OCR
RECOGNISING ACHIEVEMENT

ADVANCED SUBSIDIARY GCE

MATHEMATICS

Probability & Statistics 1

4732

Candidates answer on the Answer Booklet

OCR Supplied Materials:
- 8 page Answer Booklet
- List of Formulae (MF1)

Other Materials Required:
None

Wednesday 27 January 2010
Afternoon

Duration: 1 hour 30 minutes

* 4 7 3 2 *

INSTRUCTIONS TO CANDIDATES

- Write your name clearly in capital letters, your Centre Number and Candidate Number in the spaces provided on the Answer Booklet.
- Use black ink. Pencil may be used for graphs and diagrams only.
- Read each question carefully and make sure that you know what you have to do before starting your answer.
- Answer **all** the questions.
- Do **not** write in the bar codes.
- 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.
- **You are reminded of the need for clear presentation in your answers.**
- The total number of marks for this paper is **72**.
- This document consists of **8** pages. Any blank pages are indicated.

1 Andy makes repeated attempts to thread a needle. The number of attempts up to and including his first success is denoted by X.

(i) State two conditions necessary for X to have a geometric distribution. **[2]**

(ii) Assuming that X has the distribution Geo(0.3), find

(a) $P(X = 5)$, **[2]**

(b) $P(X > 5)$. **[3]**

(iii) Suggest a reason why one of the conditions you have given in part **(i)** might not be satisfied in this context. **[2]**

2 40 people were asked to guess the length of a certain road. Each person gave their guess, l km, correct to the nearest kilometre. The results are summarised below.

l	10–12	13–15	16–20	21–30
Frequency	1	13	20	6

(i) **(a)** Use appropriate formulae to calculate estimates of the mean and standard deviation of l. **[6]**

(b) Explain why your answers are only estimates. **[1]**

(ii) A histogram is to be drawn to illustrate the data. Calculate the frequency density of the block for the 16–20 class. **[2]**

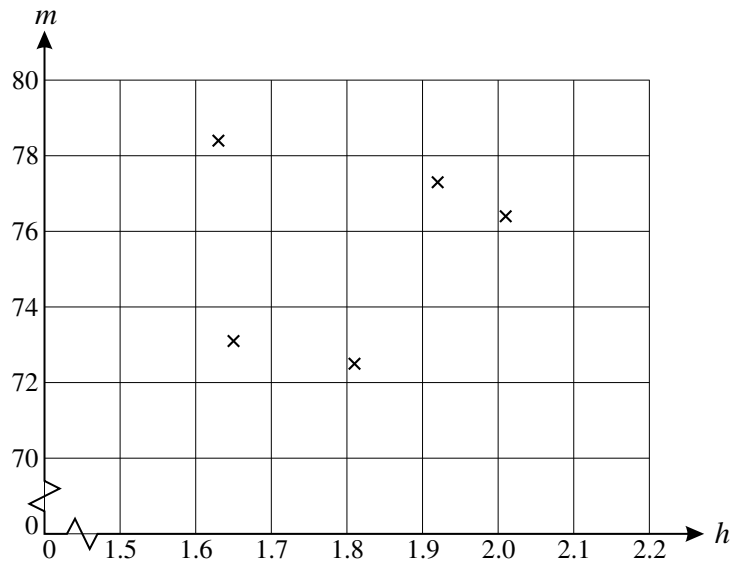
(iii) Explain which class contains the median value of l. **[2]**

(iv) Later, the person whose guess was between 10 km and 12 km changed his guess to between 13 km and 15 km. Without calculation state whether the following will increase, decrease or remain the same:

(a) the mean of l, **[1]**

(b) the standard deviation of l. **[1]**

3 The heights, h m, and weights, m kg, of five men were measured. The results are plotted on the diagram.



The results are summarised as follows.

$$n = 5 \quad \Sigma h = 9.02 \quad \Sigma m = 377.7 \quad \Sigma h^2 = 16.382 \quad \Sigma m^2 = 28\,558.67 \quad \Sigma hm = 681.612$$

