OCR RECOGNISING ACHIEVEMENT

OXFORD CAMBRIDGE AND RSA EXAMINATIONS

Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education

MATHEMATICS 2644

Probability & Statistics 4

Monday **19 JANUARY 2004** Morning 1 hour 20 minutes

Additional materials:
Answer booklet
Graph paper
List of Formulae (MF8)

TIME 1 hour 20 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 graphic 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 60.
- 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.

1 The events A and B are such that $P(A) = 0.6$, $P(B) = 0.7$ and $P(A \cap B) = 0.4$. Find

(i) $P(A|B)$, [1]

(ii) $P(A' \cap B')$. [3]

2 The discrete random variables X and Y have joint probability distribution given in the following table.

		X		
		0	1	2
Y	−1	0.3	0	0.45
	1	0	0.2	0.05

(i) Show that $\text{Cov}(X, Y) = 0$. [4]

(ii) Determine whether X and Y are independent, giving your reasoning. [3]

3 Each person in a random sample of 10 married couples was asked the same 50 questions relating to health and social issues. A correct answer scored 1 and an incorrect answer scored zero. The results are given below.

Couple	A	B	C	D	E	F	G	H	I	J
Wife's score	35	40	29	21	35	43	17	49	42	12
Husband's score	48	28	38	26	32	44	24	43	38	20

(i) Carry out the Wilcoxon matched-pairs signed-rank test, at the 10% significance level, to test whether the median scores for husbands and wives differ. [6]

(ii) State why, in this case, the Wilcoxon test is likely to be more appropriate than

(a) the sign test, [1]

(b) the t-test. [1]

(Assume that the assumptions needed to carry out the Wilcoxon matched-pairs signed-rank test can be made in this case.)

4 The continuous random variables A and B have identical distributions except that the averages may differ. Random samples of 5 observations of A and 7 observations of B are taken. The 12 observations are ranked, smallest first, and R_5 denotes the sum of the ranks of the observations of A.

(i) Show that the number of different possible rankings of the 5 ordered observations of A is 792. [1]

(ii) There are 7 rankings for which $R_5 \leqslant 18$. List these rankings, together with the corresponding values of R_5. [3]

Five randomly chosen mathematics students from Ayton University and seven from Beeton University were each given the same problem to solve. The times that they took (in minutes, correct to the nearest minute) are given below.

Ayton Students	9	12	13	15	22		
Beeton Students	16	18	19	23	26	27	30

Using a suitable non-parametric test, it was decided that the median time taken for students at Ayton University is less than that of students at Beeton University.

(iii) What can be said about the significance level of the test? [4]

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

$$f(x) = \begin{cases} \frac{1}{6}x^3e^{-x} & x > 0, \\ 0 & \text{otherwise.} \end{cases}$$

(i) Explain how you can deduce that the value of $\int_0^\infty x^3e^{-x}\,dx$ is 6. [1]

(ii) Show that the moment generating function of X is given by

$$M_X(t) = \frac{1}{6}\int_0^\infty x^3e^{-x(1-t)}\,dx. \quad [2]$$

(iii) Hence, by using the substitution $u = x(1-t)$, show that $M_X(t) = (1-t)^{-4}$. [3]

(iv) Using the moment generating function, find $E(X)$ and $Var(X)$. [4]

[Questions 6 and 7 are printed overleaf.]

6 The random variable X has a Poisson distribution with mean λ, given by

$$P(X = r) = e^{-\lambda}\frac{\lambda^r}{r!} \qquad \text{for } r = 0, 1, 2, \ldots .$$

(i) Show that the probability generating function of X is $e^{-\lambda(1-t)}$. [3]

The independent random variable Y has a Poisson distribution with mean μ, and T denotes the sum of one observation of X and one observation of Y.

(ii) Show that T has a Poisson distribution, and state the mean of T. [3]

Over any period of 24 hours the number of accidents X occurring on a certain road during daylight has a Po(1.2) distribution. During the night the number of accidents Y occurring on the same road has a Po(0.9) distribution. It may be assumed that accidents occur independently. Over a particular 24-hour period it was reported that a total of 2 accidents had occurred.

(iii) Find the conditional probability that

(a) no accident,

(b) one accident,

(c) two accidents

occurred during the daylight hours of that period. [4]

(iv) Show that the probabilities in part (iii) form a binomial distribution with $n = 2$, and state the value of p. [2]

7 The masses, in grams, of small bags of sugar have mean μ and variance 1.2. The masses, in grams, of large bags have mean 2μ and variance 1.6. The mass of a randomly chosen small bag is denoted by S and the mass of a randomly chosen large bag is denoted by L. The estimator T is defined by $T = aS + bL$, where a and b are constants. It is given that T is an unbiased estimator of μ.

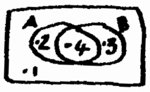
(i) Show that $a + 2b = 1$. [2]

(ii) Find the values of a and b for which T has the smallest possible variance. [6]

The mean mass of a random sample of two small bags is denoted by $\bar{S}$ and the mean mass of a random sample of two large bags is denoted by $\bar{L}$. The estimator U, where $U = a\bar{S} + b\bar{L}$, has the values of a and b found in part (ii).

