A2 Statistics Practice Paper A
60 Marks

60 Minutes

1. Judith believes that mathematical ability and chess-playing ability are related. She asks 20 randomly chosen chess players, with known British Chess Federation (BCF) ratings X, to take a mathematics aptitude test, with scores Y. The results are summarised as follows.

$$n = 20$$
$$\sum x = 3600$$
$$\sum x^2 = 660\,500$$
$$\sum y = 1440$$
$$\sum y^2 = 105\,280$$
$$\sum xy = 260\,990$$

a. Calculate the value of Pearson's product-moment correlation coefficient, r. **(2)**

b. Assume now that the data has a bivariate normal distribution. Test at the 5% significance level whether there is evidence that chess players with higher BCF ratings are better at mathematics. State clearly your null and alternate hypotheses.

(4)

(Total: 6 marks)

2. An estate agent recorded the price per square metre, p (£/m ^{2}), for 7 two-bedroom houses. He then coded the data using the coding $q = \frac{p-a}{b}$, where a and b are positive constants.

His results are shown in the table below.

p	1840	1848	1830	1824	1819	1834	1850
q	4.0	4.8	3.0	2.4	1.9	3.4	5.0

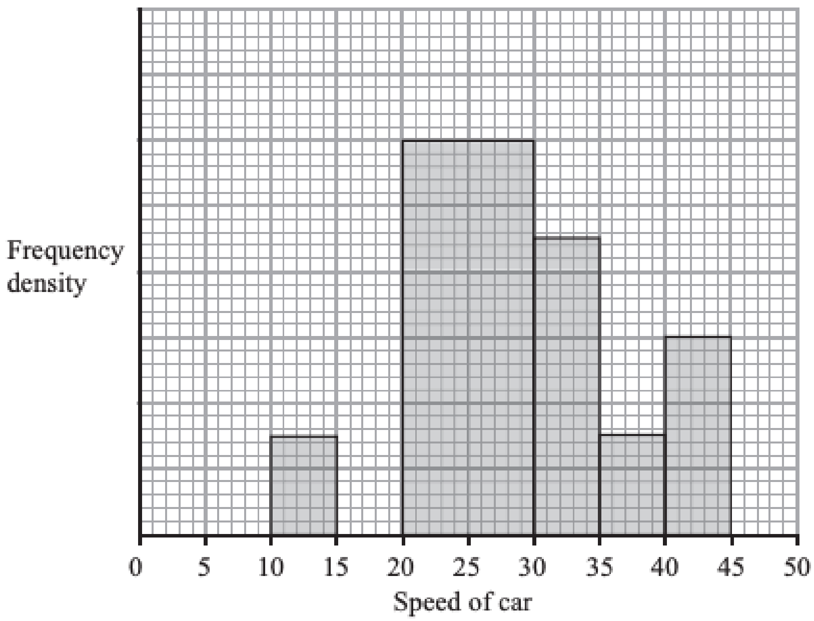
a. Find the value of a and the value of b. **(2)**

b. Given that, $S_{dd} = 1.02$ $\quad$ $S_{qq} = 8.22$ $\quad$ $S_{dq} = -2.17$

Calculate the product moment correlation coefficient between d and q **(2)**

(Total: 4 marks)

3. A policeman records the speed of the traffic on a busy road with a 30 mph speed limit. He records the speeds of a sample of 450 cars. The histogram on the next page represents the results.



a. Calculate the number of cars that were exceeding the speed limit by at least 5 mph in the sample. **(4)**

b. Estimate the value of the mean speed of the cars in the sample. **(3)**

c. Estimate, to 1 decimal place, the value of the median speed of the cars in the sample. **(2)**

(Total mark: 9)

4. A researcher measured the foot lengths of a random sample of 120 ten-year-old children. The lengths are summarised in the table below.

Foot length, l, (cm)	Number of children
$10 \leqslant l < 12$	5
$12 \leqslant l < 17$	53
$17 \leqslant l < 19$	29
$19 \leqslant l < 21$	15
$21 \leqslant l < 23$	11
$23 \leqslant l < 25$	7

a. Using interpolation, estimate the median of this distribution. **(2)**

b. Calculate estimates for the mean and the standard deviation of these data **(6)**

It can be shown that the median is 17.1, the mean is 17.1 and the standard deviation is 3.3

$$\text{Coefficient of skewness} = \frac{3(\text{mean} - \text{median})}{\text{standard deviation}}$$

c. Evaluate this coefficient and comment on the skewness of these data. **(3)**

(Total mark: 11)

5. A telesales operator is selling a magazine. Each day he chooses a number of people to telephone. The probability that each person he telephones buys the magazine is 0.1

a. Suggest a suitable distribution to model the number of people who buy the magazine from the telesales operator each day

(1)

b. On Monday, the telesales operator telephones 10 people. Find the probability that he sells at least 4 magazines.

(3)

c. Calculate the least number of people he needs to telephone on Tuesday, so that the probability of selling at least 1 magazine, on that day, is greater than 0.95

(3)

A call centre also sells the magazine. The probability that a telephone call made by the call centre sells a magazine is 0.05.

