OCR
Oxford Cambridge and RSA

Wednesday 14 June 2017 – Morning

A2 GCE MATHEMATICS

4733/01 Probability & Statistics 2

QUESTION PAPER

Candidates answer on the Printed Answer Book.

OCR supplied materials:
- Printed Answer Book 4733/01
- List of Formulae (MF1)

Other materials required:
- Scientific or graphical calculator

Duration: 1 hour 30 minutes

INSTRUCTIONS TO CANDIDATES

These instructions are the same on the Printed Answer Book and the Question Paper.

- The Question Paper will be found inside the Printed Answer Book.
- Write your name, centre number and candidate number in the spaces provided on the Printed Answer Book. Please write clearly and in capital letters.
- **Write your answer to each question in the space provided in the Printed Answer Book.** Additional paper may be used if necessary but you must clearly show your candidate number, centre number and question number(s).
- Use black ink. HB pencil may be used for graphs and diagrams only.
- Answer **all** the questions.
- Read each question carefully. Make sure you know what you have to do before starting your answer.
- Do **not** write in the barcodes.
- You are permitted to use a scientific or graphical calculator in this paper.
- 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.

INFORMATION FOR CANDIDATES

This information is the same on the Printed Answer Book and the Question Paper.

- The number of marks is given in brackets **[ ]** at the end of each question or part question on the Question Paper.
- **You are reminded of the need for clear presentation in your answers.**
- The total number of marks for this paper is **72**.
- The Printed Answer Book consists of **8** pages. The Question Paper consists of **4** pages. Any blank pages are indicated.

INSTRUCTION TO EXAMS OFFICER/INVIGILATOR

- Do not send this Question Paper for marking; it should be retained in the centre or recycled. Please contact OCR Copyright should you wish to re-use this document.

Answer **all** the questions.

1 The governors of a school wish to investigate the opinions of the parents and guardians of the pupils. The secretary of the governors distributes a questionnaire to all parents and guardians who are present at a particular Parents' Evening.

(i) Explain why this method of sampling may not give reliable results. **[1]**

(ii) Suggest a better method of sampling, using random numbers. **[2]**

2 A continuous random variable W has the distribution $N(\mu, \sigma^2)$. It is given that $P(W < 70.0) = 0.8$ and $P(W > 81.0) = 0.1$. Find the value of a if $P(W < a) = 0.1$. **[7]**

3 The number of mistakes made in one session by teams of code transmitters is known to be a random variable with the distribution $Po(\lambda)$. In the past it has been found that $\lambda = 13$. Over a holiday period a new team is used and on one randomly chosen session the new team makes 23 mistakes. Test at the 1% significance level whether the new team makes on average more than 13 mistakes in one session. **[7]**

4 The acidity of paper is measured on the numerical pH scale. It is known that the writing paper generally used by a certain author has a mean pH of 6.3. The pH, X units, of a random sample of 36 pieces of paper thought to have been used by this author was measured, and the results are summarised as follows.

$$n = 36 \qquad \sum x = 222.48 \qquad \sum x^2 = 1380.5264$$

(i) Test at the 5% significance level whether the pH of the paper from which this sample is drawn differs from 6.3. **[11]**

(ii) State where the Central Limit Theorem was used in your test in part **(i)**. **[1]**

5 One game of roulette consists of throwing a ball onto a spinning wheel, which has 37 slots, numbered 0 to 36. On each throw the ball is equally likely to come to rest in any one of the 37 slots on the wheel, independently of all other throws. The number of the slot in which the ball comes to rest is the score of the game.

(i) Use a suitable approximation to find the probability that on 74 throws of the ball a score of 0 is obtained at least four times. Justify your approximation. **[4]**

(ii) 18 of the 37 possible scores are called "impair". 148 games are played. The probability that impair is obtained at least N times is less than 0.0025. Use a suitable approximation to find the smallest possible value of N. **[6]**

6 The editor of a scientific journal receives articles, for possible publication, at random times throughout the year.

(i) State two conditions needed for the number of articles received in one randomly chosen week to be well modelled by a Poisson distribution. **[2]**

Assume now that the number of articles received in one week can be modelled by the distribution Po(2.4).

