OCR
RECOGNISING ACHIEVEMENT

OXFORD CAMBRIDGE AND RSA EXAMINATIONS

Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education

MATHEMATICS 4734

Probability & Statistics 3

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 4 printed pages.

 [Turn over

1 In order to judge the support for a new method of collecting household waste, a city council arrange‹ a survey of 400 householders selected at random. The results showed that 186 householders were i‹ favour of the new method.

(i) Calculate a 95% confidence interval for the proportion of all householders who are in favour o the new method. [5

A city councillor said he believed that as many householders were in favour of the new method a were against it.

(ii) Comment on the councillor's belief. [1

2 A particular type of engine used in rockets is designed to have a mean lifetime of at least 2000 hours A check of four randomly chosen engines yielded the following lifetimes in hours.

1896.4 2131.5 1903.3 1901.6

A significance test of whether engines meet the design is carried out. It may be assumed that lifetime have a normal distribution.

(i) Give a reason why a t-test should be used. [1

(ii) Carry out the test at the 10% significance level. [8

3 For a restaurant with a home-delivery service, the delivery time in minutes can be modelled by ‹ continuous random variable T with probability density function given by

$$f(t) = \begin{cases} \frac{\pi}{90}\sin\left(\frac{\pi t}{60}\right) & 20 \leqslant t \leqslant 60, \\ 0 & \text{otherwise.} \end{cases}$$

(i) Given that $20 \leqslant a \leqslant 60$, show that $P(T \leqslant a) = \frac{1}{3}\left(1 - 2\cos\left(\frac{\pi a}{60}\right)\right)$. [3

There is a delivery charge of £3 but this is reduced to £2 if the delivery time exceeds a minutes.

(ii) Find the value of a for which the expected value of the delivery charge for a home-delivery i £2.80. [4

4 A multi-storey car park has two entrances and one exit. During a morning period the numbers of car using the two entrances are independent Poisson variables with means 2.3 and 3.2 per minute. Th‹ number leaving is an independent Poisson variable with mean 1.8 per minute. For a randomly chose‹ 10-minute period the total number of cars that enter and the number of cars that leave are denoted b‹ the random variables X and Y respectively.

(i) Use a suitable approximation to calculate $P(X \geqslant 40)$. [6

(ii) Calculate $E(X - Y)$ and $Var(X - Y)$. [3

(iii) State, giving a reason, whether $X - Y$ has a Poisson distribution. [2

5 The continuous random variable X has cumulative distribution function given by

$$\mathrm{F}(x) = \begin{cases} 0 & x < 1, \\ \frac{1}{8}(x-1)^2 & 1 \leqslant x < 3, \\ a(x-2) & 3 \leqslant x < 4, \\ 1 & x \geqslant 4, \end{cases}$$

where a is a positive constant.

(i) Find the value of a. [2

(ii) Verify that $C_X(8)$, the 8th percentile of X, is 1.8. [2

(iii) Find the cumulative distribution function of Y, where $Y = \sqrt{X-1}$. [5

(iv) Find $C_Y(8)$ and verify that $C_Y(8) = \sqrt{C_X(8) - 1}$. [3

6 A company with a large fleet of cars compared two types of tyres, A and B. They measured th stopping distances of cars when travelling at a fixed speed on a dry road. They selected 20 cars a random from the fleet and divided them randomly into two groups of 10, one group being fitted wit tyres of type A and the other group with tyres of type B. One of the cars fitted with tyres of type broke down so these tyres were tested on only 9 cars. The stopping distances, x metres, for the tw samples are summarised by

$$n_A = 9, \quad \bar{x}_A = 17.30, \quad s_A^2 = 0.7400, \quad n_B = 10, \quad \bar{x}_B = 14.74, \quad s_B^2 = 0.8160,$$

where s_A^2 and s_B^2 are unbiased estimates of the two population variances.

It is given that the two populations have the same variance.

(i) Show that an unbiased estimate of this variance is 0.780, correct to 3 decimal places. [2

The population mean stopping distances for cars with tyres of types A and B are denoted by μ_A metre and μ_B metres respectively.

