OCR
Oxford Cambridge and RSA

A Level Mathematics B (MEI)

H640/02 MEI Pure Mathematics and Statistics

Large Data Set 1
Statistics

Question Set 4

1 Fig. 1 shows the probability distribution of the discrete random variable X.

x	1	2	3	4	5
$P(X=x)$	0.2	0.1	k	$2k$	$4k$

Fig. 1

(a) Find the value of k. **[2]**

(b) Find $P(X \neq 4)$. **[2]**

2 Fig. 2 shows the time Lorraine spent in hours, t , answering e-mails during the working day. The data were collected over a number of months.

Time in hours, t	$0 \leqslant t < 1$	$1 \leqslant t < 2$	$2 \leqslant t < 3$	$3 \leqslant t < 4$	$4 \leqslant t < 6$	$6 \leqslant t < 8$
Number of days	28	36	42	31	24	12

Fig. 2

(a) Calculate an estimate of the mean time per day that Lorraine spent answering e-mails over this period. **[1]**

(b) Explain why your answer to part **(a)** is an estimate. **[1]**

When Lorraine accepted her job, she was told that the mean time per day spent answering e-mails would not be more than 3 hours.

(c) Determine whether, according to the data in Fig. 2, it is possible that the mean time per day Lorraine spends answering e-mails is in fact more than 3 hours. **[1]**

3 Fig. 3 shows the number of times that students at a sixth form college visited a recreational mathematics website during the first week of the summer term.

Number of visits to website	0	1	2	3	4	5
Number of students	24	38	17	12	4	2

Fig. 3

(a) State the value of the mid-range of the data. **[1]**

(b) Describe the shape of the distribution. **[1]**

(c) State the value of the mode. **[1]**

4 A team called "The Educated Guess" enter a weekly quiz. If they win the quiz in a particular week, the probability that they will win the following week is 0.4, but if they do not win, the probability that they will win the following week is 0.2.

In week 4 The Educated Guess won the quiz.

(a) Calculate the probability that The Educated Guess will win the quiz in week 6. **[3]**

Every week the same 20 quiz teams, each with 6 members, take part in a quiz. Every member of every team buys a raffle ticket. Five winning tickets are drawn randomly, without replacement. Alf, who is a member of one of the teams, takes part every week.

(b) Calculate the probability that, in a randomly chosen week, Alf wins a raffle prize. **[2]**

(c) Find the smallest number of weeks after which it will be 95% certain that Alf has won at least one raffle prize. **[4]**

5 Club 65–80 Holidays fly jets between Liverpool and Magaluf. Over a long period of time records show that half of the flights from Liverpool to Magaluf take less than 153 minutes and 5% of the flights take more than 183 minutes.

An operations manager believes that flight times from Liverpool to Magaluf may be modelled by the Normal distribution.

(a) Use the information above to write down the mean time the operations manager will use in his Normal model for flight times from Liverpool to Magaluf. **[1]**

(b) Use the information above to find the standard deviation the operations manager will use in his Normal model for flight times from Liverpool to Magaluf, giving your answer correct to 1 decimal place. **[3]**

(c) Data is available for 452 flights. A flight time of under 2 hours was recorded in 16 of these flights. Use your answers to parts **(a)** and **(b)** to determine whether the model is consistent with this data. **[3]**

The operations manager suspects that the mean time for the journey from Magaluf to Liverpool is less than from Liverpool to Magaluf. He collects a random sample of 24 flight times from Magaluf to Liverpool. He finds that the mean flight time is 143.6 minutes.

(d) Use the Normal model used in part **(c)** to conduct a hypothesis test to determine whether there is evidence at the 1% level to suggest that the mean flight time from Magaluf to Liverpool is less than the mean flight time from Liverpool to Magaluf. **[7]**

(e) Identify two ways in which the Normal model for flight times from Liverpool to Magaluf might be adapted to provide a better model for the flight times from Magaluf to Liverpool. **[2]**

6 The jaguar is a species of big cat native to South America. Records show that 6% of jaguars are born with black coats. Jaguars with black coats are known as black panthers. Due to deforestation a population of jaguars has become isolated in part of the Amazon basin. Researchers believe that the percentage of black panthers may not be 6% in this population.

(a) Find the minimum sample size needed to conduct a two-tailed test to determine whether there is any evidence at the 5% level to suggest that the percentage of black panthers is not 6%. **[3]**

A research team identifies 70 possible sites for monitoring the jaguars remotely. 30 of these sites are randomly selected and cameras are installed. 83 different jaguars are filmed during the evidence gathering period. The team finds that 10 of the jaguars are black panthers.

