OCR
RECOGNISING ACHIEVEMENT

OXFORD CAMBRIDGE AND RSA EXAMINATIONS

Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education

MATHEMATICS **2641**

Probability & Statistics 1

Friday **11 JUNE 2004** Morning 1 hour 20 minutes

Additional materials:
Answer booklet
Graph paper
List of Formulae (MF8)

TIME 1 hour 20 minutes

INSTRUCTIONS TO CANDIDATES

- Write your Name, Centre Number and Candidate Number in the spaces provided on the answer booklet.
- Answer **all** the questions.
- Give non-exact numerical answers correct to 3 significant figures unless a different degree of accuracy is specified in the question or is clearly appropriate.
- You are permitted to use a graphic calculator in this paper.

INFORMATION FOR CANDIDATES

- The number of marks is given in brackets [ ] at the end of each question or part question.
- The total number of marks for this paper is 60.
- Questions carrying smaller numbers of marks are printed earlier in the paper, and questions carrying larger numbers of marks later in the paper.
- **You are reminded of the need for clear presentation in your answers.**

This question paper consists of 4 printed pages.

[Turn over

1 Two people with the same occupation were asked to give ratings out of 100 for each of five different aspects of their job. Their ratings are given in the table below.

Aspect	A	B	C	D	E
Person 1	71	63	84	57	64
Person 2	12	62	20	85	31

(i) Calculate Spearman's rank correlation coefficient for the above data. [4]

(ii) Explain what your answer to part **(i)** tells you about the two people's ratings. [1]

2 A child's game uses five bricks. One is blue, one is green, one is yellow and two are white. The five bricks are arranged in a line.

(i) How many different possible arrangements of the colours are there? [2]

(ii) Assuming that all the arrangements in part **(i)** are equally likely, find the probability that the two white bricks are at the ends of the line. [3]

3 Paul plays a game in which a fair coin is spun 4 times. If the number of heads is 0 or 1, Paul loses £5, if the number of heads is 2, Paul wins £5 and if the number of heads is 3 or 4, Paul wins £10. Let £W be the amount which Paul wins in one randomly chosen game.

(i) Show that $P(W = -5) = \frac{5}{16}$. [2]

(ii) Copy and complete the table below to show the probability distribution of W. [2]

w	-5	5	10
$P(W = w)$	$\frac{5}{16}$		

(iii) Show that $E(W) = \frac{55}{16}$. [2]

(iv) Find $Var(W)$. [3]

4 A student was experimenting with an electrical circuit in which the resistance of one component could be varied. The student increased the resistance in fixed steps from 10 units to 100 units and measured the voltage drop when a fixed current was passed through it. The table below gives the student's results.

Resistance, x units	10	20	30	40	50	60	70	80	90	100
Voltage drop, y units	149	202	253	307	353	407	443	451	550	602

$[n = 10,\ \Sigma x = 550,\ \Sigma y = 3717,\ \Sigma x^2 = 38\,500,\ \Sigma y^2 = 1\,576\,075,\ \Sigma xy = 244\,260.]$

(i) The student wished to use his experimental data to estimate the resistance which would be needed for the voltage drop to be 220 units. Calculate the equation of the appropriate regression line and use it to estimate the resistance, x, which would correspond to $y = 220$. [6]

(ii) Calculate the product moment correlation coefficient for the data and use it to comment on the reliability of the estimate found in part **(i)**. [3]

5 Andy plays a lottery game once a week for 10 weeks. He knows that he has a probability of $\frac{1}{57}$ of winning each time he plays. Let X be the number of weeks out of 10 in which Andy wins.