(ii) Stating any further assumption you need to make, calculate a 98% confidence interval for $\mu_A - \mu_B$ [5

The manufacturers of Type B tyres assert that $\mu_B < \mu_A - 2$.

(iii) Carry out a significance test of this assertion at the 5% significance level. [6

[Question 7 is printed overleaf.]

7 The continuous random variable X has probability density function given by

$$f(x) = \begin{cases} \alpha x^{-\alpha-1} & x \geqslant 1, \\ 0 & \text{otherwise,} \end{cases}$$

where α is a constant and $\alpha > 1$. This is called a Pareto distribution.

(i) Show that $\mathrm{E}(X) = \dfrac{\alpha}{\alpha - 1}$. [3]

Zipf's law states that the distribution of the population size of certain settlements should follow a Pareto distribution. The following table summarises the population distribution of a random sample of 200 settlements.

Population size in thousands (x)	$1 \leqslant x < 2$	$2 \leqslant x < 3$	$3 \leqslant x < 4$	$4 \leqslant x < 5$	$5 \leqslant x < 6$	$x \geqslant 6$
Frequency	146	33	14	5	2	0

(ii) Assuming that x has the above Pareto distribution, and given that the sample mean is 1.920, show that an estimate of α is 2.087, to 3 decimal places. [1]

For $\alpha = 2.087$, the following table gives the expected frequencies, each correct to 1 decimal place.

Population size in thousands (x)	$1 \leqslant x < 2$	$2 \leqslant x < 3$	$3 \leqslant x < 4$	$4 \leqslant x < 5$	$5 \leqslant x < 6$	$x \geqslant 6$
Expected frequency	152.9	26.9	9.1	4.1	2.2	4.8

(iii) Show how the value 26.9 for the interval $2 \leqslant x < 3$ is obtained. [4]

(iv) Carry out a test, at the 5% significance level, of whether the data supports Zipf's law. [6]

STATISTICS 3

1	(i)	$p_s \pm z\sigma_{est}$	M1		Use formula,σ involving p_s and 400.
		p_s=186/400(0.465)	A1		
		$\sigma_{est}=\sqrt{\dfrac{0.465\times0.535}{400}}$	B1		
		z=1.96	A1		
		(0.416,0.514)	A1	**5**	
	(ii)	Councillor statement implies p=0.5. CI does contain 0.5 but only juat so councillor probably correct.	B1	**1**	Any justifiable comment **Not too assertive**
2	(i)	σ^2 unknown	B1	**1**	
	(ii)	H_0: μ=2000 (or ≥), H_1:μ <2000	B1		
		$\bar{x}$ =1958.2, s=115.57	B1B1		or 1958,115.6
	EITHER:	Test statistic $=\dfrac{1958.2-2000}{115.57/2}$	M1		
		=-0.7234	A1		art –0.723
		Critical value –1.638	B1		
		Test statistic not in CR, accept H_0		M1	Or equivalent
		Accept that specification is being met	A1		Conclusion in context
	OR:	Critical region:			
		$\dfrac{\bar{x}-2000}{115.57/2}<t$	M1		
		t=-1.638		B1	
		$\bar{x}$ < 1905.2	A1		art 1900 or 1910
		As above	M1A1	**8**	Conclusion in context
3	(i)	Use of $\int_{20}^{a}$ f(t)dt	M1		With limits and f(t) substituted
		$\left[-\dfrac{2}{3}\cos\dfrac{\pi t}{60}\right]_{20}^{a}$	A1		
		AG	A1	**3**	Properly obtained
	(ii)	3×(i) + 2×(1-(i))	M1		Idea of expectation
			A1		All correct
		Equate to 2.80 and attempt to solve		M1	From equation in a, 2 or 3
		a=44.8	A1	**4**	Accept 45
					SR: ⅓(1-2cos..)= 0.8 give max3/4

