OCR
RECOGNISING ACHIEVEMENT

ADVANCED SUBSIDIARY GCE UNIT **4732/01**

MATHEMATICS

Probability & Statistics 1

FRIDAY 12 JANUARY 2007 Morning

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.

1 Part of the probability distribution of a variable, X, is given in the table.

x	0	1	2	3
$P(X = x)$		$\frac{3}{10}$	$\frac{1}{5}$	$\frac{2}{5}$

(i) Find $P(X = 0)$. [2]

(ii) Find $E(X)$. [2]

2 The table contains data concerning five households selected at random from a certain town.

Number of people in the household	2	3	3	5	7
Number of cars belonging to people in the household	1	1	3	2	4

(i) Calculate the product moment correlation coefficient, r, for the data in the table. [5]

(ii) Give a reason why it would not be sensible to use your answer to draw a conclusion about all the households in the town. [1]

3 The digits 1, 2, 3, 4 and 5 are arranged in random order, to form a five-digit number.

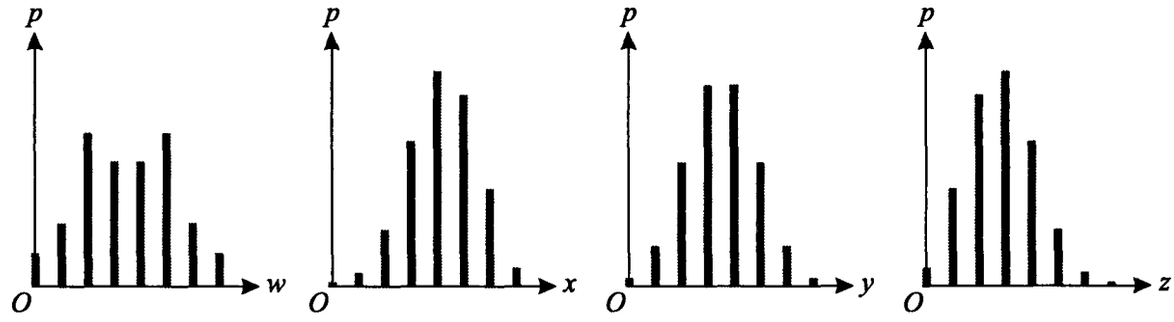
(i) How many different five-digit numbers can be formed? [1]

(ii) Find the probability that the five-digit number is

(a) odd, [2]

(b) less than 23 000. [3]

4 Each of the variables W, X, Y and Z takes eight integer values only. The probability distributions are illustrated in the following diagrams.



(i) For which one or more of these variables is

(a) the mean equal to the median, [1]

(b) the mean greater than the median? [1]

(ii) Give a reason why **none** of these diagrams could represent a geometric distribution. [1]

(iii) Which one of these diagrams could **not** represent a binomial distribution? Explain your answer briefly. [2]

5 A chemical solution was gradually heated. At five-minute intervals the time, x minutes, and the temperature, $y\,^{\circ}\mathrm{C}$, were noted.

x	0	5	10	15	20	25	30	35
y	0.8	3.0	6.8	10.9	15.6	19.6	23.4	26.7

[$n = 8$, $\Sigma x = 140$, $\Sigma y = 106.8$, $\Sigma x^2 = 3500$, $\Sigma y^2 = 2062.66$, $\Sigma xy = 2685.0$.]

(i) Calculate the equation of the regression line of y on x. [4]

(ii) Use your equation to estimate the temperature after 12 minutes. [2]

(iii) It is given that the value of the product moment correlation coefficient is close to +1. Comment on the reliability of using your equation to estimate y when

(a) $x = 17$,

(b) $x = 57$.

[2]

6 A coin is biased so that the probability that it will show heads on any throw is $\frac{2}{3}$. The coin is thrown repeatedly.

The number of throws up to and including the first head is denoted by X. Find

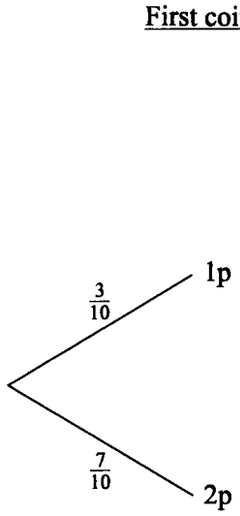
(i) $P(X = 4)$, [3]

(ii) $P(X < 4)$, [3]

(iii) $E(X)$. [2]

7 A bag contains three 1p coins and seven 2p coins. Coins are removed at random one at a time, **without** replacement, until the total value of the coins removed is **at least** 3p. Then no more coins are removed.

(i) Copy and complete the probability tree diagram. [5]



Find the probability that

(ii) exactly two coins are removed, [3]

(iii) the total value of the coins removed is 4p. [3]

8 In the 2001 census, the household size (the number of people living in each household) was recorded. The percentages of households of different sizes were then calculated. The table shows the percentages for two wards, Withington and Old Moat, in Manchester.

