

EXAMINATIONS OF THE HONG KONG STATISTICAL SOCIETY



ORDINARY CERTIFICATE IN STATISTICS, 2004

Paper I

Time Allowed: Three Hours

*Candidates may attempt **all** the questions.*

*The number of marks allotted to each question or part-question is shown in brackets.*

*The total for the whole paper is 100.*

*A pass may be obtained by scoring at least 50 marks.*

*Graph paper and Official tables are provided.*

*Candidates may use silent, cordless, non-programmable electronic calculators.*

*Where a calculator is used the **method** of calculation should be stated in full.*

This examination paper consists of 4 pages, **each printed on one side only**.

This front cover is page 1.

Question 1 starts on page 2.

There are 8 questions altogether in the paper.

1. What are the advantages and disadvantages of quota sampling, as contrasted with random sampling? (10)

2. A survey about leisure activities is planned in a community with about 3000 households. A sample of households is to be taken and it is desirable that responses are obtained from households of different sizes and types.

(i) A list of addresses for delivery of mail is available. Some addresses might contain several households. Describe three potential problems in using a list of addresses to obtain a sample of households and suggest how these might be overcome. (6)

(ii) One method of obtaining a sample is to ask interviewers to collect information, with the choice of household left to the interviewer. What instructions would you give interviewers to help ensure that the sample obtained is representative of households of different sizes and types? There is no need to repeat any suggestions you made in part (i) as regards obtaining a sample of households from a list of addresses. (6)

3. A pharmacist wishes to keep details of regular customers on a database to help ensure that customers do not take medication which might have adverse effects. Design a database suitable for this purpose. It should contain data about the following customer attributes:

- Name and address, including postcode/zipcode if there is one
- A contact telephone number
- Name of doctor
- Up to two allergies of the customer
- Medication prescribed with dates customer started and finished taking each one.

Your answer should list the fields, giving appropriate field names, together with field types and suitable widths. (12)

4. A survey is to be carried out among employees who work in a large city, to study various aspects of their travel to work. A sample of employees will be asked to complete a questionnaire. Information on where they work, how long it takes to travel from home to work, how far they travel, what method(s) of transport they usually use, how much (each part of) the journey costs, and whether there are any disadvantages in their present method(s) of travel will be sought.

Draft a short questionnaire to obtain this information, using multiple choice questions where appropriate.

(12)

Suggest three other questions that might be useful for someone seeking to improve the experience of employees in their travel to work.

(3)

5. In dietary studies, subjects are sometimes asked to keep a diary for a limited period of time in which they record what foods and drinks they have consumed, and when. Outline the advantages and disadvantages of collecting information using a diary rather than relying solely on a questionnaire.

(10)

6. Discuss the issues involved in advising someone who plans to do a survey on what size of sample they should take. Assume that the method of sampling is to be simple random.

(9)

7. A survey of farms classified into three size groups (Small, Medium and Large) is to be undertaken in a particular region. The cost of sampling each farm is the same, regardless of size, and the budget is sufficient for sampling 40 farms. The main objective of the survey is the estimation of the area under crops. The table shows the number of farms in each size group, and values of the standard deviation of the area (in hectares) under crops obtained from a sample survey undertaken last year.

<i>Size group</i>	<i>Number of farms</i>	<i>SD of area (last year)</i>
Small	203	6.4
Medium	115	11.6
Large	82	27.3

- (i) Find the numbers of farms to be sampled from each size group if the sample sizes are to be in the same ratio as the group sizes in the population of farms. (4)
- (ii) Calculate the numbers of farms to be sampled from each size group if the sample sizes are proportional to "population size times the standard deviation" in each group. (6)
- (iii) Comment on the methods of sampling described in parts (i) and (ii), drawing attention to any advantages and disadvantages. (4)
8. A library has  $N$  books (where  $N$  is known), arranged in bookcases, each of which has several shelves. A sample of about  $n$  books is to be taken in order to estimate the average age of books.
- (i) Describe in detail how each of the following might be implemented:
- (a) systematic sampling; (4)
- (b) cluster sampling. (4)
- Outline the advantages and disadvantages of each of these methods in this application. (5)
- (ii) Would stratified sampling be a suitable method in this application? Justify your answer. (5)