(b) Conduct a hypothesis test to determine whether the information gathered by the research team provides any evidence at the 5% level to suggest that the percentage of black panthers in this population is not 6%. **[7]**

7 The pre-release material includes data concerning crude death rates in different countries of the world. Fig. 7.1 shows some information concerning crude death rates in countries in Europe and in Africa.

	Europe	Africa
n	48	56
minimum	6.28	3.58
lower quartile	8.50	7.31
median	9.53	8.71
upper quartile	11.41	11.93
maximum	14.46	14.89

Fig. 7.1

(a) Use your knowledge of the large data set to suggest a reason why the statistics in Fig. 7.1 refer to only 48 of the 51 European countries. **[1]**

(b) Use the information in Fig. 7.1 to show that there are no outliers in either data set. **[3]**

The median age in Germany is 46.5 and the crude death rate is 11.42. The median age in Cyprus is 36.1 and the crude death rate is 6.62.

(c) Calculate an estimate of the number of deaths in Libya in a year. **[1]**

The median age in Germany is 46.5 and the crude death rate is 11.42. The median age in Cyprus is 36.1 and the crude death rate is 6.62.

(d) Explain why a country like Germany, with a higher median age than Cyprus, might also be expected to have a higher crude death rate than Cyprus. **[1]**

Fig. 7.2 shows a scatter diagram of median age against crude death rate for countries in Africa and Fig. 7.3 shows a scatter diagram of median age against crude death rate for countries in Europe.

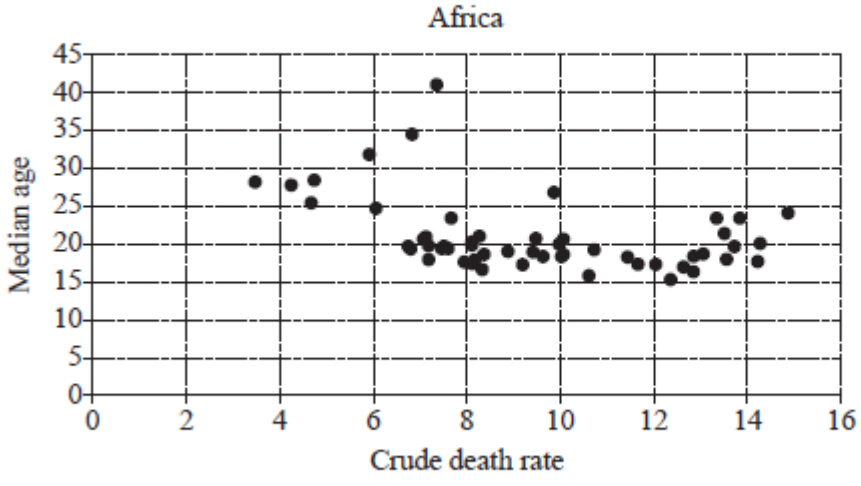


Fig. 7.2

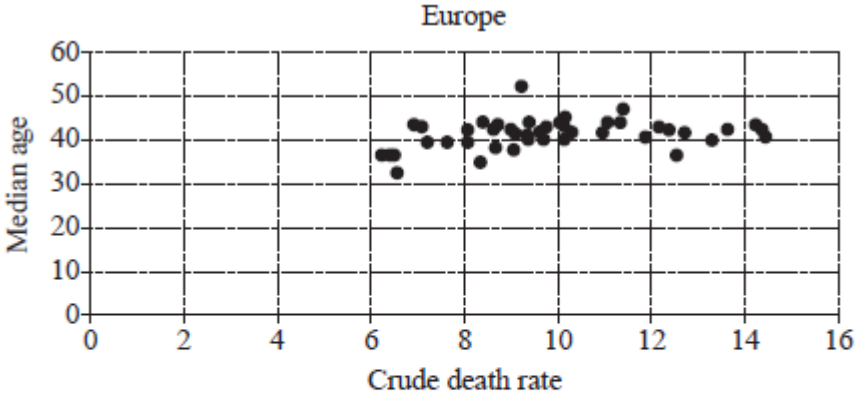


Fig. 7.3

The rank correlation coefficient for the data shown in Fig. 7.2 is −0.281 206.

The rank correlation coefficient for the data shown in Fig. 7.3 is 0.335 215.

(e) Compare and contrast what may be inferred about the relationship between median age and crude death rate in countries in Africa and in countries in Europe. **[2]**

8 **You must show detailed reasoning in this question.**

The screenshot in Fig. 8 shows the probability distribution for the continuous random variable X, where $X \sim \mathrm{N}(\mu, \sigma^2)$.