	Household size						
	1	2	3	4	5	6	7 or more
Withington	34.1	26.1	12.7	12.8	8.2	4.0	2.1
Old Moat	35.1	27.1	14.7	11.4	7.6	2.8	1.3

(i) Calculate the median and interquartile range of the household size for Withington. [3]

(ii) Making an appropriate assumption for the last class, which should be stated, calculate the mean and standard deviation of the household size for Withington. Give your answers to an appropriate degree of accuracy. [6]

The corresponding results for Old Moat are as follows.

Median	Interquartile range	Mean	Standard deviation
2	2	2.4	1.5

(iii) State one advantage of using the median rather than the mean as a measure of the average household size. [1]

(iv) By comparing the values for Withington with those for Old Moat, explain briefly why the interquartile range may be less suitable than the standard deviation as a measure of the variation in household size. [1]

(v) For one of the above wards, the value of Spearman's rank correlation coefficient between household size and percentage is -1. Without any calculation, state which ward this is. Explain your answer. [2]

9 A variable X has the distribution $B(11, p)$.

(i) Given that $p = \frac{3}{4}$, find $P(X = 5)$. [2]

(ii) Given that $P(X = 0) = 0.05$, find p. [4]

(iii) Given that $\text{Var}(X) = 1.76$, find the two possible values of p. [5]

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. Penalize over-rounding only once in paper, except qu 8(ii).

1i	$1 - (^3/_{10} + {}^1/_5 + {}^2/_5)$ $^1/_{10}$	M1 A1 2	or $(^3/_{10} + {}^1/_5 + {}^2/_5) + p = 1$
ii	$^3/_{10} + 2 \times {}^1/_5 + 3 \times {}^2/_5$ $^{19}/_{10}$ oe	M1 A1 2	$\div 4$or6 $\Rightarrow$ M0A0
Total		**4**	
2i	$\Sigma x = 20; \Sigma y = 11; \Sigma x^2 = 96; \Sigma y^2 = 31; \Sigma xy = 52$) $S_{xx} = 16$ or 3.2 $S_{yy} = 6.8$ or 1.36 $S_{xy} = 8$ or 1.6 $r = \frac{8}{\sqrt{(16 \times 6.8)}}$ or $\frac{1.6}{\sqrt{(3.2 \times 1.36)}}$ $= 0.767$ (3 sfs)	B1 B1 B1 M1 A1 5	dep $-1 \leq r \leq 1$ ft their S's (S_{xx} & S_{yy} +ve) for M1 only
ii	Small sample oe	B1f 1	
Total		**6**	
3i	120	B1 1	not just 5!
iia	3 x 4! or 72 (÷ 5!) $^3/_5$ oe	M1 A1 2	oe, eg $^{72}/_{120}$
b	Starts 1 or 21 (both) $^1/_5 + {}^1/_5 \times {}^1/_4$ $= {}^1/_4$ oe	M1 M1 A1 3	12,13,14,15, (≥2 of these incl 21, or allow 1 extra) can be implied by wking or 5x 3! or 4! + 3! (÷5!) complement: full equiv steps for Ms
Total		**6**	
4ia	**W & Y** ***oe***	B1 1	
b	**X** ***oe***	B1 1	
ii	Geo probs always decrease or Geo has no upper limit to x or $x \neq 0$	B1 1	Geo not fixed no. of values diags have fixed no of trials not Geo has +ve skew
iii	**W** Bin probs cannot fall then rise or bimodal	B1 B1dep 2	indep allow Bin probs rise then fall
Total		**5**	
5i	$\frac{2685 - \frac{140 \times 106.8}{8}}{3500 - \frac{140^2}{8}}$ or $\frac{2685 - 8 \times 17.5 \times 13.35}{3500 - 8 \times 17.5^2}$ $= {}^{136}/_{175}$ or 0.777 (3 sfs) $y - {}^{106.8}/_8 = 0.777(x - {}^{140}/_8)$ $y = 0.78x - 0.25$ or better or $y = {}^{136}/_{175}x - {}^1/_4$	M1 A1 M1 A1 4	Correct sub in any correct formula for b (incl. $(x - \bar{x})$ etc) or $a = {}^{106.8}/_8 - 0.777 \times {}^{140}/_8$ ft b for M1 ≥ 2 sfs sufficient for coeffs
ii	0.78 x 12 – 0.25 = 9.1 (2 sfs)	M1 A1f 2	M1: ft their equn A1: dep const term in equn
iiia b	Reliable Unreliable because extrapolating oe	B1 B1 2	Just "reliable" for both: B1
Total		**8**	

