OCR
RECOGNISING ACHIEVEMENT

Tuesday 17 January 2012 – Morning

A2 GCE MATHEMATICS

4733 Probability & Statistics 2

QUESTION PAPER

Candidates answer on the printed Answer Book.

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

Other materials required:
- Scientific or graphical calculator

Duration: 1 hour 30 minutes

4732980112

INSTRUCTIONS TO CANDIDATES

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

- The Question Paper will be found in the centre of 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 A random sample of 50 observations of the random variable X is summarised by

$$n = 50, \Sigma x = 182.5, \Sigma x^2 = 739.625.$$

Calculate unbiased estimates of the expectation and variance of X. **[4]**

2 The random variable Y has the distribution B(140, 0.03). Use a suitable approximation to find $P(Y = 5)$. Justify your approximation. **[5]**

3 The random variable G has a normal distribution. It is known that

$$P(G < 56.2) = P(G > 63.8) = 0.1.$$

Find $P(G > 65)$. **[6]**

4 The discrete random variable H takes values 1, 2, 3 and 4. It is given that $E(H) = 2.5$ and $Var(H) = 1.25$. The mean of a random sample of 50 observations of H is denoted by $\bar{H}$.
Use a suitable approximation to find $P(\bar{H} < 2.6)$. **[5]**

5 **(i)** Six prizes are allocated, using random numbers, to a group of 12 girls and 8 boys. Calculate the probability that exactly 4 of the prizes are allocated to girls if

(a) the same child may win more than one prize, **[2]**

(b) no child may win more than one prize. **[2]**

(ii) Sixty prizes are allocated, using random numbers, to a group of 1200 girls and 800 boys. Use a suitable approximation to calculate the probability that at least 30 of the prizes are allocated to girls. Does it affect your calculation whether or not the same child may win more than one prize? Justify your answer. **[6]**

6 The number of fruit pips in 1 cubic centimetre of raspberry jam has the distribution $Po(\lambda)$. Under a traditional jam-making process it is known that $\lambda = 6.3$. A new process is introduced and a random sample of 1 cubic centimetre of jam produced by the new process is found to contain 2 pips. Test, at the 5% significance level, whether this is evidence that under the new process the average number of pips has been reduced. **[8]**

7 **(i)** The continuous random variable X has the probability density function

$$f(x) = \begin{cases} \dfrac{1}{2\sqrt{x}} & 1 \leqslant x \leqslant 4, \\ 0 & \text{otherwise.} \end{cases}$$

Find **(a)** $E(X)$, **[3]**

(b) the median of X. **[3]**

(ii) The continuous random variable Y has the probability density function

$$g(y) = \begin{cases} \dfrac{1.5}{y^{2.5}} & y \geqslant 1, \\ 0 & \text{otherwise.} \end{cases}$$

Given that $E(Y) = 3$, show that $\text{Var}(Y)$ is not finite. **[3]**

8 In a certain fluid, bacteria are distributed randomly and occur at a constant average rate of 2.5 in every 10 ml of the fluid.

(i) State a further condition needed for the number of bacteria in a fixed volume of the fluid to be well modelled by a Poisson distribution, explaining what your answer means. **[2]**

Assume now that a Poisson model is appropriate.

(ii) Find the probability that in 10 ml there are at least 5 bacteria. **[2]**

(iii) Find the probability that in 3.7 ml there are exactly 2 bacteria. **[3]**

(iv) Use a suitable approximation to find the probability that in 1000 ml there are fewer than 240 bacteria, justifying your approximation. **[7]**

9 It is desired to test whether the average amount of sleep obtained by school pupils in Year 11 is 8 hours, based on a random sample of size 64. The population standard deviation is 0.87 hours and the sample mean is denoted by $\bar{H}$. The critical values for the test are $\bar{H} = 7.72$ and $\bar{H} = 8.28$.

(i) State appropriate hypotheses for the test, explaining the meaning of any symbol you use. **[3]**

(ii) Calculate the significance level of the test. **[4]**

(iii) Explain what is meant by a Type I error in this context. **[1]**

(iv) Given that in fact the average amount of sleep obtained by all pupils in Year 11 is 7.9 hours, find the probability that the test results in a Type II error. **[3]**

