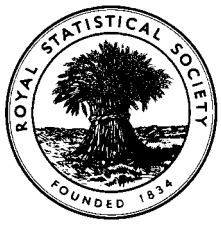
EXAMINATIONS OF THE ROYAL STATISTICAL SOCIETY
(formerly the Examinations of the Institute of Statisticians)



ORDINARY CERTIFICATE IN STATISTICS, 2000

Paper II

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.*

This examination paper consists of 7 printed pages followed by 1 blank page (page 8). This front cover is page 1. The reverse of the front cover is intentionally left blank and is page 2. Question 1 starts on page 3.

There are 8 questions altogether in the paper.

1. (i) Define the coefficient of variation of a set of data values. Describe a situation where the use of the coefficient of variation is helpful and a situation where its use might be misleading.

(3)

(ii) A set of data on the speed of vehicles has mean 32.6 mph correct to 1 decimal place. Write down the maximum and minimum values possible for this mean.

The set has standard deviation 3.0 mph correct to 1 decimal place. Write down the maximum and minimum possible values for this standard deviation.

Hence calculate the maximum and minimum values for the coefficient of variation, as a percentage correct to 1 decimal place.

(4)

2. (i) A set of 20 data values is calculated to have a mean of 36.80. It is then found that one of the values, which should have been 45.3, had been incorrectly entered as 34.5. Calculate the correct value of the mean.

(4)

(ii) The calculated standard deviation (with divisor 19) was 2.90. Show that the uncorrected sum of squares of all the data values (as originally recorded) is 27244.59.

(4)

(iii) Using the results of (i) and (ii) find the correct standard deviation.

(4)

(iv) Comment on the changes to the mean and standard deviation.

(1)

3. Users of a newly-refurbished shopping centre were surveyed to find out their views about how easy it was to manoeuvre their cars in the redesigned car-park. They were asked to record on a scale from 1 to 5 whether or not they agreed with the statement "I find it easy to manoeuvre my car in the car park", where

1=Strongly agree, 2= Agree, 3=Neither agree nor disagree, 4=Disagree, 5=Strongly disagree.

The results for a sample of 60 men and women are shown in the table.

Car Park Survey Data

Gender	*Rating*	*Gender*	*Rating*	*Gender*	*Rating*	*Gender*	*Rating*
M	3	F	1	M	1	F	2
F	4	F	2	M	1	F	2
F	5	F	5	F	5	F	5
F	4	M	3	M	3	M	2
F	2	M	3	F	3	M	2
F	3	F	5	M	1	F	1
M	1	F	5	F	4	F	5
F	1	F	4	M	3	M	2
F	3	F	4	F	1	F	2
F	2	F	5	M	2	F	4
F	2	M	3	F	2	F	2
M	1	M	1	F	2	F	4
M	3	M	3	F	1	M	1
M	2	F	5	F	2	F	2
F	3	M	3	F	4	F	5

(i) How many women and how many men took part in the survey?

(1)

(ii) Construct a frequency table for the ratings.

(3)

(iii) Construct a two-way contingency table showing the ratings broken down by gender.

(4)

(iv) Write a brief report on the conclusions to be drawn from the data, including an appropriate table of percentages.

(4)

4. A small business was started in 1995 and, as part of a five-year review, the managing director investigated the changes in the price of stationery purchased by the company over the period from 1995 to 1999. The prices charged and quantities purchased for a range of stationery items are shown in the table.

(i) Calculate

(a) a Laspeyres index of prices for 1999 (1995=100);

(b) a Paasche index of prices for 1999 (1995 = 100).

(8)

(ii) Interpret the values of these two indices and advise the Managing Director as to which of the two methods of calculation is more appropriate, in this instance, for measuring the change in stationery prices over the period 1995-99.

(3)

Prices of Stationery and Quantities Used in 1995 and 1999

Item	*Unit price (£) in 1995*	*Unit price (£) in 1999*	*Quantity purchased in 1995*	*Quantity purchased in 1999*
Box of envelopes	22.49	42.00	5	20
Ream of copier paper	8.69	13.99	10	30
Pack of folders	3.99	4.29	10	10
Box of paper clips	0.57	0.72	15	10

5. Two fair six-sided dice are rolled and the numbers X and Y on the top faces are noted. By drawing up a suitable table showing all the values taken by X and Y, or otherwise, find the probability that

(i) $X < 4$,

(ii) $X < 4$ and $Y < 3$,

(iii) $X < 4$ or $Y < 3$,

(iv) $2X + Y = 7$,

(v) $X < Y$,

(vi) the maximum of X and Y is 4 .

(12)

6. A large retailing company has 400 stores around the country. Researchers at its head office have set up a study relating weekly takings in a random sample of 12 stores to the number of people living within 20 minutes travelling time of the store. The data are shown in the table below.

(i) Using the graph paper provided, plot the data on a scatter diagram to indicate the relation between y and x, and comment on what your graph shows.

(4)

(ii) Calculate the Pearson product moment correlation coefficient, r. Interpret its value. You are given:

$\Sigma x = 254,\ \Sigma y = 7970,\ \Sigma x^2 = 5694,\ \Sigma y^2 = 5629500,\ \Sigma xy = 178330$

(5)

(iii) The slope of the linear regression line of y on x has been calculated as 30.32. Using the values provided, estimate the intercept of this regression line and plot the line on your diagram.

(4)

(iv) Predict the takings in a locality where

(a) 22,000 people live within 20 minutes travelling time of a proposed store,

(b) 35,000 people live within 20 minutes travelling time of a proposed store.

Without further calculation, indicate how confident you are that these are "good" predictions.

(5)

Takings of a random sample of 12 stores

Store	*Takings (£000)* (y)	*Population(000)* (x)
1	540	20
2	560	15
3	940	29
4	810	26
5	570	20
6	540	21
7	430	14
8	660	20
9	550	15
10	990	31
11	770	23
12	610	20

7. (i) Explain what you understand by the following terms in relation to a time series:

(a) trend,

(b) seasonal component.

(2)

(ii) Using the data in the table below and the graph paper provided, draw a line graph (i.e. a historigram) of the sales of Grainger's Department Stores, calculate the trend of sales using centred four-quarterly moving averages, plot the trend on the graph and use the graph to make a forecast of the trend for each quarter of 2000.

Discuss the reliability of your forecasts.

(15)

Sales (in £m) in Grainger's Department Stores 1995-99

Year	*Q1*	*Q2*	*Q3*	*Q4*
1995	13.4	16.6	17.0	23.7
1996	14.8	18.2	18.4	25.4
1997	15.8	19.1	19.2	26.9
1998	16.5	21.0	21.0	29.6
1999	18.4	22.6	23.2	30.4

8. Each of 200 air traffic accidents was caused by some combination of the weather (*W*), mechanical failure (*F*) or pilot error (*P*).

The table below gives the numbers of the 200 accidents in which the given factor or factors played a part.

Factors	*W*	*F*	*F* and *P*	*W* and *P*	*W* and *F*	*W* and *F* and *P*
Number	92	100	40	36	36	16

(i) Draw a suitable Venn diagram with 3 overlapping circles labelled *W*, *F* and *P*. Give the appropriate number of accidents in each of the 8 regions of your diagram.

In how many of the accidents did pilot error play a part?

(6)

(ii) Assuming that these 200 accidents can be thought of as a random sample of all accidents, estimate the probabilities that an accident is caused by 1, 2 or 3 of these factors.

(4)

BLANK PAGE