OCR
Oxford Cambridge and RSA

Wednesday 15 June 2016 – 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 bar codes.
- 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 **12** 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.

1 The results of 14 observations of a random variable V are summarised by

$$n = 14, \qquad \Sigma v = 3752, \qquad \Sigma v^2 = 1\,007\,448.$$

Calculate unbiased estimates of E(V) and Var(V). **[4]**

2 The mass, in kilograms, of a packet of flour is a normally distributed random variable with mean 1.03 and variance σ^2. Given that 5% of packets have mass less than 1.00 kg, find the percentage of packets with mass greater than 1.05 kg. **[6]**

3 The random variable F has the distribution B(40, 0.65). Use a suitable approximation to find $P(F \leqslant 30)$, justifying your approximation. **[7]**

4 It is given that $Y \sim \text{Po}(\lambda)$, where $\lambda \neq 0$, and that $P(Y = 4) = P(Y = 5)$. Write down an equation for λ. Hence find the value of λ and the corresponding value of $P(Y = 5)$. **[5]**

5 55% of the pupils in a large school are girls. A member of the student council claims that the probability that a girl rather than a boy becomes Head Student is greater than 0.55. As evidence for his claim he says that 6 of the last 8 Head Students have been girls.

(i) Use an exact binomial distribution to test the claim at the 10% significance level. **[7]**

(ii) A statistics teacher says that considering only the last 8 Head Students may not be satisfactory. Explain what needs to be assumed about the data for the test to be valid. **[1]**

6 The number of cars passing a point on a single-track one-way road during a one-minute period is denoted by X. Cars pass the point at random intervals and the expected value of X is denoted by λ.

(i) State, in the context of the question, two conditions needed for X to be well modelled by a Poisson distribution. **[2]**

(ii) At a quiet time of the day, $\lambda = 6.50$. Assuming that a Poisson distribution is valid, calculate $P(4 \leqslant X < 8)$. **[3]**

(iii) At a busy time of the day, $\lambda = 30$.

(a) Assuming that a Poisson distribution is valid, use a suitable approximation to find $P(X > 35)$. Justify your approximation. **[6]**

(b) Give a reason why a Poisson distribution might not be valid in this context when $\lambda = 30$. **[1]**

7 A continuous random variable X has probability density function

$$f(x) = \begin{cases} ax^{-3} + bx^{-4} & x \geqslant 1, \\ 0 & \text{otherwise,} \end{cases}$$

where a and b are constants.

(i) Explain what the letter x represents. **[1]**

It is given that $P(X > 2) = \frac{3}{16}$.

(ii) Show that $a = 1$, and find the value of b. **[7]**

(iii) Find $E(X)$. **[3]**

8 It is known that the lifetime of a certain species of animal in the wild has mean 13.3 years. A zoologist reads a study of 50 randomly chosen animals of this species that have been kept in zoos. According to the study, for these 50 animals the sample mean lifetime is 12.48 years and the population variance is 12.25 years^2.

(i) Test at the 5% significance level whether these results provide evidence that animals of this species that have been kept in zoos have a shorter expected lifetime than those in the wild. **[7]**

(ii) Subsequently the zoologist discovered that there had been a mistake in the study. The quoted variance of 12.25 years^2 was in fact the sample variance. Determine whether this makes a difference to the conclusion of the test. **[5]**

(iii) Explain whether the Central Limit Theorem is needed in these tests. **[1]**

9 The random variable R has the distribution $\text{Po}(\lambda)$. A significance test is carried out at the 1% level of the null hypothesis $H_0: \lambda = 11$ against $H_1: \lambda > 11$, based on a single observation of R. Given that in fact the value of λ is 14, find the probability that the result of the test is incorrect, and give the technical name for such an incorrect outcome. You should show the values of any relevant probabilities. **[6]**

