OCR
Oxford Cambridge and RSA

Friday 5 June 2015 – 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.

DC (LK) 102437/4

1 The random variable Y is normally distributed with mean μ and variance σ^2. It is found that $P(Y > 150.0) = 0.0228$ and $P(Y > 143.0) = 0.9332$. Find the values of μ and σ. **[6]**

2 A class investigated the number of dead rabbits found along a particular stretch of road.

(i) The class agrees that dead rabbits occur randomly along the road. Explain what this statement means. **[1]**

(ii) State, in this context, an assumption needed for the number of dead rabbits in a fixed length of road to be modelled by a Poisson distribution, and explain what your statement means. **[2]**

Assume now that the number of dead rabbits in a fixed length of road can be well modelled by a Poisson distribution with mean 1 per 600 m of road.

(iii) Use an appropriate formula, showing your working, to find the probability that in a road of length 1650 m there are exactly 3 dead rabbits. **[3]**

3 A continuous random variable X has probability density function

$$f(x) = \begin{cases} \frac{3}{2a^3}x^2 & -a \leqslant x \leqslant a, \\ 0 & \text{otherwise,} \end{cases}$$

where a is a constant.

(i) It is given that $P(-3 \leqslant X \leqslant 3) = 0.125$. Find the value of a in this case. **[4]**

(ii) It is given instead that $\text{Var}(X) = 1.35$. Find the value of a in this case. **[5]**

(iii) Explain the relationship between x and X in this question. **[1]**

4 A continuous random variable is normally distributed with mean μ. A significance test for μ is carried out, at the 5% significance level, on 90 independent occasions.

(i) Given that the null hypothesis is correct on all 90 occasions, use a suitable approximation to find the probability that on 6 or fewer occasions the test results in a Type I error. Justify your approximation. **[5]**

(ii) Given instead that on all 90 occasions the probability of a Type II error is 0.35, use a suitable approximation to find the probability that on fewer than 29 occasions the test results in a Type II error. **[5]**

5 **(i)** State an advantage of using random numbers in selecting samples. **[1]**

(ii) It is known that in analysing the digits in large sets of financial records, the probability that the leading digit is 1 is 0.25. A random sample of 18 leading digits from a certain large set of financial records is obtained and it is found that 8 of the leading digits are 1s. Test, at the 5% significance level, whether the probability that the leading digit is 1 in this set of records is greater than 0.25. **[7]**

6 Records for a doctors' surgery over a long period suggest that the time taken for a consultation, T minutes, has a mean of 11.0. Following the introduction of new regulations, a doctor believes that the average time has changed. She finds that, with new regulations, the consultation times for a random sample of 120 patients can be summarised as

$$n = 120, \Sigma t = 1411.20, \Sigma t^2 = 18\,737.712.$$

(i) Test, at the 10% significance level, whether the doctor's belief is correct. **[11]**

(ii) Explain whether, in answering part **(i)**, it was necessary to assume that the consultation times were normally distributed. **[1]**

7 A large railway network suffers points failures at an average rate of 1 every 3 days. Assume that the number of points failures can be modelled by a Poisson distribution. The network employs a new firm of engineers. After the new engineers have become established, it is found that in a randomly chosen period of 15 days there are 2 instances of points failures.

(i) Test, at the 5% significance level, whether there is evidence that the mean number of points failures has been reduced. **[7]**

(ii) A new test is carried out over a period of 150 days. Use a suitable approximation to find the greatest number of points failures there could be in 150 days that would lead to a 5% significance test concluding that the average number of points failures had been reduced. **[6]**

8 The random variable S has the distribution B(14, p). A significance test is carried out of the null hypothesis H_0: $p = 0.3$ against the alternative hypothesis H_1: $p > 0.3$. The critical region for the test is $S \geqslant 8$.