(i) State the distribution of X, giving the values of any parameters, and state one assumption required to use this distribution as a suitable model. [3]

(ii) Calculate

(a) $P(X = 2)$, [3]

(b) $P(X > 2)$. [3]

(iii) Write down the value of $E(X)$. [1]

6 The table below refers to the mass, m kg, of each of a sample of 60 dogs examined in a vet's surgery.

Mass, m kg	$0 \leqslant m < 5$	$5 \leqslant m < 10$	$10 \leqslant m < 15$	$15 \leqslant m < 20$	$20 \leqslant m < 30$	$30 \leqslant m < 50$
Frequency	2	7	17	19	8	7

(i) Draw a cumulative frequency graph for the data in the table. [3]

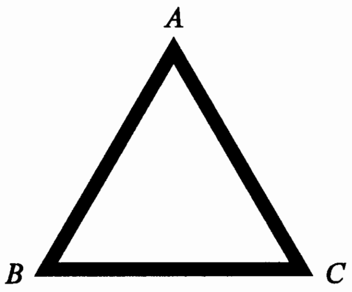
(ii) Use your cumulative frequency graph to estimate the median and the interquartile range for the data. [3]

(iii) You are now given that the minimum mass in the sample of 60 dogs was 4 kg and the maximum was 47 kg. Use your estimates from part **(ii)** to draw a box-and-whisker plot of the data. [3]

(iv) Give one feature of the data which you can deduce from a box-and-whisker plot more easily than from a cumulative frequency graph. [1]

[Question 7 is printed overleaf.]

7 Siân is involved in a game in which she runs along three paths in the form of a triangle, as in the diagram below.



When she arrives at a corner, she chooses her subsequent direction according to the following rules.

- When she is at A, she chooses path AB with probability $\frac{2}{3}$ and she chooses path AC with probability $\frac{1}{3}$.
- When she is at B, she chooses path BA with probability $\frac{3}{4}$ and she chooses path BC with probability $\frac{1}{4}$.
- When she is at C, she chooses path CA with probability $\frac{4}{5}$ and she chooses path CB with probability $\frac{1}{5}$.
- Once she has chosen a particular path she runs to the other end of the path.
- She starts at A.

(i) Show that the probability that she returns to A after choosing two paths is $\frac{23}{30}$. [3]

(ii) Find the probability that she returns to A after choosing three paths. [4]

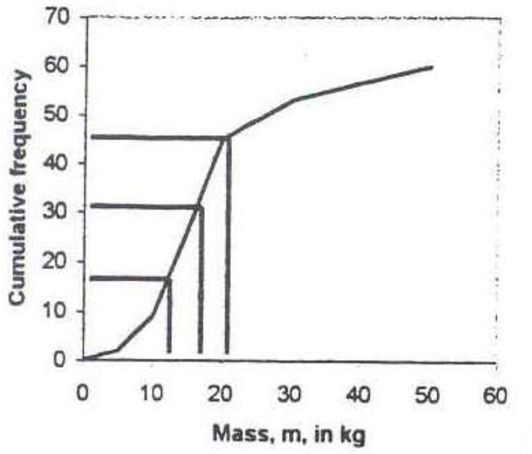
(iii) Find the probability that she is at B after four choices. [5]

