OCR
RECOGNISING ACHIEVEMENT

ADVANCED SUBSIDIARY GCE UNIT **4732/01**

MATHEMATICS

Probability & Statistics 1

TUESDAY 5 JUNE 2007 Afternoon

Time: 1 hour 30 minutes

Additional Materials: Answer Booklet (8 pages)
List of Formulae (MF1)

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.

ADVICE TO CANDIDATES

- Read each question carefully and make sure you know what you have to do before starting your answer.
- **You are reminded of the need for clear presentation in your answers.**

This document consists of **6** printed pages and **2** blank pages.

[Turn over

1 The table shows the probability distribution for a random variable X.

x	0	1	2	3
$P(X = x)$	0.1	0.2	0.3	0.4

Calculate $E(X)$ and $Var(X)$. [5]

2 Two judges each placed skaters from five countries in rank order.

Position	1st	2nd	3rd	4th	5th
Judge 1	UK	France	Russia	Poland	Canada
Judge 2	Russia	Canada	France	UK	Poland

Calculate Spearman's rank correlation coefficient, r_s, for the two judges' rankings. [5]

3 **(i)** How many different teams of 7 people can be chosen, without regard to order, from a squad of 15? [2]

(ii) The squad consists of 6 forwards and 9 defenders. How many different teams containing 3 forwards and 4 defenders can be chosen? [2]

4 A bag contains 6 white discs and 4 blue discs. Discs are removed at random, one at a time, **without** replacement.

(i) Find the probability that

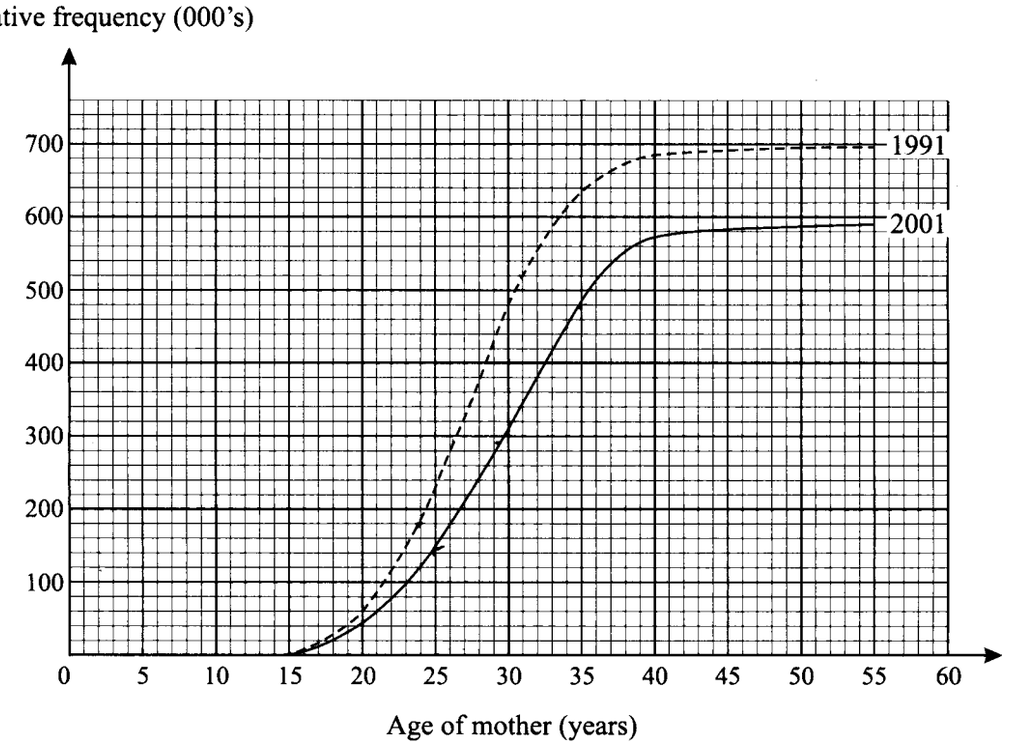
(a) the second disc is blue, given that the first disc was blue, [1]

(b) the second disc is blue, [3]

(c) the third disc is blue, given that the first disc was blue. [3]

(ii) The random variable X is the number of discs which are removed up to and including the first blue disc. State whether the variable X has a geometric distribution. Explain your answer briefly. [1]

5 The numbers of births, in thousands, to mothers of different ages in England and Wales, in 1991 and 2001 are illustrated by the cumulative frequency curves.



