OXFORD CAMBRIDGE AND RSA EXAMINATIONS

Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education

MATHEMATICS   **4732**

Probability & Statistics 1

| Wednesday | **24 MAY 2006** | Afternoon | 1 hour 30 minutes |

Additional materials:
    8 page answer booklet
    Graph paper
    List of Formulae (MF1)

TIME   1 hour 30 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 graphical 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 72.
- 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.

1 Some observations of bivariate data were made and the equations of the two regression lines were found to be as follows.

$$y \text{ on } x: \quad y = -0.6x + 13.0$$
$$x \text{ on } y: \quad x = -1.6y + 21.0$$

    (i) State, with a reason, whether the correlation between x and y is negative or positive.   [1]

    (ii) Neither variable is controlled. Calculate an estimate of the value of x when $y = 7.0$.   [2]

    (iii) Find the values of $\bar{x}$ and $\bar{y}$.   [3]

2 A bag contains 5 black discs and 3 red discs. A disc is selected at random from the bag. If it is red it is replaced in the bag. If it is black, it is not replaced. A second disc is now selected at random from the bag.

Find the probability that

    (i) the second disc is black, given that the first disc was black,   [1]

    (ii) the second disc is black,   [3]

    (iii) the two discs are of different colours.   [3]

3 Each of the 7 letters in the word DIVIDED is printed on a separate card. The cards are arranged in a row.

    (i) How many different arrangements of the letters are possible?   [3]

    (ii) In how many of these arrangements are all three Ds together?   [2]

The 7 cards are now shuffled and 2 cards are selected at random, without replacement.

    (iii) Find the probability that at least one of these 2 cards has D printed on it.   [3]

4   **(i)** The random variable X has the distribution B(25, 0.2). Using the tables of cumulative binomial probabilities, or otherwise, find $P(X \geqslant 5)$.   [2]

    (ii) The random variable Y has the distribution B(10, 0.27). Find $P(Y = 3)$.   [2]

    (iii) The random variable Z has the distribution B(n, 0.27). Find the smallest value of n such that $P(Z \geqslant 1) > 0.95$.   [3]

5 The probability distribution of a discrete random variable, X, is given in the table.

x	0	1	2	3
$P(X = x)$	$\frac{1}{3}$	$\frac{1}{4}$	p	q

It is given that the expectation, $E(X)$, is $1\frac{1}{4}$.

    (i) Calculate the values of p and q.   [5]

    (ii) Calculate the standard deviation of X.   [4]

6 The table shows the total distance travelled, in thousands of miles, and the amount of commission earned, in thousands of pounds, by each of seven sales agents in 2005.

Agent	A	B	C	D	E	F	G
Distance travelled	18	15	12	14	16	24	13
Commission earned	18	45	19	24	27	22	23

(i) (a) Calculate Spearman's rank correlation coefficient, r_s, for these data. [5]

(b) Comment briefly on your value of r_s with reference to this context. [1]

(c) After these data were collected, agent A found that he had made a mistake. He had actually travelled 19 000 miles in 2005. State, with a reason, but without further calculation, whether the value of Spearman's rank correlation coefficient will increase, decrease or stay the same. [2]

The agents were asked to indicate their level of job satisfaction during 2005. A score of 0 represented no job satisfaction, and a score of 10 represented high job satisfaction. Their scores, y, together with the data for distance travelled, x, are illustrated in the scatter diagram below.
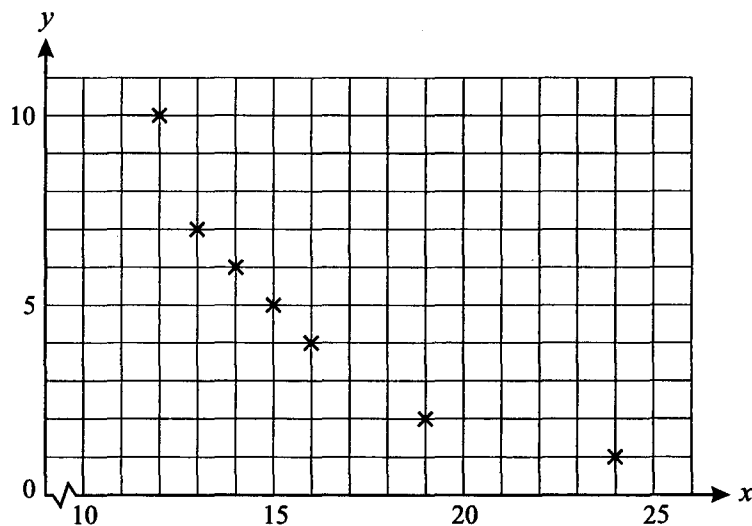
(ii) For this scatter diagram, what can you say about the value of

(a) Spearman's rank correlation coefficient, [1]

(b) the product moment correlation coefficient? [1]

[Questions 7 and 8 are printed overleaf.]

