

## EXAMINATIONS OF THE ROYAL STATISTICAL SOCIETY

### ORDINARY CERTIFICATE IN STATISTICS, 2017

#### MODULE 1 : Collection and compilation of data

**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 8 printed pages.

This front cover is page 1.

Question 1 starts on page 2.

There are 7 questions altogether in the paper.

1. College lecturers wish to use a questionnaire to explore students' views on the usefulness of a new resource room available for students within the campus. Two different sampling schemes are being considered: (1) simple random sampling of the whole student body and (2) asking students to complete the questionnaire as they leave the resource room.
  - (i) Briefly describe the steps which may be involved when implementing these two different sampling approaches and also how you might administer the questionnaire. (8)
  - (ii) For each of the two sampling approaches given above, state three benefits and three drawbacks. (12)
  
2. A research team is investigating a chronic (long term) health condition and is interested in exploring changes in patients' symptoms over a period of time. Patients are asked to complete a questionnaire asking about their symptoms each week for a period of 12 weeks.
  - (i) State why the researchers have chosen a longitudinal study design to investigate the chronic health condition. (1)
  - (ii) State three problems which may occur in a longitudinal study leading to missing values. (3)
  - (iii) Describe three different approaches for dealing with missing values when analysing data from a longitudinal study such as this. (3)

3. (a) In the context of a questionnaire, explain what is meant by an *open question*, a *closed question with a single response* and a *closed question with multiple responses*. (4)
- (b) The following three questions are an extract from a questionnaire asking about participants' experiences after attending a music festival.

What is your age? \_\_\_\_\_

Please rate the quality of the food and drinks available at the event  
(1 = poor and 5 = excellent)

	1	2	3	4	5
Food and Drinks Quality					

On which day did you attend the event?

Friday	
Saturday	
Sunday	

- (i) Change the design of the questions in order to improve the quality of the data collected. The questions should be designed to collect information on
- the age profile of participants who attended the event;
  - the quality of the food and drinks available for purchase at the event;
  - which days participants attended the event.
- (8)
- (ii) Explain how your alternative design will improve the validity of the data obtained. (7)
- (iii) Give a coding plan for your questions in part (i). This should include the variable names required in a data spreadsheet to capture all of the information collected and also, where appropriate, how you would code them. (7)

4. Explain what is meant by a *pilot survey* and discuss in detail the reasons for implementing one.

(10)

5. A budget airline wishes to know what passengers thought of a new route, by surveying 500 who have flown on it during the first three months. A stratified sampling scheme is to be implemented with strata defined by the passengers' method of booking. These are

- (A) passengers who booked flights online via the airline's own website;
- (B) passengers who booked through an online travel agency;
- (C) passengers whose flights were booked by some other method.

The sales figures show that 4500 flights were booked in the first 3 months. 3056 passengers used method (A), 1261 used method (B) and 183 used method (C).

- (i) How many customers should be sampled in each stratum if the samples are to be proportional to the total numbers of passengers who booked by each of the three methods?

(4)

- (ii) The airline also wishes to compare the annual flying frequencies of the three strata of passengers.

- (a) Discuss any issues or problems the airline may encounter when calculating the mean annual flying frequencies if using the proportional sampling approach outlined in part (i) above.

(4)

- (b) Suggest an alternative stratified sampling approach which is better for this purpose. Show how you would calculate the required sample sizes.

(3)

6. The Human Resources Manager within a large organisation wishes to explore whether absence rates differ between departments within the organisation, and whether this is related to the age and gender of employees.

(i) Suggest what available data would be suitable for this purpose, and any ethical issues that the Human Resources Manager may need to consider.

(4)

The management also plan to measure employee satisfaction through a survey and need to decide on the design.

(ii) State the advantages and the disadvantages of collecting the required data by the following methods.

(a) A *longitudinal* survey conducted over the next 10 years of a sample of employees who joined the organisation within the last year.

(6)

(b) A *cross-sectional* survey of employees who joined the organisation 1, 5 and 10 years ago.

(6)

7. Health researchers plan to survey patients who attend a hospital orthopaedic outpatient clinic in order to collect information on their health conditions. Interviewees are selected using systematic sampling and selected patients are invited to take part in the survey which is carried out in a private room in the clinic using face-to-face interviews.

Discuss in detail how different sources of error may occur in carrying out this survey.

(10)

BLANK PAGE

BLANK PAGE

BLANK PAGE