Question			Answer	Marks	Guidance
1			$\hat{\mu} = \bar{x} = 3.65$	B1	3.65 stated explicitly, *not* isw
			$S^2 = \frac{739.625}{50} - 3.65^2$ [= 1.47]	M1	Correct formula for biased estimate used, award if 1.47 seen
			$\hat{\sigma}^2 = \frac{50}{49}S^2$	M1	$n/(n-1)$ factor used, or if wrong single formula, M1 if $n-1$ divisor anywhere. Correct single formula: M2
			= 1.5	A1	Answer 1.5 or exact equivalent only
				[4]	
2			Po(4.2)	M1	Po(np) stated or implied
			$e^{-4.2}\frac{4.2^5}{5!}$ **= 0.1633**	M1	Poisson formula or tables, allow for .1944, .1144, .16(0), .1663;
				A1	Answer, a.r.t. 0.163
			n large, p small	B1	One condition } Needs Poisson. If inequalities
			or $n > 50$, $np < 5$	B1	The other condition } used, must be these, but allow $p < 0.1$ if and nothing extraneous $n > 50$ already stated
				[5]	
3			$\mu = 60$	B1	$\mu = 60$ stated or implied, can be written down
			$\frac{63.8-\mu}{\sigma} = \Phi^{-1}(0.9) = 1.282$	M1	Standardise 63.8 or 56.2 with σ, allow √ or cc errors, equate to Φ^{-1}
				B1	1.282 (or 1.281 or 1.28) seen
			$\sigma = 2.96(4)$	A1	σ, in range [2.96, 2.97], can be implied by what follows, *not* σ^2
			$1-\Phi\left(\frac{65-60}{2.964}\right) = 1 - \Phi(1.687)$	M1	Standardise 65 with their μ and σ, allow √ or cc errors
			= 0.0458	A1	Final answer, a.r.t. 0.046, c.w.o.
				[6]	
4			N(2.5, 0.025)	M1	Normal (any – can be implied by standardisation)
			$\Phi\left(\frac{2.59-2.5}{\sqrt{0.025}}\right) = \Phi(0.5692)$	A1	Mean 2.5
				A1	Variance or SD 1.25 ÷ 50 stated or used
				A1	Standardise 2.59 or 2.61, with √(1.25/50)
			= **0.7154**	A1	Answer in range [0.715, 0.716] *or* [0.736, 0.737] from 0.632
				[5]	

5	(i)	(a)	${}^6C_4\ 0.6^4\ 0.4^2$ = **0.311**[04] [= 972/3125]	M1 A1 **[2]**	This formula, allow 0.6 ↔ 0.4, or tables used correctly Final answer, exact fraction or a.r.t. 0.311
5	(i)	(b)	${}^{12}C_4 \times {}^8C_2 \div {}^{20}C_6$ [= 495×28 ÷ 38760] = **0.3576** [= 231/646]	M1 A1 **[2]**	Product of two ${}^nC_r$ divided by ${}^nC_r$, *or* ${}^6C_2\left(\frac{12}{20}\times\frac{11}{19}\times\frac{10}{18}\times\frac{9}{17}\times\frac{8}{16}\times\frac{7}{15}\right)$ Final answer, exact fraction or a.r.t. 0.358
5	(ii)		B(60, 0.6) ≈ N(36, 14.4) $1-\Phi\left(\frac{29.5-36}{\sqrt{14.4}}\right) = 1 - \Phi(-1.713)$ = **0.9567** No effect as population is large *or* yes but not by much	B1 B1 M1 A1 A1 B1 **[6]**	N(36, …) or N(24, …); 14.4 or √14.4, both from B(60, 0.6) Standardise with their *np* and √*npq* (or *npq*) Both their √*npq* and cc correct [30.5 if using 24] Answer in range [0.956, 0.957] Need all of one of these [*not* "sample"], or equiv, nothing wrong
6			$H_0: \lambda = 6.3$ [*or* μ] $H_1: \lambda < 6.3$ [*or* μ] $P(\le 2) = e^{-6.3}(1 + 6.3 + 19.845)$ = **0.0498** < 0.05 Therefore reject H_0. Significant evidence that average number of pips has been reduced.	B2 M1 A1 A1 B1 M1 A1√ **[8]**	Both: B2. One error e.g. "H_0 = 6.3", or "H_1: $\lambda \ne 6.3$", B1, but *x*, *r* etc: 0 Correct formula for at least 2 terms, can be implied by 0.0134 Fully correct formula for ≤ 2, can be implied by answer Answer, a.r.t. 0.0498 SR tables: B2 if a.r.t. 0.0506, else 0 [then can get B1M1A1] Explicitly state < 0.05, *not* from H_1: $\lambda \ne 6.3$, *or* CR ≤ 2 and explicitly state 2 in CR, needs essentially correct distribution *Not needed for final M1A1* Correct method, comparison and first conclusion Interpreted in context acknowledging uncertainty somewhere, √ on *p* etc SR: P(< 2) [0.0134] or Po(= 2) [0.0364]: B2 M1 A0 B0 M0 but *allow* "Po(= 2) = 0.0498" etc SR: Normal: B2 M1 A0 B0

