INSTRUCTION

1. This book script contains FIVE (5) questions in NINE (9) printed pages excluding the cover page.

2. This book script also contains THREE (3) ATTACHMENT printed pages from 10 – 12

3. Answer ALL questions in the spaces provided.
QUESTION 1 (15 MARKS)

a) What is the definition of sampling? (2 marks)

b) State three (3) types of probability sampling and briefly describe one (1) of them. (5 marks)

c) A researcher is interested in investigating the savings habits of blue-collar workers in the plastic industry throughout in City PQR. He selected 100 workers and study on the sample. What are should be a population, sample and element for this investigation? (3 marks)

d) The ABC Co. produces special vacuum cleaners that can be conveniently used to clean the inside of cars. About a thousand of these are produced every month with serial numbers attached to them and stored serially in a stock room. Once a month an inspector comes and does a quality control check on fifty of the units. When he certifies them to be of acceptable quality, the units are released from the stock room for sale. The production and sales managers, however, are not satisfied with the quality control check since, quite often, many of the units sold are returned by customers for various types of defects. What would be the most useful sampling plan to test the fifty units? Why? (5 marks)
QUESTION 2 (18 MARKS)

An agricultural economist wants to compare cow manure and turkey dung as fertilizers. Historically, farmers had used cow manure on their cornfields. Recently, a major turkey farmer offered to sell composted turkey dung at a favorable price. The farmers decided that they would use this new fertilizer only if there was strong evidence that productivity increased over the productivity that occurred with cow manure. The agricultural economist implemented an experiment of cow manure size $n_1 = 25$ from the population size $160$ with mean productivity is 100. Turkey dung was applied on $n_2 = 36$ from $N_2 = 200$ with the mean of 115. It is known that from the past experience the cow manure and turkey dung variance productivity is 4 and 6.25 respectively.

a) Estimate the average productivity for cow manure and turkey dung? (2 marks)

b) Place bound on the error of estimation of cow manure. (4 marks)

c) Place bound the error of estimation of turkey dung. (4 marks)
d) How many sample size of cow manure should the economist select if the population variance is the same as its estimator?  
(4 marks)

e) How many sample size of turkey dung, \( n_2 \) should the economist select if the population variance is the same as its estimator?  
(4 marks)

**QUESTION 3 (15 MARKS)**

a) Why the stratified sampling is more efficient than simple random sampling in such situation?  
(3 marks)
b) To celebrate their first anniversary, Dollah decided to buy a pair of diamond earrings for his wife, Jenab. He shown 6 pairs with marquise gems weighing approximately 2 carats per pair from $N_j = 25$. Because of difference in the colors and quantities of the stones, the prices varied from set to set. The average price was RM2990 with a sample standard deviation of RM370. He also looked at 9 pairs with pear-shaped stones of the same two-carat approximate weight from $N_2 = 20$. These earrings had an average price of RM3065 with a sample standard deviation of RM805.

i) Estimate the average price of the earrings then calculate the variances of the estimation.

(4 marks)
ii) If sample size \( n = 20 \) were given to Dollah to take a look, determine the sample size for each earrings, \( n_1 \) and \( n_2 \) using:

a. proportional allocation.  

(4 marks)

b. optimal allocation if the standard deviation for population is same as the sample.  

(4 marks)
QUESTION 4 (16 MARKS)

A researcher wants to do a survey regarding to the opinion of households in Bandar W. He wishes to select 35 households from a total population of 260 houses in the town. Firstly he assigns a number from 1 to 260 to the population frame. He used a random number to select the sample and found to be the 8th sample as the first sample. Then he select the 16th, 24th, 31st, 39th, 47th, 54th... 254th sample to complete 35 samples should be selected.

a) What is the type of sampling survey that the researcher used? (1 marks)

b) Calculate the sampling interval, k? (2 marks)

c) Suppose the objective of the survey is to determine whether households were agree or not if the motor cross circuit should be build in the town. From the survey, the researcher have the following information:

\[ \sum_{i=1}^{35} x_i = 20 \]  where \( x \) is the number house holds that said not agree. Estimate the proportion households agreed the circuit should be build in the town. Then calculate the variance of the estimation. (4 marks)

d) The manager of a large utility is interested in the average amount of time delinquent bills are over due. A systematic sample will be drawn from an alphabetic list of \( N \) overdue customers accounts. Table 1 shows the data that the manager recorded.

