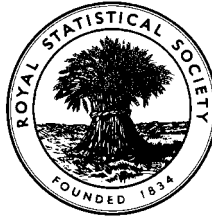# EXAMINATIONS OF THE ROYAL STATISTICAL SOCIETY
*(formerly the Examinations of the Institute of Statisticians)*



## ORDINARY CERTIFICATE IN STATISTICS, 1996

## Paper I

## Time Allowed: Three Hours

*There is no restriction on the number of questions that a candidate may attempt,
nor on the order in which they are attempted.  Candidates are not required
to answer all the questions: they should answer as many as they can.*

*The number of marks allotted to each 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.*

1. *Social Trends* is an important annual compendium of social statistics for the United Kingdom and there are similar annual publications in many other countries. Describe five specific topics which receive coverage in either *Social Trends* or another named non-UK source of social statistics. For each topic separately, give one example of a statistic presented and one way in which that information may be used. (15)

2. (a) Define the term *quota sampling*. State one advantage and one disadvantage of using quota sampling rather than stratified random sampling. (6)

   (b) A food manufacturer commissions a market researcher to interview 36 people in a supermarket. She is to ask their opinion of a new brand of fruit juice. The researcher is told that she must have:

   (i) a quarter of all interviewees aged 18 to 24 years old, the remainder aged 25 to 39;
   (ii) twice as many female interviewees as male;
   (iii) twice as many married interviewees as unmarried.

   Furthermore, these proportions must hold within each combination of both remaining variables. For example, a quarter of unmarried male interviewees must be aged 18 to 24.

   Calculate how many subjects the researcher must interview for each combination of age, sex and marital status. Present this information in the form of a table. (9)

3. Describe the difference between *observational* and *experimental* studies. Give one example of each type of study. Discuss one advantage and one disadvantage of experimental studies over observational studies. (10)

4. Describe in detail three ways of dealing with *outliers* (i.e. outlying measurements) when analysing the results of a survey. (9)

1

5. A group of politics students are asked by their teacher to construct a short questionnaire which will investigate how people's stated voting intentions in a forthcoming election are related to their sex, age and socio-economic status (as measured by employment status and salary). Three candidates X, Y and Z, are standing in the election. The questionnaire is to be sent to a random sample of voters chosen from the list of registered voters.

The questionnaire constructed by the students is given below. Write a report in which you comment on the quality of the questionnaire as a whole, and individual questions. If you identify any strengths in the students' questions, explain why you consider these good points. If you identify any weaknesses, make suggestions as to how the students might amend their questionnaire accordingly.

(19)

---

*Dear respondent,*

*We are conducting a survey of voting intentions in the forthcoming election. We would be grateful if you would complete this questionnaire and return it to us in the enclosed reply-paid envelope.*

*In particular, we want to confirm whether people of a high socio-economic status are more likely to vote for candidate X.*

*Please complete the following:*
*Sex      (1. Male      2. Female)*                                                 ..........

*Age      (in years)*                                                                            ..........

*Are you currently employed ?      (1. No    2. Yes)*                        ..........

*Have you been employed at any time in the last 12 months ?*
                                                          ....................................................

*What has been your occupation in the past 12 months ?*
                                                          ....................................................

*What is your current salary ?      (1.  < £10000   2.  > £10000)*        ..........

*Who do you intend to vote for in the forthcoming election ?*        ..........
          *(1.  Candidate X          2.  Candidate Y          3.  Candidate Z)*

6. Under stratified random sampling, it can be shown that, for a given cost, optimum precision of a particular estimate is obtained if the sampling fractions in the different strata are made proportional to the ratio of the standard deviation of responses in the stratum and the square root of the cost per unit sampled in the stratum. That is, for optimum precision,

$$n_i \text{ is made proportional to } \frac{N_i S_i}{\sqrt{c_i}}$$

where $n_i$ denotes sample size, $N_i$ denotes population size, $S_i$ denotes standard deviation, $c_i$ denotes cost per unit in the $i^{th}$ stratum.

Each year, a company surveys the level of employee stress in its four offices. This year, the company proposes to select a stratified random sample from its 5400 employees. Each of the sampled employees will be interviewed and subjected to a physical examination by trained investigators, and awarded a stress "score" of between 0 and 100. From previous years' results, it is known that there are differences in the cost of the examination per employee and in the standard deviation of stress score between offices. These differences are shown in the table below.

| Office Number | Total Number of employees $N_i$ | Standard deviation $S_i$ | Cost per employee (£) $c_i$ |
|---|---|---|---|
| 1 | 1400 | 25 | 25 |
| 2 | 500 | 30 | 25 |
| 3 | 2500 | 12 | 36 |
| 4 | 1000 | 20 | 16 |

Given that the company will spend only £ 5100 on this year's survey, calculate the sample sizes required in each stratum so that the estimate of the overall mean employee stress score has the best possible precision. (12)

7. Describe five features that you would consider desirable in any good computerised data-entry package. (10)

8. Bias in survey results may result from *non-response*.

   (i) Explain what is meant by the term *non-response*. State two reasons for non-response, giving a practical example of each. (4)

   (ii) Describe how non-response may lead to bias in survey results. (2)

   (iii) When reporting the results of a survey, why might it be important to distinguish between different types of non-response ? (2)

   (iv) Describe one method of dealing with (or reducing) bias due to non-response. (2)