(ii) Find the probability that in a randomly chosen 2-week period fewer than 5 articles are received. **[2]**

(iii) The number of articles received in one randomly chosen week is denoted by R. Given that $\mathrm{P}(R = r) = 2.5 \times \mathrm{P}(R = r + 1)$, use an algebraic method to find the value of r. **[4]**

(iv) Use an appropriate approximation to find the probability that in a randomly chosen 50-week period at least 140 articles are received. **[5]**

7 A continuous random variable X has probability density function

$$\mathrm{f}(x) = \begin{cases} \frac{1}{64}x(16 - x^2) & 0 \leqslant x \leqslant 4, \\ 0 & \text{otherwise.} \end{cases}$$

(i) Find the value of $\mathrm{E}(X)$. **[3]**

(ii) The upper quartile q is defined by the condition $\mathrm{P}(X \leqslant q) = \frac{3}{4}$.

(a) Show that q satisfies the equation

$$q^4 - 32q^2 + 192 = 0.$$ **[3]**

(b) Hence find the exact value of q. **[3]**

8 A random variable X has the distribution $\mathrm{B}(60, p)$. A hypothesis test is to be carried out, at the 5% significance level, of the null hypothesis H_0: $p = 0.95$ against the alternative hypothesis H_1: $p > 0.95$.

(i) Explain why a normal approximation cannot be used. **[1]**

(ii) Verify that the critical region for the test is $X = 60$. **[4]**

(iii) State the value(s) of p for which a Type I error could occur, and give the corresponding probability or probabilities of a Type I error. **[2]**

(iv) Find the range of values of p for which the probability that a Type II error occurs is less than 0.6. **[4]**

END OF QUESTION PAPER

1	(i)	Biased against those not at the parents' evening	B1 **1**	Reason for being biased or unrepresentative, needs more than "not all will be at the meeting", e.g. "not all will return the questionnaire" or "those at the meeting may have different opinions"	"Biased" can be implied by the reason Not *just* "not random" or "not representative", but allow "self-selecting". Ignore irrelevancies (e.g. "small sample", but withhold if definitely wrong comment seen
	(ii)	Obtain list of parents/pupils & number it 1 to n Select using random numbers, ignoring repeats/numbers outside range	B1 B1 **2**	Number (a list of) parents (sequentially) (statements in brackets can be implied) Mention use of RNs, as *only* method, *and* either "ignore repeats" or "ignore outside range" (allow "use RNs in range")	SC: Allocate *random* numbers: max B1 unless *sorted* *Not* "select numbers randomly" *Not* hat/lottery machine [RNs required by question] Allow systematic provided random start
2		$\frac{70-\mu}{\sigma}=0.842$; $\frac{81-\mu}{\sigma}=1.282$ $11 = 0.44\sigma$ $[\sigma = 25]$ $\mu = 48.93$ $81 - \mu = \mu - a$ $a = 16.9$	M1* A1 B1 dep*M1 A1 M1 A1 **7**	Stand'ise once & equate to Φ^{-1}, allow sign/√/cc errors LHS both correct, signs consistent on both sides Both z values, $\in$ [0.841, 0.842] and [1.281, 1.282] Solve to get μ or σ Either, $\mu \in$ [48.9, 49(.0)], $\sigma \in$ [25, 25.1], www Equation for a, correct signs [may involve 1.282] a in range [16.8, 17], www	"P(> 81) = P(≥ 80)" etc or "1 – 1.282" is M1A0 Can get M1A1 even if z wrong provided they <u>are</u> z Can award B1 even if signs are wrong e.g. $\mu - a = 1.282\sigma$ [a needs to be less than their μ]
3		H_0: $\lambda = 13$, H_1: $\lambda > 13$ α: P(≥ 23) = 1 – 0.9924 = 0.0076 < 0.01	B2 M1 A1 A1	One error, B1, but x, t etc or 23: B0 Find P(≥ 23, > 23, = 23, < 23, ≤ 23) from Po(13) P(≥ 23) = 0.0076 ONLY (but see SC below] Explicit comparison with 0.01 SC: 0.9924 > 0.99: M1A2 and can get last M1A1	H_0: $\mu = 13$, H_1: $\mu > 13$ gets full marks 0.9924, 0.996, 0.0040, 0.0036: M1A0A0M0A0 SC: "P(≥ 23) = 1 – P(≤ 23)": M1A0A1M1A1ft SC: 1 – 0.9970 = 0.003 or 1 – 0.9833 = 0.0167 *[from λ = 12 or 14]*: A0 but can get all other marks
		β: CR ≥ 23 Probability 0.0076 23 in CR	B1dep* dep*B1 dep*B1	Must be clearly stated [*not* just "CV = 23"] Must be seen, but allow 0.9924 [e.g. on diagram] Must be stated explicitly	SC: If *mixture* of methods: maximum (B2)M1 [or (B2)B1], M1A1, max 5/7 *Second and third B1 are independent of each other* SC: CR ≥ 24, p = 0.0093 B1* (for both) 23 not in CR, DNR B1dep, M1A1 *or* CR ≥ 22, p = 0.0061 B1* (for both) 23 in CR, reject B1dep, M1A1 *[from λ = 12 or 14]* *either could get 6/7*
		Reject H_0. There is significant evidence that the new team makes more mistakes.	M1 A1 **7**	First conclusion consistent, needs correct method Contextualised, acknowledge uncertainty	