(i) Use the summarised data to calculate the value of the product moment correlation coefficient, r. **[3]**

(ii) Comment on your value of r in relation to the diagram. **[2]**

(iii) It was decided to re-calculate the value of r after converting the heights to feet and the masses to pounds. State what effect, if any, this will have on the value of r. **[1]**

(iv) One of the men had height 1.63 m and mass 78.4 kg. The data for this man were removed and the value of r was re-calculated using the original data for the remaining four men. State in general terms what effect, if any, this will have on the value of r. **[1]**

4 A certain four-sided die is biased. The score, X, on each throw is a random variable with probability distribution as shown in the table. Throws of the die are independent.

x	0	1	2	3
$P(X = x)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$

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

The die is thrown 10 times.

(ii) Find the probability that there are not more than 4 throws on which the score is 1. **[2]**

(iii) Find the probability that there are exactly 4 throws on which the score is 2. **[3]**

5 A washing-up bowl contains 6 spoons, 5 forks and 3 knives. Three of these 14 items are removed at random, without replacement. Find the probability that

(i) all three items are of different kinds, **[3]**

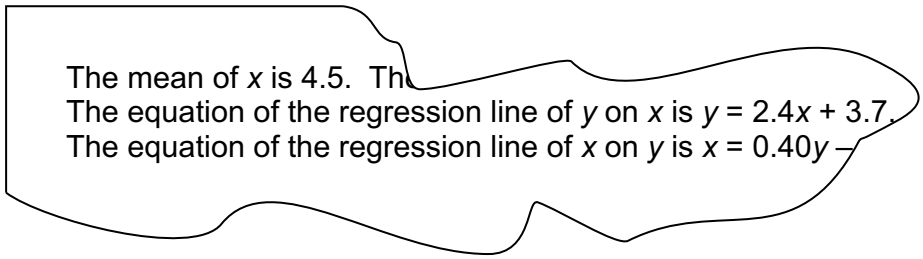
(ii) all three items are of the same kind. **[3]**

6 **(a)** A student calculated the values of the product moment correlation coefficient, r, and Spearman's rank correlation coefficient, r_s, for two sets of bivariate data, A and B. His results are given below.

$$A: \quad r = 0.9 \text{ and } r_s = 1$$
$$B: \quad r = 1 \quad \text{and } r_s = 0.9$$

With the aid of a diagram where appropriate, explain why the student's results for A could both be correct but his results for B cannot both be correct. **[3]**

(b) An old research paper has been partially destroyed. The surviving part of the paper contains the following incomplete information about some bivariate data from an experiment.



Calculate the missing constant at the end of the equation of the second regression line. **[4]**

7 The table shows the numbers of male and female members of a vintage car club who own either a Jaguar or a Bentley. No member owns both makes of car.

	Male	Female
Jaguar	25	15
Bentley	12	8

One member is chosen at random from these 60 members.

(i) Given that this member is male, find the probability that he owns a Jaguar. **[2]**

Now two members are chosen at random from the 60 members. They are chosen one at a time, without replacement.

(ii) Given that the first one of these members is female, find the probability that both own Jaguars. **[4]**

8 The five letters of the word NEVER are arranged in random order in a straight line.

(i) How many different orders of the letters are possible? **[2]**

(ii) In how many of the possible orders are the two Es next to each other? **[2]**

(iii) Find the probability that the first two letters in the order include exactly one letter E. **[3]**

9 R and S are independent random variables each having the distribution Geo(p).

(i) Find P($R = 1$ and $S = 1$) in terms of p. **[1]**

(ii) Show that P($R = 3$ and $S = 3$) $= p^2q^4$, where $q = 1 - p$. **[1]**

(iii) Use the formula for the sum to infinity of a geometric series to show that

$$\mathrm{P}(R = S) = \frac{p}{2 - p}.$$ **[5]**

4732 Probability & Statistics 1

Note: "(3 sfs)" means "answer which rounds to ... to 3 sfs". If correct ans seen to ≥ 3sfs, ISW for later rounding
Penalise over-rounding only once in paper.