(i) Find the significance level of the test, correct to 3 significant figures. **[2]**

(ii) It is given that, on each occasion that the test is carried out, the true value of p is equally likely to be 0.3, 0.5 or 0.7, independently of any other test. Four independent tests are carried out. Find the probability that at least one of the tests results in a Type II error. **[5]**

END OF QUESTION PAPER

Question			Answer	Marks	Guidance	
1			$\frac{150-\mu}{\sigma} = 2.00$ $\frac{143-\mu}{\sigma} = -1.5$ Solve to get $\mu = 146$, $\sigma = 2$	M1 A1 B1 M1 A1 A1 **6**	Standardise with σ, μ at least once, ignore cc, √ errors, equate to z Both LHS and signs of RHS correct Both z-values correct to 3 SF Correct method for solution $\mu \in$ [145.95, 146.05) www $\sigma \in$ [1.995, 2.005) or $\sigma^2 = 4$ www	z not used, e.g. equated to 0.0228 and 0.9332 or 0.5092 and 0.8246: max M0M1 One z, one not: M1A0B0 Withhold if elimination done wrongly $\sqrt{\sigma}$ or σ^2: can get M1A0B1M1A1A0 cc: M1A0B1M1A0A0
2	(i)		That they don't occur regularly or to a fixed pattern, or are unpredictable	B1 **1**	Any similar or equivalent statement, but *not* independent or equivalent Both right and wrong: B0	E.g. "no pattern": expect to be right E.g. "doesn't affect": expect to be wrong
	(ii)		Dead rabbits occur independently, i.e., one occurrence does not affect the probability of another *or* at constant average rate, i.e. mean number uniform along the whole road	B1 B1 **2**	Correct statement of principle Correct interpretation of that principle Context needed for any marks **SR:** "Constant rate" B0, correct reason can get B1 if "average" implied	*Not* "constant probability" One right, one wrong, e.g. independent + "$np < 5$, $nq < 5$": max 1 Only "Singly" stated, implied or used: max B1 Right condition but explanation shows it's wrong: B0B0
	(iii)		Po(2.75) $e^{-2.75}\frac{2.75^3}{3!} = \mathbf{0.2215}$	M1 M1 A1 **3**	Po(1650/600) attempted Correct formula, any numerical λ Answer in range [0.221, 0.222]	Needs evidence for this Must be seen Formula required, so no formula ⇒ M0A0
3	(i)		$\int_{-3}^{3}\frac{3}{2a^3}x^2dx = \left[\frac{x^3}{2a^3}\right]_{-3}^{3} = \frac{27}{a^3}$ = 0.125 so $\boldsymbol{a = 6}$	M1dep* B1 *M1 A1 **4**	Integrate, attempt at correct seen limits *somewhere* Correct indefinite integral, can be implied by, e.g. $27/a^3$ Equate, with limits, to 0.125 and solve Solve to get $a = 6$ exactly	Allow e.g. "< 3" = "≤ –4" Allow also for a^3 on top Allow 6.00 but no other decimals. *Not* ±6
	(ii)		$\mu = 0$ $\int_{-a}^{a}kx^4dx = \left[k\frac{x^5}{5}\right]_{-a}^{a} = \frac{3a^2}{5}$ = 1.35 so $\boldsymbol{a = 1.5}$	B1 M1dep* B1 *M1 A1 **5**	Stated somewhere or calculated, any a Attempt to integrate x^2f(x), limits $\pm a$ Or exact equivalent, can be implied Equate to 1.35 and solve $a = 1.5 \pm 0.005$, allow ±1.5, ignore "must be positive"	If $\mu = 0$ not mentioned anywhere, or "$-\mu$" stated [instead of "$-\mu^2$"], B0 but can get remaining 4/5 Don't need explicit $-\mu^2$ here NB: $a = 3$ is *not* MR but can get B1 for $\mu = 0$
	(iii)		x is a value [values] that X takes	B1 **1**	Ignore irrelevancies or extra wrong, unless contradictory	*Not* answers just about the *function*