[Turn over

7 In a UK government survey in 2000, smokers were asked to estimate the time between their waking and their having the first cigarette of the day. For heavy smokers, the results were as follows.

Time between waking and first cigarette	1 to 4 minutes	5 to 14 minutes	15 to 29 minutes	30 to 59 minutes	At least 60 minutes
Percentage of smokers	31	27	19	14	9

Times are given correct to the nearest minute.

   (i) Assuming that 'At least 60 minutes' means 'At least 60 minutes but less than 240 minutes', calculate estimates for the mean and standard deviation of the time between waking and first cigarette for these smokers. [6]

   (ii) Find an estimate for the interquartile range of the time between waking and first cigarette for these smokers. Give your answer correct to the nearest minute. [4]

   (iii) The meaning of 'At least 60 minutes' is now changed to 'At least 60 minutes but less than 480 minutes'. Without further calculation, state whether this would cause an increase, a decrease or no change in the estimated value of

   (a) the mean, [1]

   (b) the standard deviation, [1]

   (c) the interquartile range. [1]

8 Henry makes repeated attempts to light his gas fire. He makes the modelling assumption that the probability that the fire will light on any attempt is $\frac{1}{3}$.

Let X be the number of attempts at lighting the fire, up to and including the successful attempt.

   (i) Name the distribution of X, stating a further modelling assumption needed. [2]

In the rest of this question, you should use the distribution named in part **(i)**.

   (ii) Calculate

   (a) $P(X = 4)$, [3]

   (b) $P(X < 4)$. [3]

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

   (iv) Henry has to light the fire once a day, starting on March 1st. Calculate the probability that the first day on which fewer than 4 attempts are needed to light the fire is March 3rd. [3]

Module S1 (4732) May 2006 Mark scheme. Post-Standardisation draft
Note: "(3 sfs)" means "answer which rounds to ... to 3 sfs". If correct ans seen to $\geq$ 3sfs, ISW for later rounding
Penalise 2 sfs only once in <u>paper</u>.

1(i)	Negative, because (grad or coeff of x in 1st equn or x-value or reg coeff or B or –0.6) is negative	B1 **1**	Neg because x incr & y decr
(ii)	$x = -1.6 \times 7.0 + 21$ $x = 9.8$	M1 A1 **2**	Sub $y=7.0$ in 2nd eqn. Allow 1 sign error If sub in both must choose 2nd
(iii)	$y = –0.6(-1.6y + 21) + 13$ or similar $\bar{x} = 5$, $\bar{y} = 10$	M1 A1A1 **3**	Obtain correct eqn in 1 variable. Allow 1 num'l error Allow without bars
Total		**6**	

	In qus 2 & 3 "prod" means "product of two probabilities"		
2(i)	$^4/_7$ or 0.571 (3 sfs)	B1 **1**	
(ii)	$^5/_8 \times {}^4/_7 + {}^3/_8 \times {}^5/_8$ $= {}^{265}/_{448}$ or 0.592 (3 sfs)	M1M1 A1 **3**	M1: one correct prod or add any two prods M1: all correct
(iii)	$^3/_8 \times {}^5/_8 + {}^5/_8 \times {}^3/_7$ $= {}^{225}/_{448}$ or 0.502 (3 sfs)	M1M1 A1 **3**	M1: one correct prod or add any two prods M1: all correct
Total		**7**	

3(i)	$\dfrac{7!}{3! \times 2(!)}$ $= 420$	M1M1 A1 **3**	M1: 7!/(a factorial); or $\dots \div (3! \times 2(!))$ M1: all correct
(ii)	$\dfrac{5!}{2(!)}$ $= 60$	M1 A1 **2**	M1: 5! seen (not part of a C) or $5 \times 4!$ or 120 seen or $\dots \div 2(!)$ alone

