000 persons. For purposes of planning a community satellite to the local health department, it is desired to estimate the total number of physician visits made during a calendar year by members of the community. For this information to be useful, it should be accurate to within 10% of the true value. A small pilot survey of 10 households, conducted for purposes of gathering preliminary information, yielded the accompanying data on physician visits made during the previous calendar year (**Table Q2(d)**). Use 99.7% confidence level of estimated.

Table Q2(d) : Total number of physician visits during previous year

Household	Total number of physician visits
1	12
2	27
3	16
4	17
5	1
6	21
7	34
8	12
9	24
10	30

- (i) Using these data as preliminary information, identify the sample size needed to meet the specifications of the survey. (10 marks)

- (ii) If one person want to take 250 sample size, examine whether it is enough compare to the result of sample size form Question **Q2 (d)(i)**? Why?

(2 marks)

- Q3** (a) State the reason why variance of estimate is important? (2 marks)
- (b) State the advantage of using repeated systematic sampling. (2 marks)
- (c) Suppose that a study is planned of the level of the pesticide dieldrin, which is believed to be a carcinogen, in a 7.5 mile stretch of a particular river. To assure representativeness, a map of the river is divided into 36 zones, (see the **Figure Q3(c)** on page 8) and a 1 in 4 systematic sample of these zones is to be selected. Water samples will be drawn by taking a boat out to the geographic center of the designated zone, and drawing a grab sample of water from a depth of several centimeters below the surface level. The levels of dieldrin, in micrograms per liter, for each of these zones are shown in parentheses.
- (i) Using systematic sampling of 1 in 4, show the list of sample of zones that should be selected. (2 marks)
- (ii) Calculate the 90% confidence interval for the average level of dieldrin in this stretch of the river. (12 marks)
- (iii) Discuss the advantages that you can identify for this method of sampling the river over simple random sampling. (2 marks)

- Q4** (a) What is strata? (2 marks)
- (b) Why we used stratified sampling? (2 marks)
- (c) A sample survey of households in a community containing 1500 households is to be conducted for the purpose of determining the total number of persons over 18 years of age in the community who have one or more permanent teeth (other than third molars) missing. Since this variable is thought to be correlated with age and income, the strata shown in the accompanying **Table Q4(c)** are formed by using available population data. A stratified random sample of 100 families is to be taken.

Table Q4(c) : Population Data

Variable	Stratum			
	1	2	3	4
Age				
Mean	30	32	25	27
Standard deviation	15	15	10	10
Annual family income (x \$1000)				
Mean	15	7	15	8
Standard deviation	5	3	3	2
No. of families	300	500	100	600

- (i) Compute the number of families to be taken from each stratum if proportional allocation is used. (4 marks)
- (ii) Calculate the number of families to be taken from each stratum if optimal allocation is used based on annual family income. (6 marks)
- (iii) Calculate the number of families to be taken from each stratum if optimal allocation is used based on age. (6 marks)

Q5 (a) Why we used cluster sampling? (2 marks)

(b) Suppose that the elementary schools in a city are grouped into 30 school districts, with each school district containing four schools. Suppose that a simple one-stage cluster sample of three school districts is taken for purposes of estimating the number of school children in the city who are color-blind (as measured by a standard test), and that the accompanying data are obtained from this sample (**Table Q5(b)**).

Table Q5(b) : The number of school children in the city who are color-blind

Sample School District	School	No. Of Children	No. Of Color-Blind Children
1	1	130	2
	2	150	3
	3	160	3
	4	120	5
2	1	110	2
	2	120	4
	3	100	0
	4	120	1
3	1	89	4
	2	130	2
	3	100	0
	4	150	2

(i) Calculate and obtain a 95% confidence interval for the total number of color-blind children. (9 marks)

(ii) Calculate and obtain a 95% confidence interval for the proportion of all children who are color-blind. (9 marks)

- END OF QUESTION -

FINAL EXAMINATION

SEMESTER/SESSION: SEM I/2015/2016

PROGRAMME : 2 BWQ

COURSE NAME : SURVEY AND SAMPLING METHODS

COURSE CODE: BWB21103

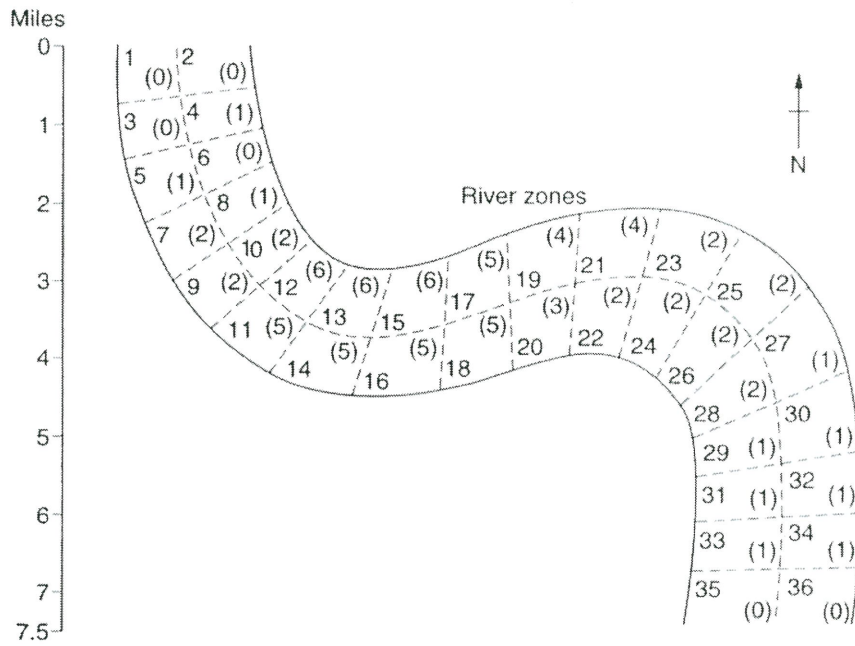


Figure Q3(c): River divided into zones