The call centre telephones 100 people every hour.

d. Using a suitable approximation, find the probability that more than 10 people telephoned by the call centre buy a magazine in a randomly chosen hour.

(3)

(Total marks: 10)

6a. State the conditions under which the normal distribution may be used as an approximation to the binomial distribution

(2)

A company sells seeds and claims that 55% of its pea seeds germinate.

To test the company's claim, a random sample of 220 pea seeds was planted.

b. State the hypotheses for a two-tailed test of the company's claim. **(2)**

c. Given that 135 of the 220 pea seeds germinated, use a normal approximation to test, at the 5% level of significance, whether or not the company's claim is justified.

(7)

(Total marks: 11)

7. Helen is studying the daily mean wind speed for Camborne using the large data set from 1987. The data for one month are summarized in Table 1 below.

Windspeed	n/a	6	7	8	9	11	12	13	14	16
Frequency	13	2	3	2	2	3	1	2	1	2

Table 1

a. Calculate the mean for these data. **(1)**

b. Calculate the standard deviation for these data and state the units. **(2)**

Month	*A*	*B*	*C*	*D*	*E*
Mean	7.58	8.26	8.57	8.57	11.57
Standard Deviation	2.93	3.89	3.46	3.87	4.64

Table 2

c. Using your knowledge of the large data set, suggest, giving a reason, which month had a mean of 11.57 **(2)**

(Total marks: 5)

8a. Explain what is meant by a population **(1)**

b. Explain what is meant by a statistic **(1)**

A researcher took a sample of 100 voters from a certain town and asked them who they would vote for in an election. The proportion who said they would vote for Dr Smith was 35%.

c. State the population and the statistic in this case. **(2)**

(Total marks: 4)

Total Marks for Paper: 60

Mark Scheme

1a	Use of calculator	**M1**
	0.4	**A1**
1b	H_0: Higher maths scores are not associated with higher BCF grading H_1: Positively associated	**B1**
	CV: 0.3783	**B1**
	0.400 > 0.3783 so reject H_0	**M1**
	Significant evidence that higher maths sores are associated with higher BCF grading.	**A1**

2a	$\frac{1840-a}{b} = 4.0$ $\frac{1848-a}{b} = 4.8$	**M1**
	$a = 1800$ $b = 10$	**A1**
2b	$r = \frac{-2.17}{\sqrt{1.02 \times 8.22}}$	**B1**
	$r = -0.749$	**B1**

3a	One large square = $\frac{40}{22.5}$ = 20 squares	**M1**
	One large square = 20 cars	**A1**
	4.5 x 20	**M1**
	= 90 cars	**A1**
3b	$mean = \frac{30 \times 12.5 + 240 \times 25 + 90 \times 32.5 + 30 \times 37.4 + 60 \times 42.5}{450} = \frac{12975}{450}$	**M1** **M1**
	$mean = 28.8$	**A1**
3c	$Q_2 = 20 + \frac{195}{240} \times 10$	**M1**
	28.1 mph	**A1**

4a	$Q_2 = 17 + \frac{60-58}{29} \times 2$	**M1**
	= 17.1	**A1**
4b	$\sum fx = 2055.5$ $\sum fx^2 = 36500.25$	**B1** **B1**
	Use of midpoints with at least one correct attempt	**B1**
	Mean = 17.1	**M1**
	$\sigma = \sqrt{\frac{36500.25}{120} - \left(\frac{2055.5}{120}\right)^2}$	**M1**
	$\sigma = 3.28$	**A1**
4c	Coefficient of skewness = 0	**M1** **A1**
	Therefore, no skew	**B1**

5a	Binomial distribution	**B1**
5b	$Y \sim B(10, 0.1)$	**B1**
	$P(Y \geq 4) = 1 - P(Y \leq 3)$	**M1**
	= 1 – 0.9872 = 0.0128	**A1**
5c	$0.9^n < 0.05$	**M1**
	$n > 28.4$	**A1**

	$n = 29$	**A1**
5d	$C \sim P_o(5)$	**B1**
	$P(C > 10) = 1 - P(C \leq 10)$	**M1**
	$= 1 - 0.9863$ $= 0.0137$	**A1**

6a	Large	**B1**
	0.5	**B1**
6b	$H_0 : p = 0.55$	**A1**
	$H_1 : p \neq 0.55$	**A1**
6c	$X \sim N(121, 54.45)$	**B1**
	$P(X \geq 134.5) = P\left(Z \geq \frac{134.5-121}{\sqrt{54.45}}\right)$	**M1** **M1** **A1**
	$P(Z \geq 1.8295)$ $= 1 - 0.9664$ $= 0.0337$	**A1**
	Accept H_0, not in CR, therefore not significant	**M1**
	The company's claim is justified	**A1**

7a	Mean = 10.2	**B1**
7b	$\sigma = 3.17$	**B1**
	Knots	**B1**
7c	September or October	**B1**
	As is it windier in autumn (o.e)	**B1**

8a	A collection of all items	**B1**
8b	A calculation based only on known data from a sample	**B1**
8c	Votes	**B1**
	Percentage/population	**B1**