END OF QUESTION PAPER

Question		Answer/Indicative content	Marks	Guidance
1		$\hat{\mu} = \bar{x} = \frac{3752}{14}$ **= 268** $\frac{1007448}{14} - \bar{x}^2$ [=136.57...] $\times \frac{14}{13}$; **= 147**(.07…)	B1 M1 M1 A1 **4**	268 only, must be stated separately, *not* isw If single formula used, give M1 for divisor 13 anywhere Multiply by 14/13 Answer, a.r.t. 147, or $\frac{1912}{13} = 147\frac{1}{13}$ MR 3572: 255.14, 7390.6 gets B0M1M1A1
2		$\frac{1.03-1.00}{\sigma} = 1.645$ [$\sigma = 0.0182\ldots \approx {}^{6}/_{329}$] $1 - \Phi\left(\frac{1.05-1.03}{\sigma}\right) = 1 - \Phi(1.0966)$ = 1 – 0.8635 = **0.1365** or 13.6(5)%	M1dep* A1 B1 *M1 M1 A1 **6**	Standardise and equate to Φ^{-1}, allow wrong sign, σ^2, 1–, cc etc All correct apart possibly from value of Φ^{-1} 1.645 seen anywhere, allow –1.645, can be implied Solve to find σ, or eliminate σ, dependent on first M1 Standardise with $\mu = 1.03$, use Φ, answer < 0.5, allow √ errors Final answer in range [0.1355, 0.137] or [13.55%, 13.7%], must be from positive σ, not from σ^2 0.1333 from $\sigma = 0.018$ is 5+A0
3		N(26, 9.1) $\Phi\left(\frac{30.5-26}{\sqrt{9.1}}\right) = \Phi(1.492)$ = **0.9322** "$np > 5$" or "n large" stated "14 > 5" or "p close to ½" stated	M1 A1 M1 A1 A1 B1 B1 **7**	Normal, mean their attempt at 40×0.65 Mean 26 and variance or SD 9.1 Standardise, their np, npq, no √n, allow cc or √ errors cc and √(their npq) correct Final answer, a.r.t. 0.932 One condition asserted Complementary condition, if "nq" must see 14 somewhere. *Not npq* [Thus: "$np > 5$, $nq > 5$": B1B0] Extra conditions, e.g. "$n > 30$": max B1B0 SC: Exact (0.935564): maximum B1B1
4		$\frac{\lambda^4}{4!}e^{-\lambda} = \frac{\lambda^5}{5!}e^{-\lambda}$ $\frac{\lambda^4}{4!} = \frac{\lambda^5}{5!}$ $\Rightarrow \boldsymbol{\lambda = 5}$ **0.175**(46)	M1 A1 M1 A1 B1 **5**	Poisson formula used [*not* just quoted] correctly once This equation or exact equivalent, needs $e^{-\lambda}$ seen somewhere Correct method for cancelling $e^{-\lambda}$ Solve to get $\lambda = 5$ only, www Probability, in range [0.175, 0.176], allow from $\lambda = 5$ from wrong working

Question		Answer/Indicative content	Marks		Guidance
5	(i)	H_0: $p = 0.55$, H_1: $p > 0.55$ R ~ B(8, 0.55) where R is the number of girls α: $P(R \geq 6) = 1 - 0.7799 = 0.2201$ > 0.1	B2 M1 A1 B1		All correct, B2. One error (e.g. ≠, wrong or no letter) B1, but r, x etc: B0 B(8, 0.55) stated or implied, e.g. N(4.4, 1.98) P(≥ 6) = 0.2201, or P(< 6) = 0.7799 Compare P(≥ 6) with 0.1 or P(< 6) with 0.9
		β: CR is ≥ 7 and 6 < 7 $p = 0.0632$	B1 A1		Correct CR stated and explicit comparison with 6 This probability seen, a.r.t. 0.0632. Award if 0.9368 seen and CR is correct. If CR not clearly stated, cannot get last M1A1
		Do not reject H_0. There is insufficient evidence that the girls are proportionately more likely to become Head Student.	M1 A1	**7**	Correct first conclusion, requires B(8, 0.55), *not* P(> 6) [= 0.0632] or P(≤ 6) [= 0.9368] or P(= 6) [= 0.1569]. Allow 0.7799 if compared with 0.9 Interpreted, in context, acknowledge uncertainty, double negative. SC: Normal: max B2 M1 SC: Two different attempts: max B2 M1 unless both correct
	(ii)	Assume that the last 8 years are a random sample of years when Head Student has been chosen	B1	**1**	Refer to random sample, allow implied by any method described Must be choosing *years*, not *students* *Not* quote conditions for random sample unless explicitly "years" Extras: ignore unless clearly wrong, in which case B0
6	(i)	Cars pass independently of one another and at constant average rate	B1 B1	**2**	"Independently", refer to cars. Not "constant rate", "constant probability". No extra conditions. Ignore all references to "singly" (which is *wrong* in this context!)
	(ii) α	P(≤ 7) – P(≤ 3) = 0.6728 – 0.1118 = **0.561**(0)	M1 A2	**3**	0.680 or 0.681: M1A0 Allow from calculator, no working 0.4491 or 0.5679: M1A1 Allow from calculator, no working
	or β	P(4) + P(5) + P(6) + P(7) = 0.1118 + 0.1454 + 0.1575 + 0.1462 = **0.561**(0)	M1 A1 A1	**3**	Correct formula for ≥ 3 probabilities from Po(6.5) added, can be implied 3, 4 or 5 correct terms (e.g. P(3) = 0.06880), can be algebraic or implied Answer, a.r.t. 0.561
	(iii)(a)	Po(30) ≈ N(30, 30) $1 - \Phi\left(\frac{35.5 - 30}{\sqrt{30}}\right) = 1 - \Phi(1.004)$ = 1 – 0.8422 = **0.1578** Normal suitable as 30 > 15	M1 A1 M1 A1 A1 B1	**6**	Normal, mean 30, stated or implied Variance or SD 30 Standardise, their λ, λ, allow wrong/no cc, var/SD errors cc, $\sqrt{\lambda}$ correct Answer, a.r.t. 0.158 [NB: 0.157 may be from exact. See below] Or "λ large", etc., but no other conditions. If numerical comparison, must involve 15. SC: Exact Poisson, 0.1574, max B1 SC: Po(30), N(15, 15): M0B1 M1A1A0 B1, max 4/6
	(b)	Cars do not pass independently/randomly, as one may be immediately followed by another	B1	**1**	Any plausible relevant explanation in context, needn't be connected to conditions, e.g. "steady stream". *Not* "several cars might pass at once". Allow explanations that might also hold for smaller λ Do not allow comment on size of λ unless explained in valid way, e.g. "λ too large so cars follow one another", but not "λ too large for Poisson"

