RECOGNISING ACHIEVEMENT

ADVANCED GCE
MATHEMATICS 4735
Probability & Statistics 4

Candidates answer on the Answer Booklet

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

Other Materials Required:
- Scientific or graphical calculator

Thursday 24 June 2010
Morning

Duration: 1 hour 30 minutes

4735

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 **4** pages. Any blank pages are indicated.

1 For the variables A and B, it is given that $\text{Var}(A) = 9$, $\text{Var}(B) = 6$ and $\text{Var}(2A - 3B) = 18$.

(i) Find $\text{Cov}(A, B)$. **[3]**

(ii) State with a reason whether A and B are independent. **[1]**

2 The probability generating function of the discrete random variable X is $\dfrac{e^{4t^2}}{e^4}$. Find

(i) $E(X)$, **[3]**

(ii) $P(X = 2)$. **[3]**

3 X_1 and X_2 are continuous random variables. Random samples of 5 observations of X_1 and 6 observations of X_2 are taken. No two observations are equal. The 11 observations are ranked, lowest first, and the sum of the ranks of the observations of X_1 is denoted by R.

(i) Assuming that all rankings are equally likely, show that $P(R \leqslant 17) = \frac{2}{231}$. **[5]**

The marks of 5 randomly chosen students from School A and 6 randomly chosen students from School B, who took the same examination, achieving different marks, were ranked. The rankings are shown in the table.

Rank	1	2	3	4	5	6	7	8	9	10	11
School	A	A	A	B	A	A	B	B	B	B	B

(ii) For a Wilcoxon rank-sum test, obtain the exact smallest significance level for which there is evidence of a difference in performance at the two schools. **[2]**

4 The moment generating function of a continuous random variable Y, which has a χ^2 distribution with n degrees of freedom, is $(1 - 2t)^{-\frac{1}{2}n}$, where $0 \leqslant t < \frac{1}{2}$.

(i) Find $E(Y)$ and $\text{Var}(Y)$. **[5]**

For the case $n = 1$, the sum of 60 independent observations of Y is denoted by S.

(ii) Write down the moment generating function of S and hence identify the distribution of S. **[2]**

(iii) Use a normal approximation to estimate $P(S \geqslant 70)$. **[3]**

5 In order to test whether the median salary of employees in a certain industry who had worked for three years was £19 500, the salaries x, in thousands of pounds, of 50 randomly chosen employees were obtained.

(i) The values $|x - 19.5|$ were calculated and ranked. No two values of x were identical and none was equal to 19.5. The sum of the ranks corresponding to positive values of $(x - 19.5)$ was 867. Stating a required assumption, carry out a suitable test at the 5% significance level. **[10]**

(ii) If the assumption you stated in part **(i)** does not hold, what test could have been used? **[1]**

6 Nuts and raisins occur in randomly chosen squares of a particular brand of chocolate. The numbers of nuts and raisins are denoted by N and R respectively and the joint probability distribution of N and R is given by

$$\mathrm{f}(n, r) = \begin{cases} c(n+2r) & n = 0, 1, 2 \text{ and } r = 0, 1, 2, \\ 0 & \text{otherwise,} \end{cases}$$

where c is a constant.

(i) Find the value of c. **[3]**

(ii) Find the probability that there is exactly one nut in a randomly chosen square. **[2]**

(iii) Find the probability that the total number of nuts and raisins in a randomly chosen square is more than 2. **[2]**

(iv) For squares in which there are 2 raisins, find the mean number of nuts. **[4]**

(v) Determine whether N and R are independent. **[2]**

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

$$\mathrm{f}(x) = \begin{cases} \dfrac{x}{2\theta^2} & 0 \leqslant x \leqslant 2\theta, \\ 0 & \text{otherwise,} \end{cases}$$

where θ is an unknown positive constant.

(i) Find $\mathrm{E}(X^n)$, where $n \neq -2$, and hence write down the value of $\mathrm{E}(X)$. **[3]**

(ii) Find

(a) $\mathrm{Var}(X)$,

(b) $\mathrm{Var}(X^2)$.