4	(i)	$\hat{\mu} = \bar{x} = 6.18$ $\hat{\sigma}^2 = \frac{36}{35}\left(\frac{1380.5264}{36} - 6.18^2\right)$ $= 0.16$ H_0: $\mu = 6.3$, H_1: $\mu \neq 6.3$ α: $z = \frac{6.18 - 6.3}{\sqrt{0.16/36}} = -1.8$, $p = 0.0359$ $-1.8 > -1.96$ or $0.0359 > 0.025$	B1 M1 M1 A1 B2 M1 A1 A1	6.18 seen somewhere Correct formula for biased estimate Multiply by 36/(36 – 1) 0.1556: M1M0A0. Allow e.g. 5.6/35 One error, B1, but x, $\bar{x}$, t : B0 Standardise, 36 needed (if omitted, no more marks in (i) 1.8 or –1.8 or a.r.t. 0.0359 Compare $-z$ with –1.96 or z with 1.96 or p with 0.025, like-with-like	Single formula: M2 if right, M1 if wrong but with 35 divisor *somewhere* H_0: $\lambda = 6.3$, H_1: $\lambda \neq 6.3$: B1 u rather than μ: B1B0 if unquestionably u, else BOD Allow 0.9641 *only* if compared with 0.95 or 0.975 Wrong or no notation (e.g. "cdfnorm"): full marks if right, M0A0A0 (M1A1) if numbers wrong in any way
		β: CV $6.3 - 1.96\sqrt{\frac{0.16}{36}} = 6.1693$ $6.18 > 6.1693$	M1 A1 A1	$6.3 - z\sqrt{(\sigma^2/36)}$, allow √ errors, cc, ± $z = 1.96$ Compare $\bar{x}$ with 6.17 (or with 6.19 from 1-tail)	$6.18 + z\sqrt{(\sigma^2/36)}$: M1 and no further marks in (i) 6.17 (and no working) can imply mark for 1.96 SC 1-tail: $6.18 > 6.19$, reject H_0, etc: M1A0A1, M1A1
		Do not reject H_0. Insufficient evidence that pH of paper is not 6.3	M1 A1ft **11**	Requires essentially correct method, 36 divisor, like-with-like, hypotheses involving 6.3 Contextualised, acknowledge uncertainty. Allow "insufficient evidence that pH hasn't changed"	Withhold A1 if no context or too assertive, e.g. "evidence that pH of paper is 6.3"
	(ii)	In comparing z with z_{crit}	B1 **1**	Or in using 1.96 for CV, etc, or "in assuming that the sample (mean) is normally distributed" (must answer "where?", mustn't leave it vague as to whether it's X or $\bar{X}$)	No extra answers. "Calculating variance" or "dividing by $\sqrt{n}$" : B0. Not "because the population is not known to be normally distributed". But allow if OK *and also* explained why it can be used.
5	(i)	B(74, 1/37) ≈ Po(2) P(≥ 4) = 1 – P(≤ 3) = 1 – 0.8571 = 0.1429 $n > 50$, (np =) $2 < 5$	M1 M1 A1 B1 **4**	Po(2) stated or implied RH tail of their Poisson Answer, a.r.t. 0.143 Both conditions, dep on Poisson Allow "n large, p small" or "n large, $np < 5$". FT on their $74 \times (1/37)$ if less than 5	If formula used, must be correct formula, and > 1 term 0.14037 (from exact binomial): 0 np condition needs their 2 seen *somewhere* If numbers used, must be compared with 50 and 5 Extra or wrong conditions (e.g. nq): B0 *If MR leading to np > 5, please consult TL*
	(ii)	B(148, 18/37) ≈ N(72, 36.973) $72 + 0.5 + 2.807 \times \sqrt{36.973}$ = 89.57 Hence $N_{min} = 90$	M1 A1 M1 B1 A1 A1 **6**	N(np, …) attempted Both parameters correct, allow √ errors [$= \frac{1368}{37}$] $72 + z\sigma$, allow σ^2 and/or no cc, do not allow $\sqrt{n}$ divisor $z = 2.807$ or 2.808 or 2.809 or a.r.t. 2.81, allow –ve [can be implied by 89.6 or 89.1 or 88.9] 89.6 or 89.1 seen or implied 90 only, www (although cc can be omitted if rounded up)	• Must see evidence of appropriate approximation • No working: 0 (approximation required) • 0.00196 implies exact binomial: 0 • BUT $\sqrt{37}$ may be $\sqrt{36.97}$, not $\sqrt{n}$ Ignore inequalities until final mark Allow from 89.07 and checked using exact binomial

