Universiti Tun Hussein Onn Malaysia

UNIVERSITI TUN HUSSEIN ONN MALAYSIA

FINAL EXAMINATION
SEMESTER I
SESSION 2013/2014

COURSE NAME	:	SURVEY AND SAMPLING METHODS
COURSE CODE	:	BWB 21103
PROGRAMME	:	2 BWQ
EXAMINATION DATE	:	DECEMBER 2013/ JANUARY 2014
DURATION	:	3 HOURS
INSTRUCTION	:	ANSWER **ALL** QUESTIONS

THIS QUESTION PAPER CONSISTS OF **EIGHT (8)** PAGES

Q1 (a) Define the elementary unit and give one example.

(2 marks)

(b) Differentiate the concepts of probability and non-probability sampling.

(4 marks)

(c) Differentiate the concepts of validity and accuracy.

(4 marks)

(d) Let N=6 and n=3. For purpose of studying sampling distribution, we assume that all population values are known. In this study, two sampling plans are proposed. In Sampling Plan 1, eight possible samples are chosen. Meanwhile, in Sampling Plan 2, only three possible samples are chosen.

Population	**Number of Element in Population**
1	99
2	104
3	156
4	130
5	189
6	173

Sampling Plan 1: {1, 3, 5} {1, 3, 6} {1, 4, 5} {1, 4, 6} {2, 3, 5} {2, 3, 6} {2, 4, 5} {2, 4, 6}
Sampling Plan 2: {1, 4, 6} {2, 3, 5} {1, 3, 5}

(i) Compute the value of population mean.

(1 marks)

(ii) For each sampling plan, compute the value of estimated mean and bias.

(8 marks)

(iii) Which sampling plan do you think is better?

(1 marks)

Q2 (a) Define a simple random sampling.

(2 marks)

(b) Give two examples of simple random sampling.

(2 marks)

(c) List all possible simple random samples of size $n=2$ that can be selected from the population $\{0, 1, 2, 3, 4\}$.

(4 marks)

(d) Medical researchers interested in investigating the average body temperature under normal conditions between men and women. Random sample from the data $N=500$ for each male and female are presented in the Table **Q2(e).**

Table Q2(e): Average body temperature under normal conditions between men and women

Body Temperatures of Male		**Body Temperatures of Female**	
96.9	97.9	97.8	98.6
97.4	98.0	98.0	98.8
97.5	98.1	98.2	98.8
97.8	98.6	98.2	99.2
97.8	98.8	98.2	99.4

(i) Compute the average of body temperatures for each male and female.

(4 marks)

(ii) Estimate the average of body temperatures for male using 95% confidence interval.

(3 marks)

(iii) Estimate the average of body temperatures for female using 95% confidence interval.

(3 marks)

(iv) In which of the confidence intervals between male and female do you have the least "confidence"?

(2 marks)

Q3 (a) Estimated means, totals, and proportions are always unbiased under systematic sampling. (True/False)

(2 marks)

(b) State the advantage of using systematic sampling method compare to simple random sampling.

(2 marks)

(c) The five possible samples are listed in Table **Q3(c)** along with estimated mean, total, and proportion of the population paramater using the systematic sampling method. Each of the five possible samples has the same probability to be selected.

Table Q3(c) : Estimates mean, total, and proportion of the population paramater

Sample	Estimated Mean	Estimated Total	Estimated Proportion
1	2.5	64	0.3
2	4.7	121	0.7
3	1.5	36	0.4
4	9.1	231	0.8
5	7.5	184	0.5

(i) Compute the mean of the sampling distribution of the estimated mean.

(2 marks)

(ii) Compute the mean of the sampling distribution of the estimated total.

(2 marks)

(iii) Compute the mean of the sampling distribution of the estimated proportion.

(2 marks)

(iv) Compute the value of standard error of each estimated mean, total, and proportion.

(9 marks)

(v) Result from simple random sampling showed that the standard errors of estimated mean, total, and proportion are 3.36, 84.11, and 0.20, respectively. Compare with the result obtained in (iv), which type of sampling method is better?

(1 mark)

Q4 (a) What is a stratified random sample?

(2 marks)

(b) State the advantage of using stratified sampling method compare to simple random sampling.

(2 marks)

(c) From a simple random serological of 1000 runners selected from 10,000 who completed the 1996 Chicago Marathon, 35 runners were found to be positive for steroids and other performance-enhancing drugs. When the sample was categorized by completion time, the results were listed in Table **Q4(c)**.

Table Q4(c): Percentage of positive for steroids and other performance-enhancing drugs

Completion Time (Hour)	No. In Sample	No. Positive for Drug	Percentage Positive
Under 2.5	100	26	26
2.5 – 4.0	500	8	1.6
Over 4.0	400	4	1.0
Total	1000	38	3.80

(i) What is the standard error of the estimated proportion positive?

(2 marks)

(ii) From inspection (without making calculations) of the rates given above, would you feel that stratification may have resulted in a substantially better estimate? Why and why not?