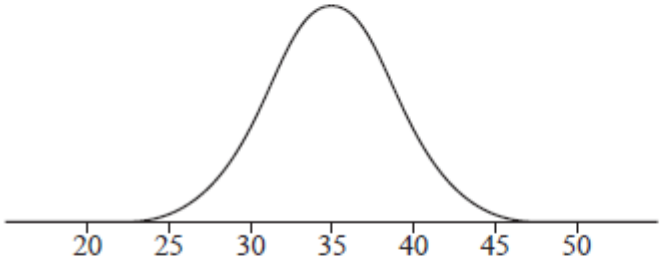


Fig. 8

The distribution is symmetrical about the line $x = 35$ and there is a point of inflection at $x = 31$.

Fifty independent readings of X are made. Show that the probability that at least 45 of these readings are between 30 and 40 is less than 0.05. **[6]**

Total Marks for Question Set 4: 59

Resource Materials

Question Set No: 4

Fig. 1

x	1	2	3	4	5
$P(X = x)$	0.2	0.1	k	$2k$	$4k$

Fig. 2

Time in hours, t	$0 \leqslant t < 1$	$1 \leqslant t < 2$	$2 \leqslant t < 3$	$3 \leqslant t < 4$	$4 \leqslant t < 6$	$6 \leqslant t < 8$
Number of days	28	36	42	31	24	12

Fig. 3

Number of visits to website	0	1	2	3	4	5
Number of students	24	38	17	12	4	2

Fig. 7.1

	Europe	Africa
n	48	56
minimum	6.28	3.58
lower quartile	8.50	7.31
median	9.53	8.71
upper quartile	11.41	11.93
maximum	14.46	14.89

Fig. 7.2

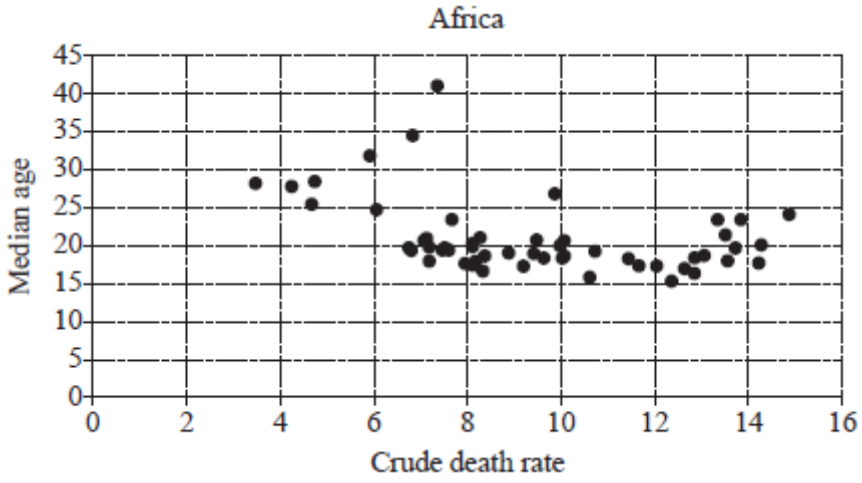


Fig. 7.3

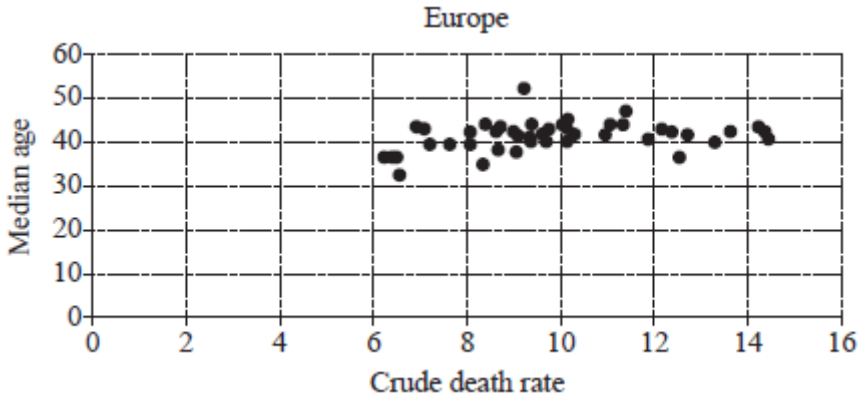
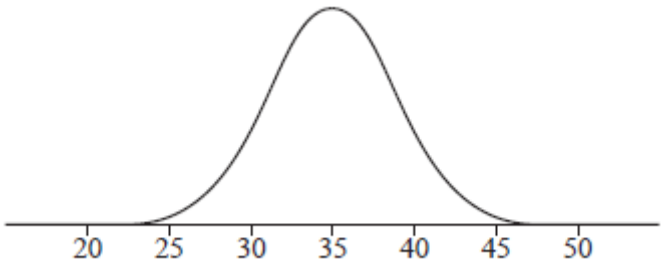

Europe
Median age
60
50
40
30
20
10
0
0
2
4
6
8
10
12
14
16
Crude death rate

Fig. 8


20
25
30
35
40
45
50