6i	Geo($^2/_3$) stated $(^1/_3)^3 \times {}^2/_3$ $= {}^2/_{81}$ or 0.0247 (3 sfs)	M1 M1 A1 3	or implied by $(^1/_3)^n \times {}^2/_3$

ii	$(^1/_3)^3$ $1 - (^1/_3)^3$ $^{26}/_{27}$ or 0.963 (3 sfs)	M1 M1 A1 3	or $^2/_3+^1/_3x^2/_3+(^1/_3)^2x^2/_3$: M2 one term omitted or extra or wrong: M1 $1 - (^1/_3)^4$ or $1- (^2/_3+^1/_3x^2/_3+(^1/_3)^2x^2/_3$):M1
iii	1 / 2/3 = 3/2 oe	M1 A1 2	
Total		**8**	
7i	$^2/_9$ or $^7/_9$ oe seen $^3/_9$ or $^6/_9$ oe seen $^1/_8$ or $^7/_8$ oe seen Correct structure All correct	B1 B1 B1 B1 B1 5	 ie 8 correct branches only, ignore probs & values including probs and values, but headings not req'd
ii	$^3/_{10}$ x $^7/_9$ + $^7/_{10}$ x $^3/_9$ + $^7/_{10}$ x $^6/_9$ $^{14}/_{15}$ or 0.933 oe	M2 A1 3	or $^3/_{10}x^7/_9$ + $^7/_{10}$ or $1 - ^3/_{10}$ x $^2/_9$ M1: one correct prod or any prod + $^7/_{10}$ or $^3/_{10}$ x $^2/_9$
iii	$^3/_{10}$ x $^2/_9$ x $^7/_8$ + $^7/_{10}$ x $^6/_9$ $^{21}/_{40}$ or 0.525 oe	M2 A1 3	M1: one correct prod cao
	No ft from diag except: with replacement: (i) structure: B1 (ii) $^{91}/_{100}$: B2 (iii) 0.553: B2		
Total		**11**	
8i	Med = 2 LQ = 1 or UQ = 4 IQR = 3	B1 M1 A1 3	cao or if treat as cont data: read cf curve or interp at 25 & 75 cao
ii	Assume last value = 7 (or eg 7.5 or 8 or 8.5) Σxf attempted $\geq$ 5 terms 2.6 or 3 sf ans that rounds to 2.6 Σx^2f or $\Sigma(x-m)^2f$ $\geq$ 5 terms $\sqrt{(\Sigma x^2f/100 - m^2)}$ or $\sqrt{(\Sigma(x-m)^2f)}/100$ fully correct but ft m 1.6 or 1.7 or 3 sf ans that rounds to 1.6 or 1.7	B1 M1 A1 M1 M1 A1 6	stated, & not contradicted in wking eg 7-9 or 7,8, 9 Not just in wking allow "midpts" in Σxf or Σx^2f dep M3 penalize > 3 sfs only once
iii	Median less affected by extremes or outliers etc (NOT anomalies)	B1 1	or median is an integer or mean not int. or not affected by open-ended interval general comment acceptable
iv	Small change in var'n leads to lge change in IQR UQ for W only just 4, hence IQR exaggerated orig data shows variations are similar	 B1 1	 for Old Moat LQ only just 1 & UQ only just 3 oe specific comment essential
v	OM % (or y) decr (as x incr) oe Old Moat	B1 B1 2	ranks reversed in OM or not rev in W NIS
Total		**13**	

9i	${}^{11}C_5 \times (1/4)^6 \times (3/4)^5$ 0.0268 (3 sfs)	M1 A1 2	or $462 \times (1/4)^6 \times (3/4)^5$
ii	$q^{11} = 0.05$ or $(1-p)^{11} = 0.05$ $\sqrt[11]{0.05}$ $q = 0.762$ or 0.7616 . . . $p = 0.238$ (3 sfs)	M1 M1 A1 A1f 4	(any letter except p)$^{11} = 0.05$ oe oe or invlog($\frac{\log 0.05}{11}$) ft dep M2
iii	$11 \times p \times (1-p) = 1.76$ oe $11p - 11p^2 = 1.76$ or $p - p^2 = 0.16$ $11p^2 - 11p + 1.76 = 0$ or $p^2 - p + 0.16 = 0$ $(25p^2 - 25p + 4 = 0)$ $(5p-1)(5p-4) = 0$ or $p = \frac{11 - \sqrt{(11^2 - 4 \times 11 \times 1.76)}}{2 \times 11}$ $p = 0.2$ or 0.8	M1 A1 A1 M1 A1 5	not $11pq = 1.76$ any correct equn after mult out or equiv with = 0 or correct fact'n or subst'n for their quad equ'n eg $p = \frac{1 \pm \sqrt{(1-4 \times 0.16)}}{2}$
Total		**11**	

Total 72 marks