EXAMINATIONS OF THE HONG KONG STATISTICAL SOCIETY

ORDINARY CERTIFICATE IN STATISTICS, 2010

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 calculators in accordance with the regulations published in the Society's "Guide to Examinations" (document Ex1).

This examination paper consists of 5 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.   (i)   Briefly explain the concepts of *selection bias* and *response rate* and why they might give rise to errors or difficulties in the interpretation of survey results.

(4)

(ii)   A chain of self-service restaurants has outlets throughout a country. A survey of customers' opinions of the restaurants is planned and the following two methods for conducting the survey in any outlet have been suggested.

A.   Leave paper questionnaires near to the pay desks with an invitation to customers to complete one.

B.   Arrange for a member of staff to interview customers at a random selection of tables as determined in the office.

Discuss how the concepts of selection bias and response rate might relate to the methods A and B above, noting the advantages and disadvantages of each suggestion.

(8)

2.   The chain of restaurants described in question 1 is in a country that is divided into eight regions, each containing both large and small urban areas. There are between ten and fifteen outlets in every region and these outlets are in both small and large urban areas. Small urban areas have only one outlet, but large urban areas have between two and five outlets.

(i)   Devise a multi-stage sampling method to select outlets in which customers are to be surveyed. State the technical method used at each stage of the scheme.

(4)

(ii)   There are between 6 and 15 tables at outlets. Describe two methods for selecting tables at which customers will be interviewed, commenting on the sample size.

(6)

3. A survey is being planned to investigate stress at work among teachers in colleges where the students are aged between 16 and 19. Design a self-completion questionnaire, with instructions on how to complete it, and a suitable covering letter, to obtain the following information.

The teacher's sex, age group, subject taught and position held in the college

What arrangements there are for flexible working

The extent to which the teacher
- feels that the working hours are excessively long
- feels that his or her work makes a valuable contribution to society
- feels valued at work
- feels stressed at work
- feels the pay is adequate

How likely the teacher is to leave the sector during the next year

Your questionnaire should include at least one closed question, at least one open question, and at least one question to which the response is to be given on a rating scale. Indicate which questions are of each type.

(13)

4. There are three colleges, A, B and C for students aged 16–19 in a city. The table below shows the number of teachers (N_i) in each college, the estimated cost (c_i) in £ of travelling to and interviewing a teacher at each college, and the standard deviation (s_i) of the number of years teachers have been at the college.

College	Number of teachers N_i	Cost of sampling one teacher c_i	SD of years at college s_i
A	307	5	7.5
B	200	10	2.8
C	103	7	5.3

(i) Using a uniform sampling fraction, calculate the required number of teachers to sample from each of the colleges to achieve an overall sample of about 150 teachers. What is the total cost of sampling teachers using these sample sizes?

(4)

(ii) Among several objectives, it is required to estimate the mean number of years teachers have spent at the colleges. The optimum allocation method of choosing sample sizes n_i when estimating a mean minimises the variance of the estimate, and, in order to achieve this minimum, the n_i must be taken proportional to $N_i s_i / \sqrt{c_i}$. Calculate the number of teachers required in the sample from each college using this method if the total cost of sampling teachers is to be no more than £1050.

(10)

(iii) Discuss, with reasons, which of the two methods of calculating sample sizes used in parts (i) and (ii) you prefer here.

(4)

5. A researcher is interested in finding out whether the views of teachers concerning stress at work at a large college are different one, two and five years after joining the college. State the advantages and disadvantages of collecting this information by means of

(i) a longitudinal study of a sample of teachers who joined the college last year,

(7)

(ii) sample surveys of teachers who joined the college one, two and five years ago.

(6)

6. Explain what is meant by a *pilot survey* and discuss in detail reasons for doing one.

(11)

7. A survey is being planned in a community with about 5000 households. A sample of households is to be taken and it is desirable that responses be 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 problems 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.

   (5)

8. It has been suggested that shoppers are more likely to buy goods that are placed at eye height than those which are low down on the shelves. To investigate this suggestion, the owner of a small supermarket has decided to place the cheaper brands of tinned goods at eye height for two weeks with the more expensive goods at the lowest position, and in the following two weeks to reverse these positions. A researcher has been asked to observe shoppers' behaviour at these shelves. What advice would you give on what to observe, how to select shoppers to observe, and how to record the information? Mention any practical problems you foresee in the collection of data.

(12)