4	(i)		B(90, 0.05) ≈ Po(4.5) P(≤ 6) from Po(4.5) = **0.8311** *n* large, *p* small, *or* $n > 50$, $np < 5$, therefore Poisson	M1 A1 M1 A1 B1 **5**	B(90, *any*) stated or implied", but *p* can be algebraic or omitted Poisson (4.5) stated or implied Allow M1A0 for 0.7029, 0.9134, 0.8436 or 0.8180 but nothing else Either pair asserted, $n = 90$, $p = 0.05$	Not just "$n = 90$ Exact (0.836055): M1A0M0A0B0, 1/5 Normal (4.5, 4.275): M1A0M0A0B0, 1/5 *Not* isw: final answer 0.1689 is M0A0 Allow "$np < 5$, n large"
	(ii)		B(90, 0.35) ≈ N(31.5, 20.475) $P(\leq 28) = \Phi\left(\frac{28.5-31.5}{\sqrt{20.475}}\right)$ $= \Phi(-0.6630)$ = **0.2537**	M1 A1 M1 A1 A1 **5**	Normal, attempt at 90×0.35 Both parameters correct, allow √ Standardise "29" using *np*, √*npq*, allow wrong/no cc, σ^2 or $\sqrt{\sigma}$ cc <u>and</u> √ correct Answer, in range [0.2535, 0.2545)	$s^2 = 819/40$ Variance *np* or *nq*, or extra √90 in SD: M0 so N(31.5, 58.5) is 1/5 NB: Exact: 0.2558, 0/5
5	(i)		Avoids (reduces) bias, or "representative" or "allows calculations to be done" or "allows reliable estimates"	B1 **1**	– unbiased (allow "fair") – representative (allow "reliable") – allows use of distribution Both right and wrong: B1	*Not:* – all equally likely to be selected – selections independent – quick/easy/cheap – random sample
	(ii)		B(18, 0.25) H_0: $p = 0.25$, H_1: $p > 0.25$ α: $P(\geq 8) = 1 - P(\leq 7)$ = **0.0569** > 0.05	M1 B2 A1 A1	B(18, 0.25) stated or used One error, B1; x or $\bar{x}$ B0; π: B2 0.0569 seen Explicit comparison with 0.05	*Any* symbol can get B2 if explicitly defined Allow 0.9431 only if "> 0.95" and vice versa. "> 8" (0.0193), "≤ 8" (0.9807) or "= 8" (.0376): max M1B2 [A0A0M0A0], 3/7
			β: CR is ≥ 9, and 8 < 9 probability 0.0193	A1dep* *A1	Correct CR and explicit comparison 0.0193 explicitly seen	
			If more than one probability seen, assume method is *β*. Note that this requires explicit comparison for either A1; but can get final M1A1			
			Do not reject H_0. Insufficient evidence that proportion of 1's is greater than 25%.	M1 A1ft **7**	Correct first conclusion, e.g. "reject H_1" Interpreted, in context, consistent with *p*, acknowledge uncertainty. FT on wrong CR/*p* *Not:* "significant evidence that proportion of 1s is 25%"	M1 needs correct method, comparison, like-with-like, ≥ 8 (or ≤ 7 but only if used consistently) Allow "change" instead of "increase" **SR**: 2-tail, max M1B1B0A1A0M1A1

6	(i)		$\bar{t} = 11.76$ $\hat{\sigma}^2 = \frac{120}{119}\left(\frac{18737.712}{120} - 11.76^2\right) = 18$ H_0: $\mu = 11.0$, H_1: $\mu \neq 11.0$ α: $z = \frac{11.76 - 11.0}{\sqrt{18/120}} =$ **1.9623** > 1.645	B1 M1 M1 A1 B2 M1 A1 A1	11.76 seen or implied Biased estimate (= 17.85) ×120/119, *or* single formula with 119 divisor Answer 18 ± 0.05 One error, B1, but $\bar{t}$, t, x etc: B0 (u: B1) Standardise with 120, ignore cc or √ errors A.r.t. (±)1.96 *or* $p \in$ [0.0245, 0.025] www Compare explicitly with (±)1.645 or 0.05, consistent with their z or p. *[Needs to be "next to" TS]*	i.e. correct single formula gets M2 If both hypotheses involve 11.76, only further mark possible is next M1 [max 5/11] 120 omitted gets no further marks [max 6/11] Ignore "N(11.76, …)" *unless* hypotheses omitted altogether, in which case treat as hypotheses in terms of 11.76
			β: CV 11.0 ± 1.645×√(18/120) = 11.637 (or 10.363) 11.76 > 11.64	M1 A1 A1	11.0 + $z\sigma/\sqrt{120}$, needs 120 and + or ± Ignore 10.363 Explicit comparison, consistent tail	If 11.76 – $z\sigma/\sqrt{120}$, give M1A0A0 M0A0 (even if correct hypotheses)
			Reject H_0. Significant evidence that the average time has changed.	M1 A1ft **11**	Correct first conclusion, allow "Accept H_1" Contextualised, acknowledge uncertainly, FT on wrong CR/z/p	Needs correct method (including 120) and comparison type, 11.0 in at least one hypothesis Allow "increase" instead of "change"
	(ii)		No, the Central Limit Theorem applies	B1 **1**	or "No, large sample". Withhold if extra wrong or irrelevant reason(s) given	Needs both "no" and reason