(iii)	$1 – {}^4/_7 \times {}^3/_6$ or $1 – {}^4C_2 / {}^7C_2$ or $1 – {}^4P_2 / {}^7P_2$
	    or ${}^3/_7 \times {}^2/_6 + {}^3/_7 \times {}^4/_6 + {}^4/_7 \times {}^3/_6$ oe
	    or ${}^3C_2 / {}^7C_2 + {}^3C_1 x^4C_1 / {}^7C_2$
	
	 $= {}^5/_7$ or 0.714 (3 sfs)
		M1M1
		
		 A1 **3**
			M1:1– prod or $1-\dots/{}^7C_2$ or $1–{}^4C_2 /\dots$ (or Ps)
			    or add 3 prods or add 2 correct prods
			    or ${}^3C_2 / {}^7C_2$ or ${}^3C_1 x^4C_1 / {}^7C_2$
			    or add ≥ 5 out of 7 correct prods
			 M1: all correct

Total		**8**	

4(i)	0.4207 or 0.421 (3 sfs) or $0.8^{25}+25 \times 0.8^{24} \times 0.2 + \dots {}^{25}C_4 \times 0.4^{21} \times 0.2^4$ 0.579(3)	B1 B1 **2**	or $1 – 0.6167$ or 0.3833 (3 sfs) or 1- (6 correct terms, 0 to 5)
(ii)	${}^{10}C_3 \times (1-0.27)^7 \times 0.27^3$ $= 0.261$ (3 sfs)	M1 A1 **2**	
(iii)	$\quad\quad\quad\quad$ Allow "=" thro'out $\quad\quad\quad\quad$ $1 – 0.73^n > 0.95$ $\quad\quad\quad\quad$ or $0.73^n < 0.05$ $0.73^9 = 0.059$ $0.73^{10} = 0.043$ $\quad$ $n\log 0.73 < \log 0.05$ oe $\quad\quad\quad\quad\quad$ $n = 10$	M1 M1 A1 **3**	or $1 - {}^nC_0 \times 0.27^0 \times 0.73^n > 0.95$ oe allow incorrect sign M1 must be correct $\quad\quad\quad\quad$ ft ($1 – 0.27$) from (ii) for M1M1 10 with incorrect sign in wking: SCB2 10 with just $0.73^9 = 0.059$: M1M1A1
Total		**7**	

5(i)	$\frac{1}{3} + \frac{1}{4} + p + q = 1$ oe	B1	
	$0 \times \frac{1}{3} + 1 \times \frac{1}{4} + 2p + 3q = 1\frac{1}{4}$ oe	B1	
	equalize coeffs, eg mult eqn (i) by 2 or 3	M1	allow one error. ft their equns
	Or make p or q subject of (i) or (ii)		subst or subtr not nec'y
	$p = \frac{1}{4}, q = \frac{1}{6}$ oe	A1A1 **5**	
(ii)	$\Sigma x^2 p$ (not /4 or /3 etc) $(= 2\frac{3}{4})$	M1	$\geq$ 2 non-zero terms correct. dep +ve result
	$- \quad (1\frac{1}{4})^2$	M1	indep if +ve result
			or $\Sigma(x-1\frac{1}{4})^2 p$
			($\geq$ 2 (non-0) terms correct): M2
			ft (i) ($0 \leq p, q < 1$) or letters p, q both M1s
	$= 1.1875$ or $1\frac{3}{16}$ oe	A1	cao
	sd $= \sqrt{}$(their 1.1875) $= 1.09$ (3 sfs)	B1f **4**	dep 1st M1 $\&\sqrt{}$(+ve no.) eg $\sqrt{2.75} = 1.66$
Total		**9**	
6(i)(a)	Ranks: 2 4 7 5 3 1 6 6 4 1 3 5 7 2	M1	$\geq$ 5 ranks correct in each set
	7 1 6 3 2 5 4 1 7 2 5 6 3 4	A1	all correct
	Σd^2	M1	dep ranks attempted even if opp orders,
	$(= 60)$		allow arith errors
	$r_s = 1 - \dfrac{6 \times 60}{7 \times 48}$	M1	Correct formula with $n = 7$, dep 2nd M1
			calc r for ranks:
			$S_{xx} = S_{yy} = 140 - 28^2/7$. $S_{xy} = 110 - 28^2/7$
			$(= 28)$ $(= -2)$
			corr subst in one corr S (any version):M1
			corr subst in $r = S_{xy} / \sqrt{(S_{xx}S_{yy})}$:M1
	$= -\frac{1}{14}$ or -0.071 (3 dps)	A1 **5**	
			-0.07 without wking: M1A1M2A0
(b)	Little (or no) connection (agreement,		No mks unless $|r_s| \leq 1$
	rel'nship) between dist and commission		ft their r_s
	Allow disagreement		Must refer to context.
			Not "little corr'n between dist and com"
		B1ft **1**	not "strong disagreement"
			Ignore other comment
(c)	Unchanged. No change in rank	B1B1 **2**	indep
(ii)(a)	$= -1$	B1 **1**	cao
(b)	Close to -1 or, eg $\approx - 0.9$	B1 **1**	not referring to "corr'n" rather than r
			allow "neg", not neg corr'n or neg skew
Total		**10**	