1 (i)	attempts at threading indep prob of succeeding in threading const	B1 B1 2	in context in context
(ii) (a)	$0.7^4 \times 0.3$ = 0.0720 (3sf)	M1 A1 2	 Condone 0.072
(b)	0.7^5 = 0.168 (3 sfs)	M2 A1 3	or 1-(0.3+0.7×0.3+0.7^2×0.3+0.7^3×0.3 +0.7^4×0.3) M1 for one term omitted or extra or wrong or 1-0.7^5 or(0.3+...+0.7^4×0.3) or 0.3, 0.7 muddle or 0.7^4 or 0.7^6 alone. 0.6 not 0.7 M0 in **(a)** M1 in **(b)** 1/3,2/3 used M1in **(a)** M1 in **(b)**
(iii)	likely to improve with practice hence independence unlikely or prob will increase each time	B1 B1 2	or thread strands gradually separate 1st B1 must be in context. hence independence unlikely or prob will decrease each time or similar Allow 'change'
Total		**[9]**	
2 (i) (a)	Use of correct midpts $\Sigma lf \div \Sigma f$ (= 706 ÷ 40) = 17.65 $\Sigma l^2 f$ (= 13050.5) $\sqrt{\frac{"13050.5"}{40} - "17.65"^2}$ (= √14.74) = 3.84 (3 sfs)	B1 M1 A1 M1 M1 A1 6	11,14,18,25.5 l within class, ≥ three lf seen [17.575,17.7] ≥ three l^2f seen ÷40,-mean²,√.Dep>0. ∑(l-17.65)²f, at least 3 M1,÷40,√ M1,3.84 A1. ÷ 4 ⇒ max B1M0A0M1M0A0
(b)	mid pts used or data grouped or exact values unknown oe	 B1 1	not "orig values were guesses"
(ii)	20 ÷ 5 = 4	M1 A1 2	condone 20 ÷ [4,5] or ans 5
(iii)	20.5th value requ'd and 1st two classes contain 14 values 16 – 20	 M1 B1 2	condone 20th oe or third class oe
(iv) (a)	increase	B1 1	
(b)	decrease	B1 1	
Total		**[13]**	
3 (i)	S_{hm} = 0.2412 S_{hh} = 0.10992 S_{mm} = 27.212 $r = \frac{S_{hm}}{\sqrt{(S_{hh}S_{mm})}}$ = 0.139 (3 sfs)	 B1 M1 A1 3	Allow x or ÷ 5 any one S correct ft their Ss
(ii)	Small, low or not close to 1 or close to 0 oe pts not close to line oe	B1 ft B1	1st B1 about value of r 2nd B1 about diag
(iii)	none or unchanged or "0.139" oe	B1 1	
(iv)	Larger oe	B1 1	
Total		**[7]**	