<table>
<thead>
<tr>
<th>Bill sampled</th>
<th>Overdue time (day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>17</td>
<td>5</td>
</tr>
<tr>
<td>27</td>
<td>2</td>
</tr>
<tr>
<td>( \vdots )</td>
<td>( \vdots )</td>
</tr>
<tr>
<td>( n )</td>
<td>4</td>
</tr>
</tbody>
</table>

\[
\sum_{i=1}^{250} y_i = 5400
\]
i) Find the sampling interval, \( k \). Then calculate the value of \( n \).  
(3 marks)

ii) Estimate of the population mean.  
(2 marks)

iii) In a similar survey conducted the previous year, the variance was found to be 100 days. Determine the sample size required to estimate average amount of time utility bills are overdue, with a bound on the error of estimation of two days.  
(4 marks)

**QUESTION 5 (16 MARKS)**

A manufacturer of band saws wants to estimate the average repair cost per month for the saws he has sold to certain industries. He cannot obtain a repair for each saw, but he can obtain the total amount spent for saw repairs and the number of saws owned by each industry. Thus, he decides to use cluster sampling, with each industry as a cluster. The manufacturer selects a simple random sample of \( n \) from 96 industries he services. Given the accompanying information, calculate the following questions:

\[
\sum_{i=1}^{20} m_i = 130 \quad \sum_{i=1}^{20} y_i = 2565 \quad \sum (y_i - \bar{y}m_i)^2 = 16065.6399
\]

Where \( m_i \) = the number of elements in cluster \( i \).  
\( y_i \) = the total of all observation in the \( i^{th} \) cluster.
a. What is the value of $n$? (1 mark)

b. Estimate the average number of saws from the sample. (2 marks)

c. Estimate the average repair cost per saw for the past month, and place a bound on the error of estimation. (hint: use the average cluster size for the population same as its estimator) (7 marks)
d. Suppose the manufacturer wants to estimate the average repair cost per saw for the next month, how many clusters should he select for this sample if he wants the bound on the error of estimation to be less than RM2?

(6 marks)
APPENDIX

List of Formulae

Simple Random Sampling

1. Estimator of the population mean, $\mu$.

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$$

2. Estimated variance of $\bar{y}$

$$\hat{\sigma}_{\bar{y}}^2 = \frac{s^2}{n} \left( \frac{N-n}{N} \right)$$

3. Bound on the error of estimator

$$B = 2 \hat{\sigma}_{\bar{y}}$$

4. Sample size required to estimate $\mu$ with a bound on the error of estimation $B$

$$n = \frac{N\sigma^2}{(N-1)D + \sigma^2}$$

$$D = \frac{B^2}{4} = 0.1350$$

Stratified Sampling

1. Estimator of the population mean

$$\bar{y}_u = \frac{1}{N} \sum_{i=1}^{k} N_i \bar{y}_i$$

2. Estimated variance of $\bar{y}_u$

$$\hat{\sigma}_{\bar{y}_u}^2 = \frac{1}{N^2} \sum_{i=1}^{k} \left( N_i \left( \frac{N_i - n_i}{N_i} \right) \left( \frac{s^2_i}{n_i} \right) \right)$$
3. Selecting sample size using proportional allocation

\[ n_i = \frac{N_i}{N} (n) \]

4. Selecting sample size using optimal allocation

\[ n_i = \frac{N_i \sigma_i}{\sum_{i=1}^{L} (N_i \sigma_i)} (n) \]

**Systematic Sampling**

1. Estimator of the population mean, \( \mu \)

\[ \hat{\mu}_y = \frac{\sum_{i=1}^{n} y_i}{n} \]

2. Estimator of the population proportion, \( p \)

\[ \hat{p}_y = \bar{y}_y = \frac{\sum_{i=1}^{n} y_i}{n} \]

3. Sample size required to estimate \( \mu \) with a bound on the error of estimation \( B \)

\[ n = \frac{N \sigma^2}{(N - 1)D + \sigma^2} \quad \quad D = \frac{B^2}{4} \]

**Cluster Sampling**

1. Average number of elements for the sample

\[ \bar{m} = \frac{\sum_{i=1}^{n} m_i}{n} \]
2. Estimator of the population mean $\mu$

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} m_i}$$

3. Estimated variance of $\bar{y}$

$$\hat{\sigma}_\bar{y}^2 = \left( \frac{N - n}{Nnm^2} \right)s_r^2$$

$$s_r^2 = \frac{\sum_{i=1}^{n} (y_i - \bar{y}m_i)^2}{n-1}$$

4. Sample size required to estimate mean with a bound on the error of estimation B

$$n = \frac{N\sigma_r^2}{ND + \sigma_r^2}$$

$$D = \frac{B^2\bar{M}^2}{4}$$