

OXFORD CAMBRIDGE AND RSA EXAMINATIONS

**Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education**

MATHEMATICS

4732

Probability & Statistics 1

Thursday **12 JANUARY 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 5 printed pages and 3 blank pages.

1 Jenny and John are each allowed two attempts to pass an examination.

(i) Jenny estimates that her chances of success are as follows.

- The probability that she will pass on her first attempt is $\frac{2}{3}$.
- If she fails on her first attempt, the probability that she will pass on her second attempt is $\frac{3}{4}$.

Calculate the probability that Jenny will pass. [3]

(ii) John estimates that his chances of success are as follows.

- The probability that he will pass on his first attempt is $\frac{2}{3}$.
- Overall, the probability that he will pass is $\frac{5}{6}$.

Calculate the probability that if John fails on his first attempt, he will pass on his second attempt. [3]

2 For each of 50 plants, the height, h cm, was measured and the value of $(h - 100)$ was recorded. The mean and standard deviation of $(h - 100)$ were found to be 24.5 and 4.8 respectively.

(i) Write down the mean and standard deviation of h . [2]

The mean and standard deviation of the heights of another 100 plants were found to be 123.0 cm and 5.1 cm respectively.

(ii) Describe briefly how the heights of the second group of plants compare with the first. [2]

(iii) Calculate the mean height of all 150 plants. [2]

3 In Mr Kendall's cupboard there are 3 tins of baked beans and 2 tins of pineapple. Unfortunately his daughter has removed all the labels for a school project and so the tins are identical in appearance. Mr Kendall wishes to use both tins of pineapple for a fruit salad. He opens tins at random until he has opened the two tins of pineapples.

Let X be the number of tins that Mr Kendall opens.

(i) Show that $P(X = 3) = \frac{1}{5}$. [4]

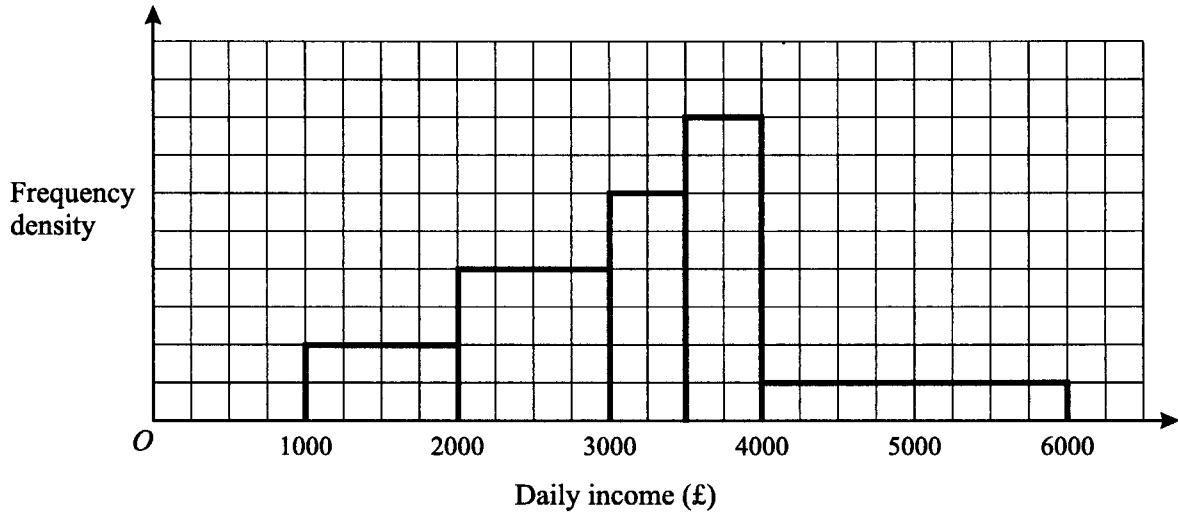
(ii) The probability distribution of X is given in the table below.

x	2	3	4	5
$P(X = x)$	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{3}{10}$	$\frac{2}{5}$

Find $E(X)$ and $\text{Var}(X)$. [5]

- 4 Each day, the Research Department of a retail firm records the firm's daily income, to be used for statistical analysis. The results are summarised by recording the number of days on which the daily income is within certain ranges.

(i)



The histogram shows the results for 300 days. By considering the total area of the histogram,

- (a) find the number of days on which the daily income was between £4000 and £6000, [4]
- (b) calculate an estimate of the number of days on which the daily income was between £2700 and £3200. [3]
- (ii) The Research Department offers to provide any of the following statistical diagrams: histogram, frequency polygon, box-and-whisker plot, cumulative frequency graph, stem-and-leaf diagram and pie chart.