6	(i)	Articles must be received independently of one another and at constant average rate	B1 B1 **2**	At least one must be contextualised for *any* marks Independent stated, allow "probability independent" Allow "uniform" rate but not "constant" rate *Not* "probability is constant"	If extras, e.g. "singly" or "randomly", then max 1 Allow "receipt of one doesn't affect receipt of another" Any implication of regularity: *can't* get second B mark
	(ii)	Po(4.8), left-hand tail P(< 5) = 0.4763	M1 A1 **2**	Po(2×2.4), stated or implied, e.g. by 0.6510	0.4946 or 0.4582: M1A0
	(iii)	$e^{-2.4}\frac{2.4^r}{r!} = 2.5\times e^{-2.4}\frac{2.4^{r+1}}{(r+1)!}$ $r + 1 = 2.5\ \ 2.4$ $r = 5$	M1* A1 *depM1 A1 **4**	One correct probability from formula Correct equation, allow from $\lambda = 4.8$ [= MR] Simplify exp and ! correctly to linear equation in r $r = 5$ only, www	SC: T&I, or tables or calculator: 0 $r + 1$!: allow if used as if it were $(r + 1)$! Wrong use of logs: M0A0
	(iv)	Po(120) ≈ N(120, 120) $1-\Phi\left(\frac{139.5-120}{\sqrt{120}}\right)$ $= 1 - \Phi(1.78) = 0.0375$	M1 A1 M1 A1 A1 **5**	Normal stated or implied, mean 50×2.4 Both parameters correct, allow √ errors Standardise, allow no/wrong cc and/or √ errors, can be implied by correct answer Both cc and correct √ Answer, anything rounding to 0.0375	If answer wrong, do not give M1 unless correct *notation* used (i.e., calculator notation such as "cdfnorm" does not qualify for M1 even if answer is recognisable)
7	(i)	$\int_0^4 \frac{1}{64}x^2(16-x^2)\mathrm{d}x =$ $\left[\frac{1}{64}\left(\frac{1}{3}.16x^3-\frac{1}{5}x^5\right)\right]_0^4$ *or* $\left[\frac{1}{12}x^3-\frac{1}{320}x^5\right]_0^4$ $\frac{32}{15}$ or 2.133…	M1 B1 A1 **3**	Attempt to integrate xf(x), correct limits *somewhere* Correct indefinite integral, aef Answer, exact or anything rounding to 2.13	Allow numerical, but if algebraic, powers must increase to score M1
	(ii)	(a) $\int_0^q \frac{1}{64}x(16-x^2)\mathrm{d}x = \frac{3}{4}$ $\frac{1}{64}\left(8q^2-\frac{1}{4}q^4\right) = \frac{3}{4}$ *or* $\frac{1}{8}q^2-\frac{1}{256}q^4 = \frac{3}{4}$ $q^4 - 32q^2 + 192 = 0$	M1 A1 A1 **3**	Correct integral, limits 0, q Correct unsimplified equation Correctly obtain given equation, www	Allow limits q, 4 if equated to ¼, not otherwise Withhold if "simplified", e.g. to $-32q^6 + 192 = 0$
		(b) $q^2 = 8$ or 24 $q = \pm\sqrt{8}$ or $\pm\sqrt{24}$ $q = \sqrt{8} = 2\sqrt{2}$	M1 A1 A1 **3**	Solve for q^2 Acknowledge other answers (or explicitly reject) $q = \sqrt{8}$ or $2\sqrt{2}$ as sole answer, allow 2.83 as sole final answer *only if* exact also seen somewhere	Numerical only: give M1 for 2.83 (2.828…) seen √8, or not ±, and no other comment: M1A0A1