1 (i)	<table><tr><td>Aspect</td><td>A</td><td>B</td><td>C</td><td>D</td><td>E</td></tr><tr><td>Person 1</td><td>71</td><td>63</td><td>84</td><td>57</td><td>64</td></tr><tr><td>Person 2</td><td>12</td><td>62</td><td>20</td><td>85</td><td>31</td></tr><tr><td>Rank 1</td><td>2</td><td>4</td><td>1</td><td>5</td><td>3</td></tr><tr><td>Rank 2</td><td>5</td><td>2</td><td>4</td><td>1</td><td>3</td></tr><tr><td>*d*</td><td>-3</td><td>2</td><td>-3</td><td>4</td><td>0</td></tr></table> $\Sigma d^2 = 9 + 4 + 9 + 16 + 0 = 38$ Spearman's rank correlation Coefficient = $1 - \frac{6 \times 38}{5 \times 24} = \frac{-9}{10} = -0.9$	B1 M1 M1 A1 4	Correct ranks (or reverse) Attempt to find *d* (or d^2) from ranked or ordered data Correct formula for Spearman *used* and $\|r\| < 1$ Correct answer -0.9 or $\frac{-9k}{10k}$ cao
(ii)	Spearman's rank correlation coefficient shows that the two people have different, opposite views, or no or little agreement when considering aspects of their job	B1 1	Comment in context, consistent with r_s value $\|r\| < 1$
2 (i)	Number of possible arrangements = $\frac{5!}{2}$ = **60**	M1 A1 2	5! or 120 seen (not in $^5C_3$) 60, cao
(ii)	Number of arrangements in which the white bricks are at each end = 3! or Number of arrangements in which both bricks are at either end = 3! × 2! Therefore P(white bricks are at each end) = $\frac{3!}{60} = \frac{6}{60}$ = **0.1** or P(both white bricks at either end) = $\frac{3! \times 2!}{60}$ = **0.2** or P(white at each end or both at either end) = 0.1 + 0.2 = **0.3**	M1 M1 A1 3	3! Seen for either case *their* 3! Divided by *their* (i) 0.1 or 0.2 or 0.3 $\frac{k}{10k}$ or $\frac{k}{5k}$ or $\frac{3k}{10k}$

3 (i)	Let X = number of heads in 4 randomly chosen spins of the coin. $P(W = -5) = P(X = 0 \text{ or } 1) = \frac{1}{16} + \frac{4}{16}$ $= \frac{5}{16}$ **AG**	M1 A1 2	Attempt to find P(X = 0) or P(X = 1) Wholly correct and convincing attempt (allow decimals)
(ii)	Distribution table for X | x | 0 | 1 | 2 | 3 | 4 | | $P(X = x)$ | $\frac{1}{16}$ | $\frac{4}{16}$ | $\frac{6}{16}$ | $\frac{4}{16}$ | $\frac{1}{16}$ | Distribution tale for W | w | -5 | 5 | 10 | | $P(W = w)$ | $\frac{5}{16}$ | $\frac{6}{16}$ | $\frac{5}{16}$ |	M1 A1 2	A clear attempt to derive $P(W = 5)$ or $P(W = 10)$ Wholly correct table
(iii)	$E(W) = (-5) \times \frac{5}{16} + 5 \times \frac{6}{16} + 10 \times \frac{5}{16}$ $= \frac{-25+30+50}{16} = \frac{55}{16} (= 3.4375)$ **AG**	M1 A1 2	Use of Σwp for *their* distribution table, at least 2 wp terms added Wholly correct method
(iv)	$E(W^2) = (-5)^2 \times \frac{5}{16} + 5^2 \times \frac{6}{16} + 10^2 \times \frac{5}{16}$ $= \frac{775}{16}$ So $\text{Var}(W) = E(W^2) - [E(W)]^2$ $= \left(\frac{775}{16}\right) - \left(\frac{55}{16}\right)^2$ $= \frac{9375}{256} = 36.62.. = 36.6$ (3 sf)	M1 M1 A1 3	Use of $\Sigma w^2 p$ for *their* distribution table, at least 2 $w^2 p$ terms added Subtracting (*their* mean)^2 $\frac{9375k}{256k}$ or a.r.t. 36.6