7	(i)	(a)	$\int_1^4 \frac{1}{2\sqrt{x}} x dx = \left[\frac{1}{3}x^{\frac{3}{2}}\right]_1^4 = 7/3$ or 2.333…	M1 B1 A1 **[3]**	Attempt to integrate xf(x), correct limits Correct indefinite integral, a.e.f. Final answer 7/3 or equiv or a.r.t. 2.33
7	(i)	(b)	$\int_1^m \frac{1}{2\sqrt{x}} dx = 0.5$ $\sqrt{m} - 1 = 0.5$ $\boldsymbol{m = 2.25}$	M1 A1 A1 **[3]**	This or complementary integral, limits needed [not "$-\infty$"], equated to 0.5, needn't attempt to evaluate This equation, any equivalent simplified form Answer 9/4 or exact equivalent only
7	(ii)		$1.5\int_1^\infty y^{-2.5} y^2 dy = 1.5\left[\frac{y^{0.5}}{0.5}\right]_1^\infty$ Upper limit gives infinite answer	M1 B1 A1 **[3]**	Attempt to integrate y^2f(y), limits 1 and ∞, allow any letter Correct indefinite integral [$=3\sqrt{y}$], ignore μ [= 3] Give correct reason, c.w.o. apart from constant, allow "$= \infty$"
8	(i)		Location of bacteria must be independent – the position of one does not affect that of another	M1 A1 **[2]**	"Found independently": M1. Allow "are independent", "singly". Context needed somewhere in answer. Correct explanation, not just of "singly", e.g. not "must not group together". No extra or wrong conditions, but allow both "singly" and "independently". Right explanation, not "independent": M1A0

Examples

α	Number of bacteria occurring in a particular volume is independent of the number in another interval of the same volume. Number in one volume occurs randomly.	M1A0
β	Bacteria are distributed independently from one another. This means that they cannot be in groups.	M1A0
γ	Position of each bacterium must be independent of the position of other bacteria. Not well modelled by Poisson if they tended to form groups, they must not be influenced by the surrounding bacteria or certain conditions (e,g, heat).	M1A0
δ	Bacteria need to be independent. The results of one cannot influence the result of another.	M1A0
ε	Bacteria must occur independently, so the state of one bacterium has no effect on any other bacteria.	M1A0
ς	Probability of bacteria must be independent, they cannot affect the probability of another bacterium occurring.	M1A1
η	Bacteria must occur independently, so if one occurs it can't cause more to appear.	M1A1

8	(ii)		$1 - P(\le 4)$ [= 1 – 0.8912] **= 0.1088**	M1 A1 **[2]**	Allow M1 for 1 – .9580 [= 0.042] or wrong λ. 0.8912 etc: M0 0.109 or 0.1088 or better
8	(iii)		Po(0.925) $e^{-0.925}\dfrac{0.925^2}{2!}$ **= 0.169**(64)	M1 M1 A1 **[3]**	Po(0.925) stated or implied [37/40] Correct Po formula for $r = 2$, any λ, can be implied by: Answer 0.17(0) or 0.1696 or better
8	(iv)		Po(250) $\lambda > 15$ or λ large [or μ] N(250, 250) $\Phi\left(\dfrac{239.5-250}{\sqrt{250}}\right)=1-\Phi(0.664)$ **= 0.2533**	B1 B1 M1* A1√ Dep*M1 A1√ A1 **[7]**	Po(250) stated or implied Either of these N, mean their 100 × 2.5 … … variance (or SD) their mean Standardise, allow wrong or no cc and/or no √ or σ^2, needs A1 Continuity correction and √ correct Final answer a.r.t. 0.253, c.w.o.
9	(i)		H_0: $\mu = 8$; H_1: $\mu \ne 8$ where μ is the population mean amount of sleep obtained by Year 11 pupils	B2 B1 **[3]**	Both, B2. One error, B1, allow *x*/*r*/*t* here, but not $\overline{H}$ Need "population" or equivalent, but allow "average amount of sleep obtained by Year 11 pupils". Allow "μ is population mean".
9	(ii)		$\Phi\left(\dfrac{0.28}{0.87/\sqrt{64}}\right) = \Phi(2.575)$ 2 × (1 – above) = **0.01** or **1%**	M1 A1 M1 A1 **[4]**	Standardise, with √*n* or *n*, allow cc, √ errors $z = 2.575$ or 2.57 or 2.58, can be implied by, e.g., 0.005 or 0.995 Correct handling of tails Answer 0.01 or 1% correct to 2 SF, c.w.o.
9	(iii)		Rejecting H_0 when $\mu = 8$	B1 **[1]**	Or equivalent, some mention of context, *not* "probability of …"
9	(iv)		$\Phi\left(\dfrac{8.28-7.9}{0.87/\sqrt{64}}\right)-\Phi\left(\dfrac{7.72-7.9}{0.87/\sqrt{64}}\right)$ $= \Phi(3.494) - \Phi(-1.655)$ [= 0.99976 – (1 – 0.951) or 1 – …] **= 0.951**	M1 M1 A1 **[3]**	Find P(between 7.72 and 8.28 | $\mu = 7.9$), allow 1 – 2×P(1 tail) (need attempt to find correct region, *not* isw – i.e., *not* ans 0.049) Correct handling of tails, needn't attempt to evaluate, needs 64 Final answer, a.r.t. 0.951. SR: One tail only used: M1M0A0. 0.951 from no working: B2