Which one of these statistical diagrams would most easily enable managers to

- (a) read off the median and quartile values of the daily income, [1]
- (b) find the range of the top 10% of values of the daily income? [1]
- 5 Andrea practises shots at goal. For each shot the probability of her scoring a goal is $\frac{2}{5}$. Each shot is independent of other shots.
- (i) Find the probability that she scores her first goal
- (a) on her 5th shot, [2]
- (b) before her 5th shot. [3]
- (ii) (a) Find the probability that she scores exactly 1 goal in her first 5 shots. [3]
- (b) Hence find the probability that she scores her **second** goal on her 6th shot. [2]

- 6 An examination paper consists of two parts. Section A contains questions A1, A2, A3 and A4. Section B contains questions B1, B2, B3, B4, B5, B6 and B7.

Candidates must choose three questions from section A and four questions from section B. The order in which they choose the questions does not matter.

(i) In how many ways can the seven questions be chosen? [3]

(ii) Assuming that all selections are equally likely, find the probability that a particular candidate chooses question A1 but does **not** choose question B1. [3]

(iii) Following a change of syllabus, the form of the examination remains the same except that candidates who choose question A1 are not allowed to choose question B1. In how many ways can the seven questions now be chosen? [3]

- 7 Past experience has shown that when seeds of a certain type are planted, on average 90% will germinate. A gardener plants 10 of these seeds in a tray and waits to see how many will germinate.

(i) Name an appropriate distribution with which to model the number of seeds that germinate, giving the value(s) of any parameters. State any assumption(s) needed for the model to be valid. [4]

(ii) Use your model to find the probability that fewer than 8 seeds germinate. [2]

Later the gardener plants 20 trays of seeds, with 10 seeds in each tray.

(iii) Calculate the probability that there are at least 19 trays in each of which at least 8 seeds germinate. [4]

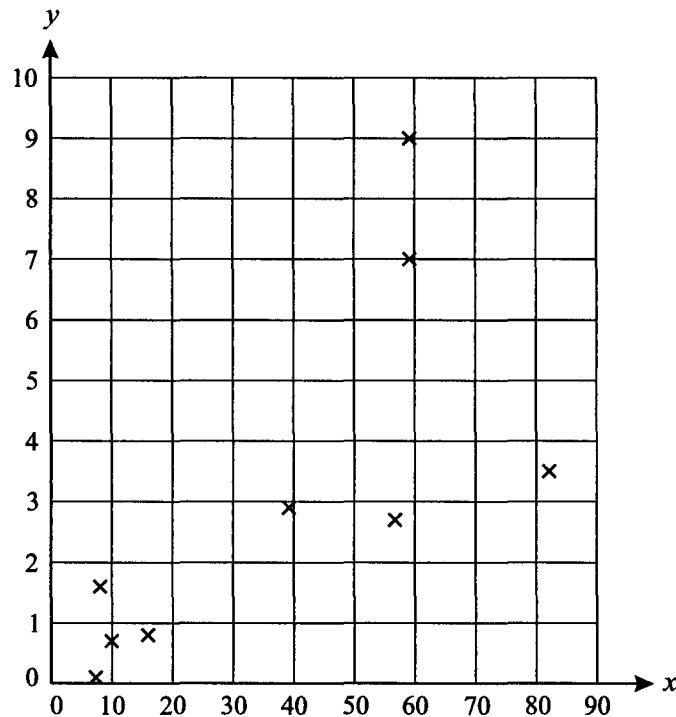
- 8 The table shows the population, x million, of each of nine countries in Western Europe together with the population, y million, of its capital city.

	Germany	United Kingdom	France	Italy	Spain	The Netherlands	Portugal	Austria	Switzerland
x	82.1	59.2	59.1	56.7	39.2	15.9	9.9	8.1	7.3
y	3.5	7.0	9.0	2.7	2.9	0.8	0.7	1.6	0.1

$$[n = 9, \Sigma x = 337.5, \Sigma x^2 = 18959.11, \Sigma y = 28.3, \Sigma y^2 = 161.65, \Sigma xy = 1533.76.]$$

- (i) (a) Calculate Spearman's rank correlation coefficient, r_s . [5]
 (b) Explain what your answer indicates about the populations of these countries and their capital cities. [1]
- (ii) Calculate the product moment correlation coefficient, r . [2]

The data are illustrated in the scatter diagram.