Question		Answer/Indicative content	Marks	Guidance
7	(i)	[x represents a] possible value(s) taken by X	B1 **1**	Must refer to, or imply, both x and X or "the random variable" Ignore extra unless definitely wrong
	(ii)	$\int_2^\infty ax^{-3}+bx^{-4}\mathrm{d}x=\left[-\frac{a}{2x^2}-\frac{b}{3x^3}\right]_2^\infty=\frac{a}{8}+\frac{b}{24}$ *or* $\int_1^\infty ax^{-3}+bx^{-4}\mathrm{d}x=\left[-\frac{a}{2x^2}-\frac{b}{3x^3}\right]_1^\infty=\frac{a}{2}+\frac{b}{3}$ *or* $\int_1^2 ax^{-3}+bx^{-4}\mathrm{d}x=\left[-\frac{a}{2x^2}-\frac{b}{3x^3}\right]_1^2=\frac{3a}{8}+\frac{7b}{24}$ $\frac{a}{2}+\frac{b}{3}=1$ or $\frac{a}{8}+\frac{b}{24}=\frac{3}{16}$ or $\frac{3a}{8}+\frac{7b}{24}=\frac{13}{16}$ Solve to get $a=1$ $b=\frac{3}{2}$	B1 M1 M1 M1 A1 A1 B1 **7**	Correct indefinite integral [from any set of limits or none] Integrate and substitute limits to obtain one expression Integrate and substitute limits to obtain a second expression The limits must be two of (1, ∞), (1, 2) or (2, ∞), allow (3, ∞) for "≥ 2" Equate two expressions from definite integrals to 1 or $\frac{3}{16}$ or $\frac{13}{16}$ as appropriate, and attempt to solve Both equations correct, any equivalent simplified form, can be implied ["simplified" = one a term, one b term, one number term] Correctly show $a = 1$ **AG**, www Correct value of b obtained from at least one correct equation SC: One equation only: M1B1 M0M0A0 A0B1, max 3/7 Two equations, assume $a = 1$, solve for b, checked in other equation: 7/7
	(iii)	$\int_1^\infty ax^{-2}+bx^{-3}\mathrm{d}x=\left[-\frac{a}{x}-\frac{b}{2x^2}\right]_1^\infty$ $\left\{=a+\frac{b}{2}\right\}$ $=1\frac{3}{4}$	M1 B1ft A1 **3**	Integrate xf(x), limits 1 and ∞ seen somewhere Correct indefinite integral, their b, can be implied by correct answer *Expect to see* $\int_1^\infty x^{-2}+\frac{3}{2}x^{-3}\mathrm{d}x=\left[-\frac{1}{x}-\frac{3}{4x^2}\right]_1^\infty$ Correctly obtain 1¾ or a.r.t. 1.75 www, allow from calculator