7(i)				Correct (149.5)	With 150	Tot = 2000
	Midpoints attempted ≥ 2 classes	M1				
	Σxf / 100 or Σxf / Σf attempted ≥ 2 terms	M1		2720.5/100	2725/100	Allow Ms
	x within class, not class width					
						& poss As
	Mean = 27.2 (to 3 sfs) (not 27.25)					
	art 27.2 from fully correct wking	A1		27.2	27.25	
				240702.25	242050	
	$\Sigma x^2 f$ or $\Sigma(x - \bar{x})^2 f$ ≥ 2 terms	M1		40.82	40.96	
	$\sqrt{(\Sigma x^2 f /100 - \bar{x}^2)}$ or $\sqrt{((\Sigma x - \bar{x})^2 f /100)}$ or $/\Sigma f$			allow class widths for 2nd M1 only		
	fully corr method, not $\sqrt{\text{neg}}$	M1				
	= 40.5 to 41.1 (3 sfs)	A1	**6**			
(ii)	Recog LQ in 1st class <u>&</u> UQ in 3rd class	B1				
	<u>Graph</u>: | <u>Interp</u>:					
	Attempt 25(.25)th value | LQ = 3.0 to 4.3					
	Attempt 75(.75)th value | UQ =27 to 29	M1		both nec'y		
	Subtract	M1		dep B1or M1		
	IQR = 23 or 24 or 25	A1	**4**	integer. dep M2		
(iii)(a)	Increase	B1	**1**			
(b)	Increase	B1	**1**	Ignore "probably" etc		
(c)	No change	B1	**1**			
Total			**13**			
8(i)	Geometric.	B1		In context. Not "events,. trials, outcomes" . Ignore extra		
	Each attempt (or result or try) indep	B1	**2**			
(ii)(a)	$(^2/_3)^3 \times {}^1/_3$	M2		$(^2/_3)^2 x^1/_3$ or $(^2/_3)^4 \times {}^1/_3$: allow other numerical "p" $(0<p<1)$:M1		
	= $^8/_{81}$ or 0.0988 (3 sfs)	A1	**3**			
(b)	$(^2/_3)^3$	M1		not $(^2/_3)^3$ x ….		
	$1 - (^2/_3)^3$	M1		or $^1/_3 + {}^2/_3 x^1/_3 + (^2/_3)^2 x^1/_3$ M2		
				$1 - (^2/_3)^4$ or $1 - ("q")^4$ M1		
				or 3 terms, with 2 correct M1		
				or 3 correct terms + 1 extra M1		
				or "p" + "qp" + "$q^2 p$" M1		
				or 1 – sum of 3 correct terms M1		
	= $^{19}/_{27}$ or 0.704 (3sfs)	A1	**3**	"p" means num value, not $^1/_3$		
(iii)	3	B1f	**1**	or $^1/_{"p"}$		
(iv)	$1 - {}^{19}/_{27}$ | (1 – 0.7037) or 0.2963	M1		ft (b) for M1M1 must see method if ft		
	$(^8/_{27})^2 \times {}^{19}/_{27}$ | 0.2963^2 x 0.7037	M1		Allow figs rounded to 2 sfs for M1M1		
	= $^{1216}/_{19683}$ | = 0.0618 (3 sfs)	A1	**3**	cao. allow art 0.0618 or 0.0617		
Total			**12**			

Total 72 marks