4 (i)	Since x is a controlled variable, only the y on x line is appropriate	M1		Use of y on x line
	$S_{xy} = \Sigma xy - \frac{(\Sigma x)(\Sigma y)}{n} =$ $= 244260 - \frac{550 \times 3717}{10} = 39825$ $S_{xx} = \Sigma x^2 - \frac{(\Sigma x)^2}{n}$ $= 38500 - \frac{550^2}{10} = 8250$ $\bar{x} = 55,\ \bar{y} = 371.7$			
	$b = \frac{S_{xy}}{S_{xx}} = \frac{39825}{8250} = 4.82727...$	M1		$\frac{S_{xy}}{S_{xx}}$ used
	$a = \bar{y} - b\bar{x}$ = 371.7 – 4.82727… 55 = 106.2			May be implied if calculator routine is used
	Equation of line is	M1		Using $= \bar{y} - b\bar{x}$ with *their* b
	$y = 4.82727\ldots x + 106.2$ $y = 4.83x + 106$ (3 sf)	A1		$y = 4.83x + 106.2$, or correct equivalent (does **not** need to be in the form $y = a + bx$)
	Estimated value of x $= \frac{(220 - 106.2)}{4.82727...}$ = 23.57438 **= 23.6 (3 sf)**	M1		Substitute y = 220 into *their* equation
		A1	6	a.r.t. 23.6
				Use of the x on y line
		M1		For $\frac{S_{xy}}{S_{xx}}$ used
		M1		Using $\bar{x} - b'\bar{y}$ with *their b′*
		A1		$x = -21.1 + 0.205y$
		M1		Substitute y = 220 into *their* equation
(ii)	$S_{yy} = \Sigma y^2 - \frac{(\Sigma y)^2}{n}$ $= 1576075 - \frac{3717^2}{10}$ = 194466.1			
	$r = \frac{S_{xy}}{\sqrt{S_{xx} \times S_{yy}}}$	M1		Calculator of formula correctly used or equivalent (may be implied)
	$= \frac{39825}{\sqrt{8250 \times 194466.1}}$ = 0.99427… = **0.994 (3 sf)**	A1		Correct answer, a.r.t. 0.994
	This is a very high positive correlation so the estimate is likely to be reliable	B1	3	Comment consistent with *their r* value, provided |r| < 1

5 (i)	$X \sim B(10, \frac{1}{57})$ Independence: whether Andy wins a particular lottery game is independent of whether he has won any other game. Two possible outcomes: for each game Andy either wins or loses.	B1) B1) B1 3	Binomial stated n = 10 and $p = \frac{1}{57}$ stated clearly One valid comment in context
(ii)(a)	$P(X = 2) = {}^{10}C_2 \times \left(\frac{1}{57}\right)^2 \times \left(\frac{56}{57}\right)^8$ = 0.0120217633 **= 0.012**	M1 M1 A1 3	Their ${}^nC_2 \times p^2 \times (1 - p)^{n-2}$ used Wholly correct method a.r.t. 0.012
(b)	$P(X > 2)$ $= 1 - P(X = 0) - P(X = 1) - P(X = 2)$ = 1 – [0.83778… + 0.14960… + 0.01202…] = 0.00059074… **= 0.000591 (3 sf)**	M1 M1 A1 3	1 – [P(X=0)+P(X=1)+P(X=2)] with at least 2 probs attempted Wholly correct method a.r.t. 0.0006
(iii)	$E(X) = np = 10 \times \frac{1}{57} = \frac{10}{57}$ = 0.175438… = **0.175 (3 sf)**	B1 1	$= \frac{10k}{57k}$ or a.r.t. 0.175

<table>
<tr><td>6 (i)</td><td>
<table>
<tr><td>Mass, m in kg</td><td>Cumulative frequency</td></tr>
<tr><td>m < 5</td><td>2</td></tr>
<tr><td>m < 10</td><td>9</td></tr>
<tr><td>m < 15</td><td>26</td></tr>
<tr><td>m < 20</td><td>45</td></tr>
<tr><td>m < 30</td><td>53</td></tr>
<tr><td>m < 50</td><td>60</td></tr>
</table>