4 (i)	$(0\times\frac{1}{2})+1\times\frac{1}{4}+2\times\frac{1}{8}+3\times\frac{1}{8}$ $=\frac{7}{8}$ or 0.875 oe $(0\times\frac{1}{2})+1\times\frac{1}{4}+2^2\times\frac{1}{8}+3^2\times\frac{1}{8}$ (= $1\frac{7}{8}$) $-(\text{“}\frac{7}{8}\text{”})^2$ $=\frac{71}{64}$ or 1.11 (3 sfs) oe	M1 A1 M1 M1 A1 5	$\geq$ 2 non-zero terms seen If ÷3 or 4 M0M0M1(poss) $\geq$ 2 non-zero terms seen dep +ve result M1 all4 $(x-0.875)^2$ terms seen. M1 mult p,$\sum$ A1 1.11
(ii)	Bin stated or implied 0.922 (3 sfs)	M1 A1 2	Eg table or $\frac{1}{4}^n\times\frac{3}{4}^m$ (n+m=10,n,m≠ 1) or10C4 or 5(or 4 or 6) terms correct
(iii)	n = 10 & $p=\frac{1}{8}$ stated or implied $^{10}C_4\times\frac{7}{8}^6\times\frac{1}{8}^4$ = 0.0230 (3 sfs)	M1 M1 A1 3	condone 0.023
Total		**[10]**	
5 (i)	$\frac{6}{14}\times\frac{5}{13}\times\frac{3}{12}$ × 3! oe $=\frac{45}{182}$ or 0.247 (3 sfs)oe	M1 M1 A1 3	$^6C_1\times{}^5C_1\times{}^3C_1$ $\div{}^{14}C_3$ With repl M0M1A0
(ii)	$\frac{6}{14}\times\frac{5}{13}\times\frac{4}{12}+\frac{5}{14}\times\frac{4}{13}\times\frac{3}{12}+\frac{3}{14}\times\frac{2}{13}\times\frac{1}{12}$ $=\frac{31}{364}$ or 0.0852 (3 sf)	M2 A1 3	$^6C_3+{}^5C_3+{}^3C_3$ M1 for any one ($\div{}^{14}C_3$)M1 all 9 numerators correct. With repl M1$(6/14)^3+(5/14)^3+(3/14)^3$
Total		**[6]**	
6 (a)	A: diag or explanation showing pts close to st line, always increasing B:Diag or expl based on r=1=>pts on st line =>r(s)=1	B1 B1 B1 3	. Diag or expl based on r(s)≠ 1=>pts not on st line =>r≠ 1 r=1=>pts on st line&r(s)≠ 1=>pts not on st line B1B1 r=1=>r(s)=1 B2
(b)	$\bar{y}$ = 2.4 × 4.5 + 3.7 = 14.5 4.5 = 0.4 × “14.5”- c c = 1.3 a’=x-b’y :-14.5 M1A1; then a’=4.5-0.4x14.5=-1.3 M1A1	M1 A1 M1 A1 4	Attempt to sub expression for y x=0.96x+1.48-c oe sub x=4.5 and solve c=1.3 14.5 M1A1.(y-3.7)/2.4=0.4y-c and sub14.5 M1 c=1.3 A1
Total		**[7]**	
7 (i)	$^{25}/_{37}$	B2 2	B1 num, B1 denom 25/37xp B1
(ii)	$\frac{15}{23}$ seen or implied $\times\frac{39}{59}$ seen or implied $=\frac{585}{1357}$ or 0.431 (3 sfs) oe	M1 M2 A1 4	M1 num, M1 denom Allow M1 for 39/59x or + wrong p
Total		**[6]**	

8 (i)	$^{5!}/_{2}$ = 60	M1 A1 2	Allow 5P3
(ii)	4! = 24	M1 A1 2	Allow 2×4!
(iii)	$^{2}/_{5} \times {}^{3}/_{4}$ or3/5×2/4 × 2 = $^{3}/_{5}$ oe	M1 M1 A1 3	allow M1 for $^{2}/_{5} \times {}^{3}/_{5} \times 2$ or $^{12}/_{25}$ or (6×3!)÷**(i)** M2 or 3!÷**(i),**6÷**(i)**,(6+6)÷**(i)**,6k÷**(i)** or 6×6 or 36 or 1-correct answer M1 (k,integer ≤ 5)
Total		**[7]**	
9 (i)	p^2	B1 1	
(ii)	$(q^2p)^2$ oe =AG	B1 1	
(iii)	r=q² a/(1-r) used $(S_\infty =)\ \dfrac{p^2}{1-q^2}$ $= \dfrac{p^2}{1-(1-p)^2}$ p/(2-p) AG	B1 M1 A1 M1 A1 5	May be implied With a=p²and r=q²or q ⁴ Attempt to simplify using p+q=1 correctly. Dep on r = q² or q⁴ $\dfrac{(1-q)^2}{(1-q)(1+q)}$ or p²/p(1+q) Correctly obtain given answer showing at least one intermediate step.
P2Total		**[7]**	

Total 72 marks