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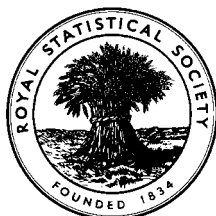
(formerly the Examinations of the Institute of Statisticians)

ORDINARY CERTIFICATE IN STATISTICS, 1999

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 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.*

1. The Office of National Statistics in the UK publishes the Retail Prices Index (RPI) every month. Similar consumer price indices are produced in other countries. The RPI is a weighted index formed from the prices of a large number of goods and services.

   (i) State the main use of such an index and suggest two groups of people who might be particularly interested in its value.

   (3)

   (ii) Outline the ways in which the values of the prices and weights used to form the Index could be determined.

   (6)

   (iii) A newspaper reports that "The value of the RPI in January 1996 was 150.2 (January 1987 = 100) and its value in January 1997 was 154.4".
   (a) Explain the meaning of "January 1987 = 100".
   (b) Calculate the percentage increase in the RPI between January 1996 and January 1997.

   (4)

2. There are a number of decisions to be made when conducting and analysing the results of a survey using a questionnaire. For each of the decisions below, outline one advantage and one disadvantage of the proposed method as opposed to the alternative.

   (i) Personal interview as opposed to a telephone survey.

   (ii) Offering a reward, announced in advance, as opposed to not doing so.

   (iii) Using the same respondents for a second survey on the same topic as opposed to selecting a new set of respondents.

   (iv) Including late returns in the analysis as opposed to ignoring them.

   (8)

3.	A Transportation Research Group is setting up a nation-wide database of road traffic accidents. Its aim is to co-ordinate the collection of data as part of an attempt to suggest measures for reducing the number and severity of accidents.

(i)	Design a simple form to collect data on 10 variables which you consider to be relevant.

(10)

(ii)	For two of your variables, list all possible values which could be entered on the database and also give a suitable missing value code in each case.

(4)

4.	The postal questionnaire shown on the following page was used as part of a national survey targetting the main shopper in households. Study it carefully and give six examples of good practice in its design and six examples of where the design could be improved.

(12)

<div align="center">**NATIONAL SHOPPING SURVEY**</div>

<u>**Instructions**</u>
1) Please have this questionnaire filled out by the main shopper in your household.
2) Please give the answers that apply to you by putting a "**X**" in the appropriate boxes.
3) Please return your completed survey in the reply-paid envelope enclosed.

1.   **ABOUT YOU**
1. Gender   Male •   Female •

2. What is your date of birth?   • •   • •   19• •
                                                     Day Month   Year

3. Which group best describes your combined household income?
1•	Up to £5,000	**2**•	£5,000 - £10,000
3•	£10,000 - £20,000	**4**•	£20,000 - £30,000
5•	£30,000 - £50,000	**6**•	£50,000 +

4. How many children have you got living at home?

5. Please indicate the age and gender of all children living at home:

	Girl	**Boy**		**Girl**	**Boy**		**Girl**	**Boy**
0-12 M	•	•	6Y	•	•	12Y	•	•
1Y	•	•	7Y	•	•	13Y	•	•
2Y	•	•	8Y	•	•	14Y	•	•
3Y	•	•	9Y	•	•	15Y	•	•
4Y	•	•	10Y	•	•	16Y	•	•
5Y	•	•	11Y	•	•	17+Y	•	•

2.   **YOUR SHOPPING**
When doing your main shopping:
1. Why do you buy where you do?
•	Distance	•	Range of products
•	Convenience	•	Prices
•	Quality of products	•	Customer service
•	Parking facilities	•	Other (Please state)
			_____

2. How far is the shop from your home?

3. What do you spend on groceries?
•	Under £10	•	£10 - £19.99
•	£20 - £29.99	•	£30 - £39.99
•	£40 - £49.99	•	£50 - £74.99
•	£75 - £99.99	•	£100 or more

4. How many packets of the following cereals do you buy in a week?
Corn flakes	1•	2•	3•	4•
Bran/High fibre	1•	2•	3•	4•
Muesli/oat	1•	2•	3•	4•
Children's cereals	1•	2•	3•	4•
Wheat biscuits	1•	2•	3•	4•
Other cereals	1•	2•	3•	4•

5. How many bottles of whisky are purchased by your household in one year?
       • 0   • 1   • 2-3   • 4-6   • 7+

Thank you for taking time to answer this questionnaire.
One last thing - please print your name and address below and we will send you some shopping coupons within 6 - 16 weeks.

Turn over

5.   A college of higher education wishes to obtain an estimate of the mean time spent each week by students using the computer facilities within the college. It is believed that this time may show wide variation between departments so a stratified random sampling method is to be used. There are also differences in the cost of obtaining the data from the students in the various departments because of scattered locations and varying modes of study. Details are:

Department	Number of students	Cost (£) per student
Business and Management	1200	0.81
Science and Technology	600	0.36
Art and Design	600	0.64

(i)   It has been decided to select a sample containing 160 students in total. Find the required sample size in each stratum if proportional allocation is used. What is the total cost of collecting the data using this allocation?

(5)

(ii)   Suppose, instead, that the college has decided that the maximum amount it can spend on this survey is £100. 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 ith stratum.

Estimates of the standard deviations of the weekly time spent on the computers by students in each department have been obtained from a pilot study. They are: Business and Management, 2.7 hours; Science and Technology, 1.8 hours; Art and Design, 1.6 hours.

Calculate the sample sizes required in each stratum so that the estimate of overall mean time has the best possible precision within the funds available.

(10)

6. Answers to questions in questionnaires generally need to be *coded* before they are used.

   (i)   State the three stages at which this coding can be carried out.

   (3)

   (ii) Outline the advantages and disadvantages of coding at each stage.

   (6)

   (iii) Give two rules for a well-designed question with pre-coded answers.

   (2)

   (iv) Design such a question to determine the education level of a respondent.

   (3)

7. The paperback edition of Moser and Kalton's *Survey Methods in Social Investigation* contains 488 pages of main text. Each page has a maximum of 44 lines and each line contains a maximum of 15 words.

   You are asked to count the number of letters in a simple random sample of 100 words chosen from the book. Suggest an economical method of carrying out the sampling using Table IV of your statistical tables. Your description should include details of

   (i)   selecting which numbers to use from the table;

   (ii) choosing a random page, line and word from the book;

   (iii) eliminating those choices which do not correspond to a word.

   (12)

   Demonstrate your method using two numeric examples from the table.

   (4)

8. Give, with explanation, two distinct examples of practical situations where the selection of a simple random sample is not feasible. In each case, suggest what type of sampling method would overcome these difficulties and state one disadvantage of this proposed method.

   (8)

BLANK PAGE