OCR
RECOGNISING ACHIEVEMENT

ADVANCED GCE

MATHEMATICS

Probability & Statistics 2

4733

Candidates answer on the Answer Booklet

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

Other Materials Required:
None

Friday 15 January 2010
Afternoon

Duration: 1 hour 30 minutes

4733

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 The values of 5 independent observations from a population can be summarised by

$$\Sigma x = 75.8, \quad \Sigma x^2 = 1154.58.$$

Find unbiased estimates of the population mean and variance. **[4]**

2 A college has 400 students. A journalist wants to carry out a survey about food preferences and she obtains a sample of 30 pupils from the college by the following method.

- Obtain a list of all the students.
- Number the students, with numbers running sequentially from 0 to 399.
- Select 30 random integers in the range 000 to 999 inclusive. If a random integer is in the range 0 to 399, then the student with that number is selected. If the number is greater than 399, then 400 is subtracted from the number (if necessary more than once) until an answer in the range 0 to 399 is selected, and the student with that number is selected.

(i) Explain why this method is unsatisfactory. **[2]**

(ii) Explain how it could be improved. **[1]**

3 In a large town, 35% of the inhabitants have access to television channel C. A random sample of 60 inhabitants is obtained. Use a suitable approximation to find the probability that 18 or fewer inhabitants in the sample have access to channel C. **[6]**

4 80 randomly chosen people are asked to estimate a time interval of 60 seconds without using a watch or clock. The mean of the 80 estimates is 58.9 seconds. Previous evidence shows that the population standard deviation of such estimates is 5.0 seconds. Test, at the 5% significance level, whether there is evidence that people tend to underestimate the time interval. **[7]**

5 The number of customers arriving at a store between 8.50 am and 9 am on Saturday mornings is a random variable which can be modelled by the distribution Po(11.0). Following a series of price cuts, on one particular Saturday morning 19 customers arrive between 8.50 am and 9 am. The store's management claims, first, that the mean number of customers has increased, and second, that this is due to the price cuts.

(i) Test the first part of the claim, at the 5% significance level. **[7]**

(ii) Comment on the second part of the claim. **[1]**

6 The continuous random variable X has the distribution $N(\mu, \sigma^2)$.

(i) Each of the three following sets of probabilities is impossible. Give a reason in each case why the probabilities cannot both be correct. (You should not attempt to find μ or σ.)

(a) $P(X > 50) = 0.7$ and $P(X < 50) = 0.2$ **[1]**

(b) $P(X > 50) = 0.7$ and $P(X > 70) = 0.8$ **[1]**

(c) $P(X > 50) = 0.3$ and $P(X < 70) = 0.3$ **[1]**

(ii) Given that $P(X > 50) = 0.7$ and $P(X < 70) = 0.7$, find the values of μ and σ. **[4]**

7 The continuous random variable T is equally likely to take any value from 5.0 to 11.0 inclusive.

(i) Sketch the graph of the probability density function of T. **[2]**

(ii) Write down the value of $\mathrm{E}(T)$ and find by integration the value of $\mathrm{Var}(T)$. **[5]**

(iii) A random sample of 48 observations of T is obtained. Find the approximate probability that the mean of the sample is greater than 8.3, and explain why the answer is an approximation. **[6]**

8 The random variable R has the distribution $\mathrm{B}(10, p)$. The null hypothesis H_0: $p = 0.7$ is to be tested against the alternative hypothesis H_1: $p < 0.7$, at a significance level of 5%.

(i) Find the critical region for the test and the probability of making a Type I error. **[3]**

(ii) Given that $p = 0.4$, find the probability that the test results in a Type II error. **[3]**

(iii) Given that p is equally likely to take the values 0.4 and 0.7, find the probability that the test results in a Type II error. **[2]**

9 Buttercups in a meadow are distributed independently of one another and at a constant average incidence of 3 buttercups per square metre.

(i) Find the probability that in 1 square metre there are more than 7 buttercups. **[2]**

(ii) Find the probability that in 4 square metres there are either 13 or 14 buttercups. **[3]**

(iii) Use a suitable approximation to find the probability that there are no more than 69 buttercups in 20 square metres. **[5]**

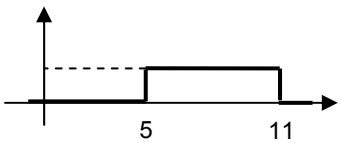
(iv) (a) Without using an approximation, find an expression for the probability that in m square metres there are at least 2 buttercups. **[2]**

(b) It is given that the probability that there are at least 2 buttercups in m square metres is 0.9. Using your answer to part **(a)**, show numerically that m lies between 1.29 and 1.3. **[4]**

4733 Probability & Statistics 2

Penalise over-specified answers (> 6 SF) first time but only once per paper.
Use Ⓐ or Ⓒ to annotate "over-assertive" or "no context" respectively