[4]

(iii) Find $\mathrm{E}(X_1 + X_2 + X_3)$ and $\mathrm{E}(X_1^2 + X_2^2 + X_3^2)$, where X_1, X_2 and X_3 are independent observations of X. Hence construct unbiased estimators, T_1 and T_2, of θ and $\mathrm{Var}(X)$ respectively, which are based on X_1, X_2 and X_3. **[6]**

(iv) Find $\mathrm{Var}(T_2)$. **[2]**

8 For the events L and M, $\mathrm{P}(L \mid M) = 0.2$, $\mathrm{P}(M \mid L) = 0.4$ and $\mathrm{P}(M) = 0.6$.

(i) Find $\mathrm{P}(L)$ and $\mathrm{P}(L' \cup M')$. **[3]**

(ii) Given that, for the event N, $\mathrm{P}\big(N \mid (L \cap M)\big) = 0.3$, find $\mathrm{P}(L' \cup M' \cup N')$. **[3]**

1(i)	Var(2A – 3B) = 4Var(A) + 9Var(B) – 12Cov(A,B) $\Rightarrow$ 18 = 36 + 54 – 12Cov(A,B) $\Rightarrow$ Cov(A ,B) = 6	M1 A1 A1 **3**	Correct formula**.** Allow one error Substitute relevant values CAO
(ii)	Since Cov(A, B) ≠ 0, A and B are not independent	B1 ft **1** **(4)**	Must have a reason. ft Cov≠ 0
2(i)	$G'(t) = 8te^{4t^2}/e^4$ E(X) = G′(1) = 8	M1A1 A1 **3**	M1 for ct^2/e^4
(ii)	EITHER: $G(t) = e^{-4}(1 + 4t^2+\ldots)$ P(X=2) = coefficient of t^2 = $4e^{-4}$ or $4/e^4$ or 0.0733 OR $G''(t) = (8+64t^2)e^{4t^2-4}$ P(X=2) = $\frac{1}{2}G''(0) = 4e^{-4}$ or $4/e^4$ or 0.0733	M1A1 A1 **3** M1A1 A1 **(6)**	Expand in powers of t M1 for reasonable attempt at M"(t)
3(i)	Number of different rankings $^{11}C_5$ =462 For $R \le 17$: 1+2+3+4+5 = 15 1+2+3+4+6=16 1+2+3+5+6=17 1+2+3+4+7=17 P($R \le 17$) = 4/462 = 2/231 AG	M1 A1 B2 A1 **5**	Number of selections of 5 from 11 B1 for 2 or 3 correct
(ii)	W = 17 P($W \le 17$) = $\frac{2}{231}$ Smallest SL = $\frac{400}{231}$ %	M1 A1ft **2** **(7)**	 Allow $\frac{4}{231}$; ft $\frac{2}{231}$, but must be exact
4(i)	EITHER**:** (α) $M'(t) = n(1-2t)^{-½n-1}$ E(Y) = M'(0)=n $M''(t) = n(n+2)(1-2t)^{-½n-2}$ Var(Y) = $n(n+2) - n^2 = 2n$ OR: $M(t) = 1 + nt + \frac{1}{2}n(n+2)t^2$ E(Y) = n Var(Y) = $n(n+2) - n^2 = 2n$	M1 A1 A1 M1 A1 **5** M1A1A1 A1 A1 **5**	Correct form for M1 Ft similar M′(t) $M''(0) - (M'(0))^2$
(ii)	MGF = $(1-2t)^{-30}$ χ^2 distribution with 60 d.f.	B1 B1 **2**	From $[(1-2t)^{-1/2}]^{60}$
(iii)	E(S) = 60, Var(S) = 120 Using CLT, Probability =1 – Φ(10/√120) = 0.181	B1ft M1 A1 **3** **(10)**	From **(i)** Correct tail: allow cc