4	(i)	Use Poisson distribution	M1		Po(5.5) or Po(55) seen
		With μ=55	B1		
		σ^2=55	A1		
		$(39.5-55)/\sqrt{55}$	A1		Standardising, with ,without or wrong cc
		-2.09	A1		
		art 0.982		A1	**6**
	(ii)	E(X- Y)=37	B1√		ft μ above
		Var(X-Y)=55+18		M1	
		=73	A1√	**3**	ft μ above
	(iii)	EITHER: Expectation not equal to variance OR: X-Y could be negative OR: Difference of two Poisson variables could have a negative expectation		M1	Any one
		So X-Y does not have a Poisson distn	A1	**2**	
5	(i)	EITHER: Use $\frac{1}{8}(3-1)^2=a(3-2)$ OR: a(4-2)=1		M1	Continuity of F
		a=½	A1	**2**	
	(ii)	$F(1.8)=\frac{1}{8}(0.8)^2$ =0.08	M1		Appropriate use of F
		$C_X(8)=1.8$	A1	**2**	
	(iii)	$G(y)=P(Y\le y)=P((X-1)^{1/2}\le y)$	M1		
		$=P(X\le y^2+1)$		A1	
		$=F(y^2+1)$	A1		
		$G(y)=\begin{cases}\frac{1}{8}y^4 & (0\le y\le\sqrt{2}),\\ \frac{1}{2}(y^2-1) & (\sqrt{2}<y\le\sqrt{3}).\end{cases}$	A1		
		Ignore others, A1 for both ranges of y	B1	**5**	
	(iv)	Use G(y) to find $C_Y(8)$	M1		
		Obtain √(0.8)	A1		
		Correct verification	B1	**3**	

6	(i)	$s^2=(8\times0.7400+9\times0.8160)/17$	M1		Formula for pooled estimate
		$=0.7802$	A1		At least 4DP shown
		0.780 AG		2	
	(ii)	Assumes braking distances have normal distributions	B1		
		Use $\overline{x}_A-\overline{y}_B\pm t\sigma$	M1		Must be t value
		$t=2.567$	A1		
		$\sigma=\sqrt{[0.7802(1/10+1/9)}$ (0.40584)	B1		Allow 0.780
		(1.518,3.602)	A1	5	art (1.52, 3.60)
	(iii)	$H_0{:}\mu_A-\mu_B=2$, $H_1{:}\mu_A-\mu_B>2$	B1		For both hypotheses
		Use of CV, 1.740	B1		
	EITHER:	Test statistic=$(2.56-2)/\sigma$	M1		Standardising, σ as above
		$=1.38$	A1		Rounding to 1.38
	OR:	Critical region			
		$\overline{x}_A-\overline{x}_B>2+1.74\times0.4054$	M1		
		$=2.7054$	A1		2.70 or 2.71
		Indication that test statistic is not in critical region	M1		Not from different signs test statistic
		and Insufficient evidence to accept claim and H_1	A1	6	critical value. A1 dep on correct H_0
7	(i)	Use $\int_1^\infty x\alpha x^{-\alpha-1}dx=(\int_1^\infty \alpha x^{-\alpha}dx)$	M1		
		$\left[\dfrac{-\alpha x^{-\alpha+1}}{\alpha-1}\right]_1^\infty$	B1		Correct limits not required
		$=\alpha/(1-\alpha)$ AG	A1	3	Properly obtained
	(ii)	$\alpha/(1-\alpha)=1.92$ giving 2.087 AG	B1	1	
	(iii)	Integral of $2.087x^{-3.087}$ from 2 to 3	M1		
		$[-x^{-2.807}]_2^3$	A1		
		×200	A1		
		Obtain AG	A1	4	Evidence required
	(iv)	Combine last 3 cells	B1		
		$X^2=6.9^2/152.9+6.1^2/26.9$ $+4.9^2/9.1+4.1^2/11.18$	M1		Accept one error
			A1		All correct
		$=5.847\ldots$	A1		art 5.8
		Use CV 5.991	B1 √		ft number of sells used.
		Accept that data supports Zipf's law	B1	6	
		SR: From 6 cells: B0M1A1 (for 9.34) then B1 for 9.488, B1 Max 4/6			