(i) In which of these two years were there more births? How many more births were there in this year? [2]

(ii) The following quantities were estimated from the diagram.

Year	Median age (years)	Interquartile range (years)	Proportion of mothers giving birth aged below 25	Proportion of mothers giving birth aged 35 or above
1991	27.5	7.3	33%	9%
2001				18%

(a) Find the values missing from the table. [5]

(b) Did the women who gave birth in 2001 tend to be younger or older or about the same age as the women who gave birth in 1991? Using the table and your values from part **(a)**, give two reasons for your answer. [3]

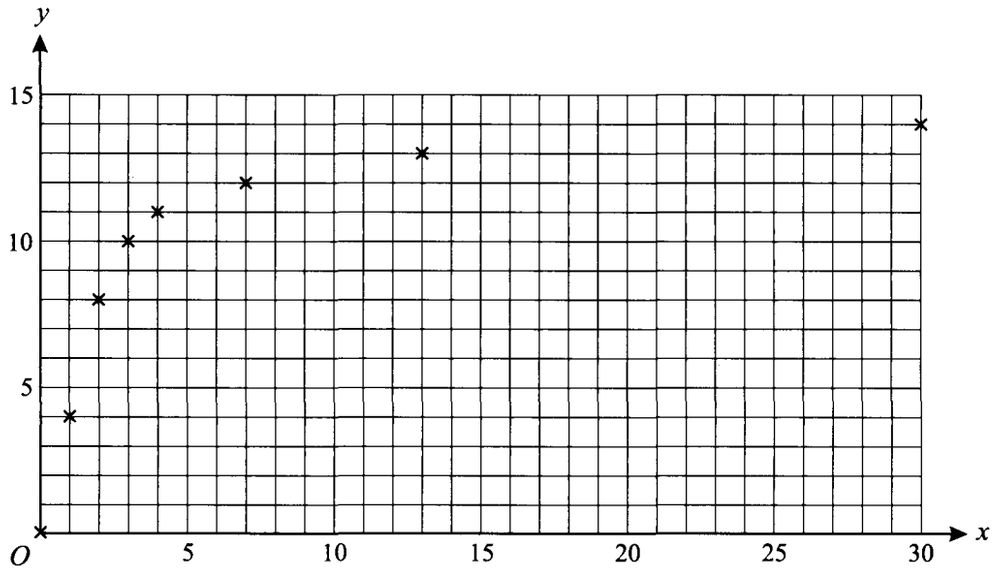
[Turn over

6 A machine with artificial intelligence is designed to improve its efficiency rating with practice. The table shows the values of the efficiency rating, y, after the machine has carried out its task various numbers of times, x.

x	0	1	2	3	4	7	13	30
y	0	4	8	10	11	12	13	14

$[n = 8,\ \Sigma x = 60,\ \Sigma y = 72,\ \Sigma x^2 = 1148,\ \Sigma y^2 = 810,\ \Sigma xy = 767.]$

These data are illustrated in the scatter diagram.



(i) (a) Calculate the value of r, the product moment correlation coefficient. [3]

(b) Without calculation, state with a reason the value of r_s, Spearman's rank correlation coefficient. [2]

(ii) A researcher suggests that the data for $x = 0$ and $x = 1$ should be ignored. Without calculation, state with a reason what effect this would have on the value of

(a) r, [2]

(b) r_s. [2]

(iii) Use the diagram to estimate the value of y when $x = 29$. [1]

(iv) Jack finds the equation of the regression line of y on x for all the data, and uses it to estimate the value of y when $x = 29$. Without calculation, state with a reason whether this estimate or the one found in part **(iii)** will be the more reliable. [2]

7 On average, 25% of the packets of a certain kind of soup contain a voucher. Kim buys one packet of soup each week for 12 weeks. The number of vouchers she obtains is denoted by X.