</td><td>M1

M1

A1

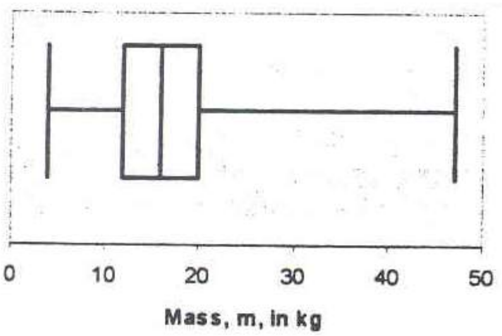
3</td><td>At least one correct cumulative frequency seen, other than 2

At least 4 points correct with the correct (u.c.b., cum freq)

Wholly correct diagram</td></tr>
<tr><td>6 (ii)</td><td>From the graph

Reading from the CF axis at 15 (or 15.25) for Q_1, at 30 (or 30.5) for Q_2, at 45 (or 45.75) for Q_3</td><td>M1

A1

A1
3</td><td>Correct method for either the median or for a quartile
Their Q_2 from their CF curve, provided u.c.b's used
Their IQR from their CF curve</td></tr>
<tr><td>6 (iii)</td><td>

</td><td>M1

M1

A1
3</td><td>A recognisable attempt at a boxplot

At least 4 from 4: their Q_1: their Q_2 their Q_3: 47 correctly plotted

Wholly correct diagram</td></tr>
<tr><td>6 (iv)</td><td>Comment on skewness, range, IQR, 5 summary numbers, max and/or min values, symmetry</td><td>B1

1</td><td>One valid feature of data which can be deduced more easily from a boxplot, but do not allow median and/or quartiles.</td></tr>
</table>

7 (i)	Possible routes: ABA → prob $= \frac{2}{3}\times\frac{3}{4}$ ACA → prob $= \frac{1}{3}\times\frac{4}{5}$ P(back at A) $= \frac{1}{2}+\frac{4}{15}=\frac{15}{30}+\frac{8}{30}$ $= \frac{23}{30}$ **AG**	M1 M1 A1 3	One correct product seen Both correct routes identified (letters, probs, tree diagram) **and** one correct product. No other routes allowed. Wholly convincing and correct
7 (ii)	Possible routes= ABCA or ACBA So prob $= \frac{2}{3}\times\frac{1}{4}\times\frac{4}{5}+\frac{1}{3}\times\frac{1}{5}\times\frac{3}{4}$ $= \frac{2}{15}+\frac{1}{20}=\frac{8}{60}+\frac{3}{60}$ $= \frac{11}{60}$ or **0.183.. = 0.183** (3 sf)	M1 M1 M1 A1 4	One correct route identified Both correct routes identified and one correct product Wholly correct method (no other routes) $\frac{11k}{60k}$ or a.r.t. 0.183
7 (iii)	Possible routes ACBCB → $\frac{1}{3}\times\frac{1}{5}\times\frac{1}{4}\times\frac{1}{5}$ ACBAB → $\frac{1}{3}\times\frac{1}{5}\times\frac{3}{4}\times\frac{2}{3}$ ACACB → $\frac{1}{3}\times\frac{4}{5}\times\frac{1}{3}\times\frac{1}{5}$ ABACB → $\frac{2}{3}\times\frac{3}{4}\times\frac{1}{3}\times\frac{1}{5}$ ABCAB → $\frac{2}{3}\times\frac{1}{4}\times\frac{4}{5}\times\frac{2}{3}$ $= \frac{1}{300}+\frac{1}{30}+\frac{4}{225}+\frac{1}{30}+\frac{4}{45}=\frac{53}{300}=$ 0.176666. = **0.177** (3 s.f.)	M1 M1 M1 M1 A1 5	At least 4 correct routes chosen 2 correct routes identified and one correct 4-termed product 3 correct products all products correct and added (no other routes) $\frac{53k}{300k}$ or a.r.t. 0.177
7 (iii)	ALITER: (i) $\times\frac{1}{3}\times\frac{1}{5}+$ (ii) $\times\frac{2}{3}+\frac{1}{3}\times\frac{1}{5}\times\frac{1}{4}\times\frac{1}{5}$		