5(i)	Assumes salaries symmetrically distributed H_0: *m*(edian) = 19.5, H_1: *m*(edian)≠ 19.5 *P* = 867 (or 408) Using normal approximation μ = ¼ ×50×51 (= 637.5) σ^2 = 50×51×101/24 (= 10731.25) z = (*a* – 637.5)/√10731.25 Use *a* = 866.5 = 2.211, or 2.215 or 2.220 (– from 408) Compare their *z* with 1.96 and reject H_0 There is sufficient evidence at the 5% SL that the median salary differs from £19 500	B1 B1 M1 A1 A1 M1 A1 A1 M1 A1 ft **10**	In context For both ; not μ ; accept words *a*=866.5, 867, 867.5 (or 408.5, 408, 407.5) Or *p*-value rounding to 0.026 or 0.027 Compare with 0.05 or equivalent ft *z* Or find critical region
(ii)	Use sign test when salary distribution is skewed	B1 **1** **(11)**	
6(i)	*N*: 0, 1, 2 *R* = 0: 0, *c*, 2*c* *R* = 1: 2*c*, 3*c*, 4c *R* = 2: 4*c*, 5*c*, 6*c* Total 27*c* = 1 $c = \frac{1}{27}$	B1 M1 A1 **3**	 Calculate 9 probs in terms of *c*
(ii)	9*c*/27*c* $= \frac{1}{3}$	M1 A1 ft **2**	Marginal probability AEF; ft *c*
(iii)	P(*N* + *R* > 2) = 15*c*/27*c* = $\frac{5}{9}$	M1 A1 ft **2**	 AEF; ft *c*
(iv)	P(*R*=2) = $\frac{15}{27}$ P(*N* | *R*=2): $p_0 = \frac{4}{15}$, $p_1 = \frac{1}{3}$, $p_2 = \frac{2}{5}$ E(*N* |*R*=2)= $1\times\frac{1}{3}+2\times\frac{2}{5}$ $= \frac{17}{15}$	M1 A1 ft A1 ft A1 **4**	Using conditional probabilities One value; ft values in **(i)** All values Or 1.13
(v)	Eg P(*N* = 0 and *R* = 0) = 0 P(*N*=0)×P(*R*=0) = $\frac{6}{27}\times\frac{3}{27} \neq 0$ So *N* and *R* are not independent	M1 A1 **2** **(13)**	Or from conditional probs M0 from *N*=1 with *R*=1 or 2 All correct

7(i)	$\int_0^{2\theta} \frac{x^{n+1}}{2\theta^2}\,dx = \left[\frac{x^{n+2}}{2(n+2)\theta^2}\right]$ $= 2^{n+1}\theta^n/(n+2)$ $E(X) = 4\theta/3$	M1 A1 B1 ft **3**	Correct integral AEF B0 if not 'deduced'
(ii)	$Var(X) = 2\theta^2 - (4\theta/3)^2 = 2\theta^2/9$ $Var(X^2) = E(X^4) - (E(X))^2$ $= 16\theta^4/3 - 4\theta^4 = 4\theta^4/3$	M1A1ft M1A1ft **4**	ft **(i)** with no n ft **(i)** with no n
(iii)	$E(\sum X_i) = 3\times 4\theta/3$ $= 4\theta$ $T_1 = ¼\Sigma X_i$ $E(\sum X_i^2) = 3\times 2\theta^2$ $=6\theta^2$ $T_2 = (\sum X_i^{\ 2})/27$	M1 A1 ft A1 ft M1 A1 ft A1 ft **6**	ft with no n ft with no n or θ ft with no n ft with no n or θ
(iv)	$Var(T_2) = 1/27^2 \times 3 \times Var(X^2)$ $= 4\theta^4/729$	M1 A1 **2** **(15)**	 CAO
8(i)	$P(L\cap M) = P(L\mid M)P(M) = 0.12$ and $P(L) = P(M\cap L)/P(M\mid L)$ = 0.12/0.4 = 0.3 $P(L'\cup M') = P\left[(L\cap M)'\right]$ $=1-P(L\cap M)$ = 1 – 0.2 × 0.6 = 0.88	*M1* A1 B1 **3**	
(ii)	$P(N\mid L\cap M) = 0.3$ $\Rightarrow P(N\cap L\cap M) = 0.3\times 0.12$ =0.036 $P(L'\cup M'\cup N') = 1-0.036 = 0.964$	M1 A1 A1 **3** **[6]**	