(i) State two conditions needed for X to be modelled by the distribution B(12, 0.25). [2]

In the rest of this question you should assume that these conditions are satisfied.

(ii) Find $P(X \leqslant 6)$. [2]

In order to claim a free gift, 7 vouchers are needed.

(iii) Find the probability that Kim will be able to claim a free gift at some time during the 12 weeks. [1]

(iv) Find the probability that Kim will be able to claim a free gift in the 12th week but not before. [4]

8 **(i)** A biased coin is thrown twice. The probability that it shows heads both times is 0.04. Find the probability that it shows tails both times. [3]

(ii) Another coin is biased so that the probability that it shows heads on any throw is p. The probability that the coin shows heads exactly once in two throws is 0.42. Find the two possible values of p. [5]

9 **(i)** A random variable X has the distribution Geo$(\frac{1}{5})$. Find

(a) $E(X)$, [2]

(b) $P(X = 4)$, [2]

(c) $P(X > 4)$. [2]

(ii) A random variable Y has the distribution Geo(p), and $q = 1 - p$.

(a) Show that $P(Y \text{ is odd}) = p + q^2p + q^4p + \ldots$. [1]

(b) Use the formula for the sum to infinity of a geometric progression to show that

$$P(Y \text{ is odd}) = \frac{1}{1+q}.$$ [4]

Note: "3 sfs" means an answer which is equal to, or rounds to, the given answer. If such an answer is seen and then later rounded, apply ISW.

1	$(0\times0.1) + 1\times0.2 + 2\times0.3 + 3\times0.4$ $= 2(.0)$ $(0^2\times0.1) +1\times0.2 + 2^2\times0.3 + 3^2\times0.4$ (= 5) $- 2^2$ $= 1$	M1 A1 M1 M1 A1 5	$\geq$ 2 non-zero terms correct eg ÷ 4: M0 $\geq$ 2 non-zero terms correct ÷ 4: M0 Indep, ft their μ. Dep +ve result $(-2)^2\times0.1+(-1)^2\times0.2+0^2\times0.3+1^2\times0.4$:M2 $\geq$ 2 non-0 correct: M1 ÷ 4: M0
Total		**5**	
2	UK Fr Ru Po Ca 1 2 3 4 5 or 5 4 3 2 1 4 3 1 5 2 2 3 5 1 4 Σd^2 (= 24) $r_s = 1 - \frac{6\times\text{"24"}}{5\times(5^2-1)}$ $= -\frac{1}{5}$ or -0.2	M1 A1 M1 M1 A1 5	Consistent attempt rank other judge RCFUP 3 5 2 1 4 3 1 4 5 2 1 2 3 4 5 5 4 3 2 1 All 5 d^2 attempted & added. Dep ranks att'd Dep 2nd M1 $\frac{43-15^2/5}{\sqrt{((55-15^2/5)(55-15^2/5))}}$ Corr sub in $\geq$ 2 S's M1 All correct: M1
Total		**5**	
3i	${}^{15}C_7$ or $\frac{15!}{7!8!}$ 6435	M1 A1 2	
ii	${}^6C_3 \times {}^9C_4$ or $\frac{6!}{3!3!}\times\frac{9!}{4!5!}$ 2520	M1 A1 2	Alone except allow ÷ ${}^{15}C_7$ Or ${}^6P_3 \times {}^9P_4$ or $\frac{6!}{3!}\times\frac{9!}{5!}$ Allow ÷ ${}^{15}P_7$ NB not $\frac{6!}{3!}\times\frac{9!}{4!}$ 362880
Total		**4**	
4ia	$\frac{1}{3}$ oe	B1 1	B↔W MR: max (a)B0(b)M1M1(c)B1M1
b	P(BB) + P(WB) attempted $= \frac{4}{10}\times\frac{3}{9} + \frac{6}{10}\times\frac{4}{9}$ or $\frac{2}{15} + \frac{4}{15}$ $= \frac{2}{5}$ oe	M1 M1 A1 3	Or $\frac{4}{10}\times\frac{3}{9}$ OR $\frac{6}{10}\times\frac{4}{9}$ correct NB $\frac{4}{10}\times\frac{4}{10} + \frac{6}{10}\times\frac{4}{10} = \frac{2}{5}$: M1M0A0
c	Denoms 9 & 8 seen or implied $\frac{3}{9}\times\frac{2}{8} + \frac{6}{9}\times\frac{3}{8}$ $= \frac{1}{3}$ oe	B1 M1 A1 3	Or $\frac{2}{15}$ as numerator Or $\frac{\frac{2}{15}}{\frac{4}{10}}$ Or $\frac{\frac{4}{10}\times\frac{6}{9}\times\frac{3}{8} + \frac{4}{10}\times\frac{3}{9}\times\frac{2}{8}}{\text{above} + \frac{6}{10}\times\frac{5}{9}\times\frac{4}{8} + \frac{6}{10}\times\frac{4}{9}\times\frac{3}{8}}$ May not see wking
ii	P(Blue) not constant or discs not indep, so no	B1 1	Prob changes as discs removed Limit to no. of discs. Fixed no. of discs Discs will run out Context essential: "disc" or "blue" NOT fixed no. of trials NOT because without repl Ignore extra
Total		**8**	