8	(i)	$nq = 3 < 5$, or p not close to ½ and n not large enough	B1 **1**	www. 3 and 5 must be seen if inequality used (No need to mention Poisson – ignore any mention)	Withhold if extra wrong statements seen but ignore irrelevant statements (e.g. np). Do *not* allow $npq = 2.85$
	(ii)	$P(X = 60) = 0.95^{60} = 0.046$ < 0.05 $P(X \geq 59) = 0.046 + 0.1455$ [= 0.19155] > 0.05	B1dep* dep*B1 B1dep† dep†B1 **4**	In range [0.0460, 0.0461], *or* [0.9539, 0.954] Correct tail explicitly compared In range [0.191, 0.192], *or* [0.808, 0.809] Correct tail explicitly compared (no final conclusion needed)	0.145 or 0.855 qualifies for these two marks SC: Po(3): $P(X = 60) = 0.0498$ B1dep* < 0.05, dep*B1 $P(X \geq 59) = 0.1991$ B1dep† > 0.05: dep†B1
	(iii)	$p = 0.95$ 0.046	B1 B1ft **2**	Question requires this to be stated FT on their 0.046 from (ii)	*NOT* 0.05. NB: if Po(3) used in (ii), 0.0498 gets B1
	(iv)	$P(\leq 59) < 0.6$ $1 - p^{60} < 0.6$ $p^{60} > 0.4$ $0.985 < p \leq 1$	M1 A1 A1 A1 **4**	Can be implied Range required, allow any combination of $<$/$\leq$ Withhold if more than 5 sf seen	$p^{60} > 0.6$ or $p^{60} < 0.6$: can get M1 if $P(\leq 59) < 0.6$ stated explicitly, otherwise 0 SC: T&I or tables: $p > 0.985$ or better, B4, else 0 Allow $p > 0.958$ SC: $60 - R \sim \text{Po}(\lambda)$: $1 - P(\leq 0) < 0.6$ M1 $e^{-\lambda} > 0.4$ A1 $(\lambda < 0.916)$ $60(1 - p) < 0.916$ M1, needs M1A1 $p > 0.985$ A1