7	(i)	α	$H_0 : \lambda = 5$ (1 or ⅓), $H_1 : \lambda < 5$ (1 or ⅓) Po(5) P(≤ 2) = **0.1247** > 0.05	B2 M1 A1 A1	One error, B1, except *t*, *x* etc: 0. Allow μ Stated or implied	H_0: $\lambda_0 = 5$, H_1: $\lambda_1 < 5$ is one error Can be implied by N(5, 5) but no more marks < 2 (0.0404) or = 2 (0.0843) or ≥ 2 (0.9596): no further marks. > 2 (0.8753) and compare with 0.95, OK, but must be fully consistent for more than 3 marks
		β	CR is ≤ 1 and compare 2 explicitly $p = 0.0404$	A1* dep*A1		
			If more than one probability seen, assume method is β. Note that this requires explicit comparison for either A1; but can get final M1A1			
			Do not reject H_0. Insufficient evidence of reduction in mean number of points failures	M1 A1ft **7**	Correct first conclusion, needs Po(5) Contextualised, acknowledge uncertainty, FT on wrong CR/*p*/comparison value	M1 needs correct method and comparison
	(ii)		Po(50) ≈ N(50, 50) 50 – 0.5 – 1.645×√50 = 37.87 so maximum number is **37**	M1 M1 A1 M1 A1 A1 **6**	Stated or implied Normal, mean 150λ Variance/sd same $50 - z\times\sqrt{50}$, *not* $\div\sqrt{n}$, allow √ error, any cc $z = 1.645$, need √50, allow 50 + 0.5 or 50 37 only, from cc, *not* final answer 37.86, but allow corrected to 37 after check Exact Poisson (0.0473 ⇒ 38) is 1/6	Allow if variance = √50 etc. Ignore $50 + z\times\sqrt{50}$. √ errors here is M1A0 38(.36) from no cc is probably 5/6, but beware that exact Poisson also gives 38
8	(i)		1 – P(≤ 7) = 0.0315 or 3.15%	M1 A1 **2**	Clearly stated, or implied by any of 0.0083, 0.0933, 0.0103, 0.0576, 0.0744 Ignore subsequent "therefore 5%" etc	Other answers 0/2 unless "1 – P(≤ 7)" explicitly seen
	(ii)		P(Type II error | $p = 0.5$) = 0.6047 P(Type II error | $p = 0.7$) = 0.0933 ⅓ × 0.6047 + ⅓ × 0.0933 [= 0.2327] P(none of 4) = $1 - (1 - 0.2327)^4$ **= 0.653**	M1 A1 M1 M1 A1 **5**	Explicit "P(≤ 7) for $p = 0.5$" [or 0.7] or 0.2120, 0.7805: M1 Both correct ⅓ × one prob + ⅓ × other prob (*not* 0.9685) $1 - (1 - \text{ans})^4$ or equivalent binomial, not ⅓ Allow in range [0.653, 0.654]	Tail consistent. Any "1 –" errors: M0 Ignore 0.9685 Independent Independent *[not* $\frac{16}{81}$ or $\frac{65}{81}$] (all three M marks independent)