5i	1991 100 000 to 110 000	B1 ind B1 ind 2	Or fewer in 2001 Allow digits100 to 110
iia	Median = 29 to 29.9 Quartiles 33 to 34, 24.5 to 26 = 7.5 to 9.5 140 to 155 23 to 26.3%	B1 M1 A1 M1 A1 5	 Or one correct quartile and subtr NOT from incorrect wking ×1000, but allow without Rnded to 1 dp or integer 73.7 to 77% : SC1
b	Older Median (or ave) greater } % older mothers greater oe} % younger mothers less oe}	B1 B1 B1 3	Or 1991 younger Any two Or 1991 steeper so more younger: B2 NOT mean gter Ignore extra
Total		**10**	

6ia	Correct subst in ≥ two S formulae $\dfrac{767-\frac{60\times72}{8}}{\sqrt{(1148-\frac{60^2}{8})(810-\frac{72^2}{8})}}$ or $\dfrac{227}{\sqrt{698}\sqrt{162}}$ = 0.675 (3 sfs)	M1 M1 A1 3	Any version All correct. Or $\dfrac{767\text{-}8x7.5x9}{\sqrt{((1148\text{-}8x7.5^2)(810\text{-}8x9^2))}}$ or correct substn in any correct formula for r
b	1 y always increases with x or ranks same oe	B1 B1 2	+ve grad thro'out. Increase in steps. Same order. Both ascending order Perfect RANK corr'n Ignore extra NOT Increasing proportionately
iia	Closer to 1, or increases because nearer to st line	B1 B1 2	Corr'n stronger. Fewer outliers. "They" are outliers Ignore extra
b	None, or remains at 1 Because y still increasing with x oe	B1 B1 2	Σd^2 still 0. Still same order. Ignore extra NOT differences still the same. NOT ft (i)(b)
iii	13.8 to 14.0	B1 1	
iv	(iii) or graph or diag or my est Takes account of curve	B1 B1 2	Must be clear which est. Can be implied. "This est" probably ⇒ using equn of line Straight line is not good fit. Not linear. Corr'n not strong.
Total		**12**	
7i	P(contains voucher) constant oe Packets indep oe	B1 B1 2	Context essential NOT vouchers indep
ii	0.9857 or 0.986 (3 sfs)	B2 2	B1 for 0.9456 or 0.946 or 0.997(2) or for 7 terms correct, allow one omit or extra NOT 1 – 0.9857 = 0.0143 (see (iii))
iii	(1 – 0.9857) = 0.014(3) (2 sfs)	 B1ft 1	Allow 1- their (ii) correctly calc'd
iv	B(11, 0.25) or 6 in 11 wks stated or impl $^{11}C_6 \times 075^5 \times 0.25^6$ (= 0.0267663) P(6 from 11) × 0.25 = 0.00669 or 6.69 x 10^{-3} (3 sfs)	B1 M1 M1 A1 4	or $0.75^a \times 0.25^b$ ($a + b = 11$) or $^{11}C_6$ dep B1
Total		**9**	