(2 marks)

(d) The following data are available for 1989 from 6 health maintenance organizations (HMOs) in a medium-size city. Table **Q4 (c)** shows the number of personnel providing patient cares and number of patient encounters during 1989 among six HMOs.

As a time-sharing and cost-sharing device, it is proposed that a sample of two of these HMOs be taken for purposes of estimating the total number of patient encounters in 1989 among these six HMOs.

Table Q4 (c): Number of personnel providing patient care and number of patient encounters during 1989 among six HMOs

HMO	No. Of Physicians Providing Patient Care	No. Of Patient Encounter
1	10	22 000
2	6	14 000
3	4	10 200
4	30	70 000
5	7	15 000
6	3	5 000

(i) Enumerate all simple random samples of size 2 and estimate the total number of patient encounters in 1989 among the six HMOs.

(5 marks)

(ii) What is the standard error of the estimated total from the sample design in (i).

(3 marks)

(iii) Group HMOs 1, 2, 3, 5, and 6 into 1 stratum and HMO 4 into 2 stratum. From these two strata, enumerate all stratified random samples of size 2. What is the standard error of the estimated total as obtained by this sampling plan?

(4 marks)

Q5 (a) State two disadvantage of using cluster sampling.

(2 marks)

(b) A simple one-stage cluster sample was taken of 10 hospitals in a mid-western state from a population of 33 hospitals that received state and federal funds to upgrade their emergency medical service. The number of patients hospitalized for trauma conditions and the number of discharged dead are shown in the Table **Q5(b)**.

Table Q5(b) : The number of patients hospitalized for trauma conditions and the number discharged dead

Hospital	Total No. Of Patients Hospitalized for Trauma Conditions	Total No. Of Discharged Dead Among All Patients Hospitalized for Trauma Conditions
1	560	4
2	190	4
3	260	2
4	370	4
5	190	4
6	130	0
7	170	9
8	170	2
9	60	0
10	110	1

(i) Estimate and give a 95% confidence interval for the total number of patients discharged dead among all persons hospitalized for trauma condition.

(9 marks)

(ii) Estimate and give a 95% confidence interval for the total number of patients hospitalized for trauma condition among the 33 hospitals.

(9 marks)

- END OF QUESTION -

FINAL EXAMINATION

SEMESTER/SESSION: SEM I/2013/2014 PROGRAMME : 2 BWQ
COURSE NAME : SURVEY AND SAMPLING METHODS COURSE CODE: BWB21103

Chapter 2: $\bar{X} = \frac{\sum_{i=1}^{N} X_i}{N}$, $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$, $B(\hat{d}) = E(\hat{d}) - d$

Chapter 3: $\bar{x} \pm z_{1-\left(\frac{\alpha}{2}\right)} \sqrt{\frac{N-n}{N}} \left(\frac{s_x}{\sqrt{n}}\right)$, $z_{0.975} = 1.96$

$$s_x^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

Chapter 4: $E(\bar{x}) = \frac{\sum_{i=1}^{n} x_i}{n}$, $E(x') = \frac{\sum_{i=1}^{n} x_i}{n}$, $E(p_y) = \frac{\sum_{i=1}^{n} y_i}{n}$

$$SE(\bar{x}) = \sqrt{\frac{1}{n}\sum[\bar{x} - E(\bar{x})]^2} \text{ , } SE(x') = \sqrt{\frac{1}{n}\sum[x' - E(x')]^2}$$

$$SE(p_y) = \sqrt{\frac{1}{n}\sum[p_y - E(p_y)]^2}$$

Chapter 5: $\widehat{SE}(p_{y,str}) = \sqrt{\sum_{h=1}^{L}\left(\frac{N_h}{N}\right)^2 \left(\frac{p_{hy}(1-p_{hy})}{n_h - 1}\right)\left(\frac{N_h - n_h}{N_h}\right)}$, $x' = \left(\frac{N}{n}\right)x$

$$E(x'_{str}) = \sum_{h=1}^{L} x'_h \text{ , } \widehat{SE}(x'_{str}) = \sqrt{\sum_{h=1}^{L}\left(\frac{N_h^2 s_{hx}^2}{n_h}\right)\left(\frac{N_h - n_h}{N_h}\right)}$$

Chapter 6: $x'_{clu} = \left(\frac{M}{m}\right)x$, $\widehat{SE}(x'_{clu}) = \left(\frac{M}{\sqrt{m}}\right)\hat{\sigma}_{1x}\sqrt{\frac{M-m}{M-1}}$,

$$\hat{\sigma}_{1x} = \sqrt{\left[\frac{\sum_{i-1}^{m}(x_i - \bar{x}_{clu})^2}{m-1}\right]}\sqrt{\frac{M-1}{M}} \text{ ,}$$

$$x'_{clu} - 1.96 \times \widehat{SE}(x'_{clu}) \leq X \leq x'_{clu} + 1.96 \times \widehat{SE}(x'_{clu})$$