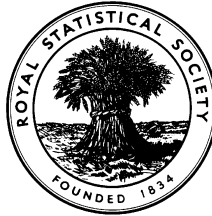**EXAMINATIONS OF THE ROYAL STATISTICAL SOCIETY**

(formerly the Examinations of the Institute of Statisticians)

ORDINARY CERTIFICATE IN STATISTICS, 2006

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.

Questions 1 to 3 carry 41 marks in total, and relate to the following situation.

Farmers in a developing country are to be interviewed as part of a project investigating the control of a pest XX which attacks a particular food crop. Farms are small and family run, and some are a considerable distance from the nearest village. No sample frames or official data are available for the population of the country, and there are no reliable large-scale maps, but basic geographical and agricultural data are available on the 50 villages in which public meetings, to generate awareness of the project, are to be held. Fifteen of these villages are to be chosen before the public meetings to provide a sample of farmers to interview. The numbers of farmers required in each of these villages are fixed beforehand.

Two methods, A and B, are proposed for selecting the farmers from the 15 chosen villages. In method A, a random sample of farmers is to be selected from those attending the public meetings in the selected villages, and the farmers selected are to be interviewed at the end of the meeting. In method B, interviewers are to travel round the selected villages and nearby areas and to interview farmers at systematic samples of farms.

1. (i) What general criteria should be considered when selecting the 15 villages in which farmers are to be interviewed?

(3)

(ii) Discuss any advantages and disadvantages of method A. Where possible suggest ways in which the disadvantages might be overcome.

(7)

(ii) Discuss any advantages and disadvantages of method B. Where possible suggest ways in which the disadvantages might be overcome.

(6)

2. Draft a questionnaire, which will be completed by an interviewer, to obtain the following information from farmers. Your questionnaire should include detailed instructions to the interviewer.

- Name, age and highest educational level of all members of the household, including the farmer

- Relationship of every member of the household to the farmer

- The number of hours a week that each member of the household, including the farmer, works on the farm at the time of year when the interview is held

- The types of livestock owned by the farmer, and the numbers of each type

- The crops grown by the farmer and the area under each crop

- Whether the farmer thinks the risk of pest XX is high, medium, or low

- What measures the farmer takes to control pest XX

(13)

Turn over

3.   The information obtained from farmers, as detailed in question 2, is to be stored in electronic form.

(i)   It is thought that farmers' households have between one and five members (including the farmer). What problems might occur in storing the information relating to the different household members in electronic form as a result of this variation in household sizes?

(4)

(ii)   The information relating to livestock, crops, pest risk and pest control is to be stored in a separate file, in order to produce tables and diagrams for each village. Explain how the responses to the questionnaire will help determine what variables are required. State the *type* of each variable, and, where appropriate, indicate the coding system to be used.

(8)

Question 4 carries 14 marks in total.

4.   A psychologist is studying the behaviour of children while they attend nursery school. One of the main objectives of the study is the estimation of the mean number of minutes per hour that children interact with one another. The psychologist plans to observe individual children for one hour each in three schools and to record the number of minutes, m, in the hour that the child under observation interacts with other children. The cost of observing a child is the same in each school, and the budget is sufficient to observe 20 children altogether. The table shows the number of children at each school, and the standard deviations of m obtained from a preliminary study in these schools.

Nursery school	Number of children	SD of m (preliminary study)
1	25	10.2
2	30	7.7
3	37	8.9

(i)   How many children should be sampled from each nursery school, if the sample sizes are to be in approximately the same ratio as the numbers of children in the three schools?

(4)

(ii)   How many children should be sampled from each school if the sample sizes are to be approximately proportional to "population size times the standard deviation of m" in each school?

(6)

(iii)   How might you choose between the two methods of sampling described in parts (i) and (ii) here?

(4)

3

Questions 5 to 8 carry 45 marks in total, and relate to the following situation.

A survey research organisation has been commissioned to undertake a study of how people in a particular industrialised country use their time. The country is divided into several regions, each of which is divided into a large number of administrative districts, with each district classified as urban or rural. For each district, data on the total number of adults, the numbers in full-time employment, and the numbers in part-time employment, are available for the previous year.

5. It is required to make comparisons of time use for adults in the three groups of full-time employment, part-time employment and the remainder (neither in full-time nor part-time employment), and by whether they live in rural or urban districts.

   (i) Describe how each of stratified sampling and cluster sampling might be implemented to choose a sample of administrative districts from which a sample of adults will be interviewed face to face, so that these comparisons can be made. (Consideration of choice of sample size is not required.)

   (5)

   (ii) Which of the two methods of part (i) do you prefer here, and why?

   (3)

6. Two ways of selecting a sample of adults to be interviewed about their use of time are under consideration.

   • Simple random sampling from a register of voters that was compiled a year ago and which shows names and addresses

   • Quota sampling

   Describe the non-sampling errors that might occur in each of these two selection methods in this application.

   (8)

4

7.   A sample of adults is to be asked to record their activities for each 10-minute period during one week. For this they will be given a diary in which these 10-minute periods are listed for every day. Respondents are asked to record both the primary activity and, where applicable, a secondary activity. (For example, the primary activity might be driving, with a secondary activity of listening to the vehicle's radio.)

    (i)   What are the advantages of collecting time-use data by a diary, over collecting this information by a questionnaire?

(5)

    (ii)   What are the disadvantages of collecting time-use data by a diary, as opposed to collecting this information by a questionnaire?

(7)

8.   The time use of children aged between 5 and 16 is also to be studied.

    (i)   It has been suggested that, in the case of younger children, an observational study would be a suitable method of collecting information. Discuss the practicalities of this method here, and whether or not reliable results are likely to be obtained.

(6)

    (ii)   The firm conducting the survey decided to publish a questionnaire for older children on the web.

        (a)   Suggest and comment on two ways in which a sample of children might be obtained.

(6)

        (b)   Explain why it is particularly important to pilot the questionnaire in this case.

(5)