8i	$\sqrt{0.04}$ (= 0.2) $(1 - \text{their } \sqrt{0.04})^2$ = 0.64	M1 M1 A1 3	
ii	$1 - p$ seen $2p(1-p) = 0.42$ or $p(1-p) = 0.21$ oe $2p^2 - 2p + 0.42(= 0)$ or $p^2 - p + 0.21(= 0)$ $\frac{2 \pm \sqrt{((-2)^2 - 4\times0.42)}}{2\times2}$ or $\frac{1 \pm \sqrt{((-1)^2 - 4\times0.21)}}{2\times1}$ or $(p-0.7)(p-0.3)=0$ or $(10p-7)(10p-3)=0$ $p = 0.7$ or 0.3	B1 M1 M1 M1 A1 5	M1 for either $2pq = 0.42$ or $pq = 0.21$ Allow $pq=0.42$ or opp signs, correct terms any order (= 0) oe Correct Dep B1M1M1 Any corr subst'n or fact'n Omit 2 in 2nd line: max B1M1M0M0A0 One corr ans with no or inadeq wking: SC1 eg $0.6 \times 0.7 = 0.42 \Rightarrow p = 0.7$ or 0.6 $p^2 + 2pq + q^2 = 1$ B1 $p^2 + q^2 = 0.58$ } $p = 0.21/q$ } $p^4 - 0.58p^2 + 0.0441 = 0$ M1 corr subst'n or fact'n M1 $1 - p$ seen B1 $2p(1-p) = 0.42$ or $p(1-p) = 0.21$ M1 $p^2 - p = -0.21$ $p^2 - p + 0.25 = -0.21 + 0.25$ oe } M1 OR $(p-0.5)^2 - 0.25 = -0.21$ oe } $(p-0.5)^2 = 0.04$ M1 $(p-0.5) = \pm0.02$ $p = 0.3$ or 0.7 A1
Total		**8**	
9ia	$1 / {}^1/_5$ = 5	M1 A1 2	
b	$({}^4/_5)^3 \times {}^1/_5$ $= {}^{64}/_{625}$ or 0.102 (3 sfs)	M1 A1 2	
c	$({}^4/_5)^4$ $= {}^{256}/_{625}$ or a.r.t 0.410 (3 sfs) or 0.41	M1 A1 2	or $1 - ({}^1/_5 + {}^4/_5\times{}^1/_5 + ({}^4/_5)^2\times{}^1/_5 + ({}^4/_5)^3\times{}^1/_5)$ NOT $1 - ({}^4/_5)^4$
iia	$P(Y=1) = p$, $P(Y=3) = q^2p$, $P(Y=5) = q^4p$	B1 1	$P(Y=1)+P(Y=3)+P(Y=5) = p + q^2p + q^4p$ p, $p(1-p)^2$, $p(1-p)^4$ q^{1-1}, q^{3-1}, q^{5-1} or any of these with $1-p$ instead of q "Always q to even power $\times$ p" Either associate each term with relevant prob Or give indication of how terms derived $\geq$ two terms
b	Recog that c.r. = q^2 or $(1-p)^2$ $S_\infty = \frac{p}{1-q^2}$ or $\frac{p}{1-(1-p)^2}$ $P(\text{odd}) = \frac{1-q}{1-q^2}$ $= \frac{1-q}{(1-q)(1+q)}$ Must see this step for A1 $(= \frac{1}{1+q}$ **AG**)	M1 M1 M1 A1 4	or eg $r = q^2p/p$ $(= \frac{p}{2p-p^2}) = \frac{p}{p(2-p)}$ $(= \frac{1}{2-p}) = \frac{1}{2-(1-q)}$