Question		Answer/Indicative content	Marks	Guidance
8	(i)	H_0: $\mu = 13.3$, H_1: $\mu < 13.3$	B2	Both correct: B2. One error [e.g. p, ≠, no symbol] B1, but x, $\bar{x}$ etc B0
	α:	$z = \frac{12.48 - 13.3}{\sqrt{12.25/50}} = -1.6566$ [$p = 0.0488$]	M1 A1	Standardise with √50, allow √ errors, allow cc, allow 13.3 – 12.48 z in range [–1.66, –1.65], or p in range [0.04875, 0.0489], allow 0.9512 only if consistent
		[12.25/50 = 0.245] < –1.645 [$p < 0.05$]	B1	Compare with –1.645, allow +1.6566 with +1.645, or p with 0.05/0.95 as consistent
	β:	CV $13.3 - 1.645\sqrt{\frac{12.25}{50}} = 12.4857\ldots$	M1 B1	13.3 – $z\sigma/\sqrt{50}$, any recognisable z, allow √ errors etc, ignore 13.3 + … $z = 1.645$
		12.48 < CV	A1	Compare 12.49 (or better) with 12.48, ignore 13.3 + … SC: 2-tailed, 12.33 gets B1B0 M1B0A1ft M1A1
		Reject H_0. Significant evidence that animals in zoos have shorter expected lifetime	M1 A1ft **7**	Consistent, needs √50, like-with-like comparison, hypotheses *not* 12.48 Contextualised, acknowledge uncertainty, their z SC1: 2-tailed: can get B1 M1A1B0 M1A1 max 5/7 SC2: No √50: can get B2 M0A0 B1 M0 max 3/7 SC3: $\bar{x}$ and μ confused consistently: can get B0 M1A1 B1 M0 SC4: 50/49 used in (i): can get B2 M1A0B1 M1A1 (6) in (i), M1 in (ii)
	(ii)	$\hat{\sigma}^2 = \frac{50}{49} \times 12.25$ [= 12.5]	M1	Multiply 12.25 by 50/49, allow √ etc, allow if done in part (i) but then 0
		$z = \frac{12.48 - 13.3}{\sqrt{12.5/50}} = -1.64$ [$p = 0.0505$]	M1 A1	Standardise with √50 Obtain a.r.t. –1.64, allow +1.64 if consistent with (i).
		> –1.645 [$p > 0.05$]	B1	Compare with same CV as in (i)
		Opposite conclusion	A1ft **5**	State opposite conclusion (ft), any form, allow $\bar{x}/\mu$ here , needs M1M1 *Identical mark scheme for method β, CV 12.4775* SC1: 50 omitted consistently in both: M1M0A0B1A1 max 3/5 SC2: no √50 in (i), √50 but not 50/49 in (ii): M0M1A0B1A1 max 3/5
	(iii)	Yes as population not known to be normal	B1 **1**	Not "n large" unless "Yes, not known normal, but n large so can use" No wrong extras, e.g. "depends on whether it's sample or population"
9		P(≤ 19 | $\lambda = 11$) = 0.9907	M1	Attempt to find critical region from $\lambda = 11$, allow even if tail wrong e.g. P(≤ 3) = 0.0049, allow from $\lambda = 10$ or 12
		so critical region ≥ 20	A1	Critical region is ≥ 20. [CV 19 or 20 can imply first M1] [If only one probability shown, assume this is CR]
		P(≥ 20) = 0.0093	A1	Probability 0.9907 or 0.0093 seen (even if CR is wrong) [from $\lambda = 11$]
		P(≤ 19 | $\lambda = 14$) = **0.9235**	M1 A1	Find P(not in CR | $\lambda = 14$), must now be LH tail, e.g. 0.8826 Answer in range [0.923, 0.924]
		Type II error	B1 **6**	"Type II error" stated, allow "Type 2" SC1: P(≤ 19) = 0.9907 so CR is ≥ 19: M1A0A1M1A0B1 max 4/6 SC2: $\lambda = 14$ used throughout, e.g. P(≥ 23) = 0.0093: max B1