(iii) Show that U is an unbiased estimator of μ, and find $\dfrac{\text{Var}(U)}{\text{Var}(T)}$. [3]

S4 Jan 2004

(1) 

$P(A|B) = 4/7$ (1)

$P(A' \cap B') = 1/10$. (3)

(2) $Cov(X,Y) = E(XY) - \mu_x \mu_y$

$= (1 \times 0.2 + -2 \times 0.45 + 2 \times 0.05)$

$- (1 \times 0.2 + 2 \times 0.5)(-1 \times 0.75 + 1 \times 0.25)$

$= -0.6 - 1.2 \times -0.5 = \underline{0}$. (4)

Not independent. eg.

$P(X=0 \cap Y=-1) = 0.3 \neq 0.3 \times 0.75$. (3)

(3) H_0: median of diffs $= 0$

H_1: " " " $\neq 0$

d_i: 13 −12 9 5 −3 1 7 −6 −4 8

signed ranks: +10 −9 +8 +4 −2 +1 +6 −5 −3 +7

$P = 36$, $Q = 19$ so $\underline{T = 19}$

critical value $= 10$, this is well outside the tail so conclude no significant difference between m/bw scores. (6)

Sign test discards information about size of score differences, whereas this weights them appropriately. (1)

Don't know distribution of d's is Normal (1)

(4) 5 can occupy 5 out of 12 rankings; neglecting different orders, $^{12}C_5$ is no of ways, $= \underline{792}$. (1)

	R_s
1 2 3 4 5	15
1 2 3 4 6	16
1 2 3 4 7	17
1 2 3 4 8	18
1 2 3 5 6	17
1 2 3 5 7	18
1 2 4 5 6	18

(3)

ranks: Ayton 1 2 3 4 8

Barton 5 6 7 9 10 11 12

So $W_{(R_m)} = 18$.

Given this was found to be in the tail, significance level must be $\geq \frac{7}{792}$ i.e. 0.88% so conventionally, $\geq 1\%$. (4)

(5) Total prob $= 1 = \int_0^\infty \frac{1}{6} x^3 e^{-x}$

So $6 = \int_0^\infty x^3 e^{-x} dx$. (1)

$M_X(t) = E(e^{tX}) = \int \frac{1}{6} x^3 e^{-x} \cdot e^{tx} dx$

$= \int_0^\infty \frac{1}{6} x^3 e^{-x(1-t)} dx$. (2)

$u = x(1-t)$

$\rightarrow \frac{du}{dx} = 1-t$

$M = \int_0^\infty \frac{1}{6} \frac{u^3}{(1-t)^3} \cdot e^{-u} \cdot \frac{du}{1-t}$

(for $|t|<1$)

$= \frac{1}{(1-t)^4} \int_0^\infty \frac{1}{6} u^3 e^{-u} du$

↳ which we know is 1

So $M_X(t) = (1-t)^{-4}$. (3)

$M'_X(t) = 4(1-t)^{-5}$, $M''_X(t) = 20(1-t)^{-6}$

So $\underline{E(X) = 4}$, $\underline{V(X) = 4}$ (4)

(6) $G_X(t) = E(t^X) = e^{-\lambda} + \lambda e^{-\lambda} t + \frac{\lambda^2 e^{-\lambda}}{2!} t^2 + \ldots$

$= e^{-\lambda}\left(1 + \lambda t + \frac{(\lambda t)^2}{2!} + \ldots\right) = e^{-\lambda} e^{\lambda t}$

$= \underline{e^{-\lambda(1-t)}}$. (3)

$G_{X+Y}(t) = G_X \times G_Y = e^{-\lambda(1-t)} \times e^{-\mu(1-t)}$

$= e^{-(\lambda+\mu)(1-t)}$ which is $P(\lambda+\mu)$. (3)

$P(X+Y) = 2 = \frac{2.1^2 e^{-2.1}}{2!} = 0.270$.

(i) $P(X=0 | X+Y=2) = \frac{P(X=0 \cap Y=2)}{P(X+Y=2)}$

$= \frac{e^{-1.2} \times \frac{0.9^2 e^{-0.9}}{2!}}{0.270\ldots} = \frac{0.9^2}{2.1^2} = \underline{0.184}$

similarly, (ii) $P(X=1 | X+Y=2) = \underline{0.490}$

(iii) $= \underline{0.327}$ (4)

Probs were $\frac{0.9^2}{2.1^2}$, $\frac{2 \times 1.2 \times 0.9}{2.1^2}$, $\frac{1.2^2}{2.1^2}$

which are q^2, $2pq$, p^2

So $\sim B(2, \frac{4}{7})$. (2)

↖ p

(7) $E(T) = E(aS+bL)$

$= aE(S) + bE(L)$

$= a\mu + 2b\mu = \mu$ (unbiased)

$\therefore\ a + 2b = 1.$ (2)

$V(T) = V(aS + bL)$

$= a^2 \times 1.2 + b^2 \times 1.6$

$= 1.2(1-2b)^2 + 1.6b^2.$

$\frac{dV}{db} = 1.2 \times 2(1-2b) \times -2 + 3.2b$

$= 0$ for ~~max~~/min. ← happy parabola!

$\therefore\ b = 3/8,\ a = 1/4.$ (6)

$E(U) = E(a\bar{S} + b\bar{L})$

$= \frac{1}{4}\cdot\mu + \frac{3}{8}\cdot 2\mu = \mu$ so unbiased ✓

$\frac{V(U)}{V(T)} = \frac{a^2 \times \frac{1.2}{2} + b^2 \times \frac{1.6}{2}}{a^2 \times 1.2 + b^2 \times 1.6} = \underline{\frac{1}{2}}$ (3)