- (iii) By considering the diagram, state the effect on the value of the product moment correlation coefficient, r , if the data for France and the United Kingdom were removed from the calculation. [1]
- (iv) In a certain country in Africa, most people live in remote areas and hence the population of the country is unknown. However, the population of the capital city is known to be approximately 1 million. An official suggests that the population of this country could be estimated by using a regression line drawn on the above scatter diagram.
- (a) State, with a reason, whether the regression line of y on x or the regression line of x on y would need to be used. [2]
- (b) Comment on the reliability of such an estimate in this situation. [2]

1(i)	$\frac{2}{3} + \text{prod of 2 P's}$ or $1 - \text{prod of 2 P's}$ $\frac{2}{3} + \frac{1}{3} \times \frac{3}{4}$ or $1 - \frac{1}{3} \times \frac{1}{4}$ $= \frac{11}{12}$ or 0.917 (3 sfs)	M1 M1 A1	3	or $\frac{1}{3} \times \frac{3}{4}$ or $\frac{1}{3} \times \frac{1}{4}$
(ii)	$\frac{1}{3} \times p$ $\frac{2}{3} + \frac{1}{3} \times p = \frac{5}{6}$ oe $p = \frac{1}{2}$	M1 M1 A1	3	or $\frac{1}{3}(1-p)$ or $\frac{1}{3}(1-p) = 1 - \frac{5}{6}$ SW: $\frac{1}{2} \times \frac{1}{3} = \frac{1}{6}$ M2A0, unless clear this is a check
Total			6	
2(i)	124.5, 4.8	B1B1	2	for 4.8 allow "same"
(ii)	mean smaller or generally smaller or means similar or hts similar oe More widely spread or varied oe	B1f B1f	2	Assume 2 nd referred to unless clear 1 st or less consistent or gter dispersion or further from mean, gter variance Not "range" greater Allow opposite if ft (i)
(iii)	$(\text{"124.5"} + 2 \times 123)/3$ $= 123.5$	M1 A1	2	or $(50 \times \text{"124.5"} + 100 \times 123)/150$ cao
Total			6	
3(i)	$\frac{3}{5} \times \frac{2}{4} \times \frac{1}{3}$ or $\frac{2}{5} \times \frac{3}{4} \times \frac{1}{3}$ $\times 2$ or + $\frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} + \frac{2}{5} \times \frac{3}{4} \times \frac{1}{3}$ $= \frac{1}{5}$ AG	M1 M1 M1 A1	4	or $\frac{1}{10}$ <u>from tree</u> add 2 equal products of 3 probs all correct Must see correct working NB incorrect methods eg $\frac{3}{5} \times \frac{2}{4} \times \frac{2}{3}$
(ii)	Σxp $= 4$ $\Sigma x^2 p (= 17)$ $- \mu^2$ $= 1$	M1 A1 M1 M1 A1	5	≥ 3 terms added. Allow arith errors. ≥ 3 terms added. Allow arith errors Indep if +ve result $\Sigma(x-\mu)^2 p$ M2; 3 terms: M1 dep +ve result Σxp & $\Sigma x^2 p$, if ÷ eg 4: M0A0 (- μ^2 poss M1)
Total			9	

4(i)(a)	Total area = 60 sqs Recog that total area reps 300 $8 \times 300/60$ = 40	M1 M1 M1 A1	4	Attempt total area, eg 15000 or 15 cm ² eg 1 squ = 5 or 15000 ÷ (300 or 50) or 2000/50 cao
(b)	Splitting classes 1.2x4x5 or 0.8x6x5 oe 48	M1 M1 A1	3	or 0.3x16x5 or 0.4x12x5 or 24 NB other correct eg 2x4x5 + $\frac{4}{5}x2x5$ Alt method: estimate: 46-50 SC B1
(ii)(a)	Box & whisker	B1	1	
(b)	Cum freq diag	B1	1	
Total			9	
5(i)(a)	$(\frac{3}{5})^4 \times \frac{2}{5}$ = 0.0518 (3sfs) or $\frac{162}{3125}$ oe	M1 A1	2	Allow index 3 or 5
(b)	$(\frac{3}{5})^4$ $1 - (\frac{3}{5})^4$ = 0.870 (3 sfs) or $\frac{544}{625}$ oe	M1 M1 A1	3	$\frac{2}{5} + \frac{3}{5}x\frac{2}{5} + (\frac{3}{5})^2x\frac{2}{5} + (\frac{3}{5})^3x\frac{2}{5}$: M2 (1 extra or omit or wrong: M1) Allow $1 - (\frac{3}{5})^3$ or $1 - (\frac{3}{5})^5$
(ii)(a)	B($5, \frac{2}{5}$) stated $5 \times \frac{2}{5} \times (\frac{3}{5})^4$ or 0.3370 – 0.0778 = 0.259 (3 sfs) or $\frac{162}{625}$ oe	M1 M1 A1	3	or $({}^5C_a \text{ or } {}^5C_b)x(\frac{2}{5})^ax(\frac{3}{5})^b$ & $a + b = 5$
(b)	“0.259” $\times \frac{2}{5}$ = 0.104 (3 sfs) or $\frac{324}{3125}$ oe	M1 A1f	2	eg ft: (a) 0.0518 → 0.0207 (a) 0.922 → 0.369
Total			10	
6(i)	${}^4C_3 \times {}^7C_4$ = 140	M1M1 A1	3	M1 either comb. 140/330: M1M1
(ii)	${}^3C_2 \times {}^6C_4$ or $\frac{{}^3C_2}{{}^4C_3}$ or $\frac{{}^6C_4}{{}^7C_4}$ $\frac{{}^3C_2 \times {}^6C_4}{\text{“140”}}$ or $\frac{3}{4} \times (1 - \frac{4}{7})$ = $\frac{9}{28}$ oe or 0.321 (3 sfs)	M1 M1 A1	3	or ${}^3C_2(x...)/\text{“140”}$ or $(...x){}^6C_4/\text{“140”}$ or $({}^3C_2+{}^6C_4)/\text{“140”}$ or $(3+15)/\text{“140”}$ or $\frac{3}{4}$ or $1 - \frac{4}{7}$ seen all correct
(iii)	${}^3C_2 \times {}^6C_4$ (or i x ii) or $({}^3C_3x){}^7C_4$ or 45 or 35 or $\frac{1}{4}x{}^4C_3x{}^7C_4$ or $\frac{3}{4}x{}^4C_3x{}^6C_4$ ${}^3C_2x{}^6C_4 + ({}^3C_3x){}^7C_4$ or “140” – ${}^3C_2x{}^6C_3$ = 80	M1 M1 A1ft	3	1 correct prod or “140” – any prod or $\frac{1}{4}x{}^4C_3x{}^7C_4 + \frac{3}{4}x{}^4C_3x{}^6C_4$ ft only “140”
Total			9	

7(i)	Binomial $n = 10, p = 0.9$ Each seed equally likely germ or P(germ) same for all seeds oe Seeds independent oe	B1 B1 B1 B1	4	Both requ'd. Ignore $q = 0.1$ or seeds grown in same conditions Context nec'y for each B1
(ii)	0.0702 (3 sfs)	B2	2	0.07 or 0.2639: B1 Σ or $1-\Sigma$: 1 term extra or omit or wrong: M1
(iii)	1 - "0.0702" $0.9298^{20} + {}^{20}C_1 \times 0.0702 \times 0.9298^{19}$ = 0.585 (3 sfs)	M1 M1M1 A1	4	Or 0.9298 or 0.93(0) seen M1 each term cao eg ft (ii) 0.2639 \rightarrow (iii) 0.0178 from correct wking: M3A0 $0.0702^{20} + {}^{20}C_1 \times 0.9298 \times 0.0702^{19}$ (= 2.25×10^{-21}): SC M1M1 NB ft (ii) for all M mks. But if 0.1, 0.9 used, must be clear using (ii) rounded
Total			10	

8(i)(a)	Ranks 1 2 3 4 5 6 7 8 9 9 8 7 6 5 4 3 2 1 3 2 1 5 4 7 8 6 9 7 8 9 5 6 3 2 4 1 $\Sigma d^2 (= 16)$ $r_s = 1 - \frac{6 \times \text{their } 16}{9 \times (9^2 - 1)}$ $= 0.867$ (3 sfs) or $\frac{13}{15}$ oe	M1 A1 M1dep M1dep A1 5	Attempt ranks, same dir'n Correct ranks Dep ranks attempted Correct formula with $n = 9$, dep M1M1
(b)	Countries with larger pops tend to have larger capital pops. oe	B1ft 1	or ft (a) Must <u>interp</u> & refer to context. Not "Gd corr'n country & cap pops" Not "Gd agree't country & cap pops" Not "Gd rel'nship country & cap pops" Not "proportional"
(ii)	$\frac{1533.76 - (337.5 \times 28.3)/9}{\sqrt{((18959.11 - 337.5^2/9)(161.65 - 28.3^2/9))}}$ $= 0.698$ (3 sfs)	M1 A1 2	(= $472.51/\sqrt{(6302.86 \times 72.66)}$) Or correct subst in 2 "S" formulae, any version No wking: 0.7 M0A0; 0.70: M1A0
(iii)	Increase	B1 1	or nearer to 1
(iv)(a)	x on y Est country pop from cap or x from y oe	B1ind B1ind 2	y indep or known or given or x unknown or x dep on y oe
(b)	any indication different context, eg "Africa", "remote areas" unreliable	B1 B1dep 2	or reliable because r (or r_s) high: B1 or unreliable because r (or r_s) not hi: B1 "accurate": B0
Total		13	

Total 72 marks