1		$\hat{\mu} = \bar{x} = 15.16$ $\hat{\sigma}^2 = \frac{5}{4}s^2$ $= \mathbf{1.363}$	B1 M1 M1 A1	**4**	15.16 or 15.2 as answer only Use $\frac{\Sigma x^2}{5} - \bar{x}^2$ [=1.0904] Multiply by 5/4, or equiv for single formula Final answer 1.36 or 1.363 only, *not* isw
2	(i)	Not all equally likely – those in range 0 to 199 more likely to be chosen	M1 A1	**2**	Not all equally likely stated or implied Justified by reference to numbers, no spurious reasons
	(ii)	Ignore random numbers greater than 799, or 399	B1	**1**	Any valid resolution of this problem, no spurious reasons
3		B(60, 0.35) ≈ N(21, 13.65) $\Phi\left(\frac{18.5-21}{\sqrt{13.65}}\right) = \Phi(-0.6767)$ $= 1 - 0.7507$ $= \mathbf{0.2493}$	M1 M1 A1 M1 A1 A1	**6**	B(60, 0.35) stated or implied N(21, ...) Variance or SD = 13.65 Standardise, their *np* and √*npq* or *npq*, wrong or no cc Both √*npq* and cc correct Answer, a.r.t. 0.249
4		$H_0 : \mu = 60$; $H_1 : \mu < 60$ (α) $z = \frac{58.9-60}{\sqrt{5^2/80}} = -1.967$ < -1.645	B2 M1 A1 B1		Both correct, B2 B1 for one error, but not *x*, *t*, $\bar{x}$ or $\bar{t}$ Standardise 58.9 & √80, allow – or √ errors *z*, art –1.97 or *p* in range [0.024, 0.025] Explicit comparison with –1.645 or 0.05, or +1.645 or 0.95 if 1.967 or 0.976 used
	or:	(β) $c = 60 - 1.645 \times \frac{5}{\sqrt{80}} = 59.08$ 58.9 < 59.08	M1 B1 A1✓		60 – *z*×5/√80, any $z = \Phi^{-1}$, allow √ errors or ±, not just +; *z* = 1.645 and compare 58.9 59.1 or better, ✓on wrong *z*
		Reject H_0 Significant evidence that people underestimate time	M1 A1✓	**7**	Correct first conclusion, needs essentially correct method including √80 or 80 Contextualised, uncertainty acknowledged SR: μ = 58.9: B0M1A0B1 max 2/7 SR: 2-tail: max 5/7
5	(i)	$H_0 : \lambda = 11.0$ $H_1 : \lambda > 11.0$ (α) P(≥ 19) = 1 – 0.9823 = 0.0177 < 0.05	B2 M1 A1 B1		Allow μ. Both correct, B2 One error: B1, but not *C*, *x* etc Find P(≥ 19) [or P(< 19) if later 0.95] art 0.0177 [0.9823, ditto] Compare 0.05 [0.95 if consistent], needs M1
		(β) CR ≥ 18, P(≥ 18) = 0.0322 19 > 18	M1 A1 B1		CR or CV 16/17/18/19 stated or clearly implied, but not < 18 and 0.0322 both seen, allow 0.9678 Explicit comparison with 19, needs M1
		Reject H_0 Significant evidence of an increase in number of customers	M1 A1✓	**7**	Needs essentially correct method & comparison Contextualised, uncertainty acknowledged SR: Normal, or P(= 19) or P(≤ 19) or P(> 19): First B2 only.
	(ii)	Can't deduce cause-and-effect, or there may be other factors	B1	**1**	Conclusion needed. No spurious reasons. If "DNR" in **(i)**, "couldn't deduce even if…"

6	(i)	(a) Probabilities don't total 1	B1	**1**	Equivalent statement
		(b) P(> 70) must be < P(> 50)	B1	**1**	Equivalent statement
		(c) P(> 50) = 0.3 $\Rightarrow \mu < 50$ P(< 70) = 0.3 $\Rightarrow \mu > 70$	B1	**1**	Any relevant valid statement, e.g. "P(< 50) = 0.7 but P(< 50) must be < P(< 70)"
	(ii)	$\mu = 60$ by symmetry	B1		$\mu = 60$ obtained at any point, allow from Φ
		$\frac{10}{\sigma} = \Phi^{-1}(0.7) = 0.524(4)$	M1		One standardisation, equate to Φ^{-1}, not 0.758
		$\sigma = 10/0.5243$	B1		$\Phi^{-1} \in [0.524, 0.5245]$ seen
		= **19.084**	A1	**4**	σ in range [19.07, 19.1], e.g. 19.073
7	(i)		M1 A1	**2**	Horizontal line Evidence of truncation *[no need for labels]*
	(ii)	$\mu = 8$	B1		8 only, cwd
		$\int_5^{11} \frac{1}{6}t^2\,dt = \left[\frac{1}{18}t^3\right]_5^{11}$ [= 67]	M1		Attempt $\int kt^2\,dt$, limits 5 and 11 seen
			B1		$k = 1/6$ stated or implied
		$- 8^2$	M1		Subtract their (non-zero) mean^2
		= **3**	A1	**5**	Answer 3 only, *not* from MF1
	(iii)	N(8, 3/48)	M1		Normal stated or implied
		$1 - \Phi\left(\frac{8.3 - 8}{\sqrt{3/48}}\right) = 1 - \Phi(1.2)$	A1		Mean 8
			A1		Variance their (non-zero) **(ii)**/48
		= 1 – 0.8848	M1		Standardise, $\sqrt{n}$, ignore sign or $\sqrt{}$ errors. cc: M0
		= **0.1151**	A1		Answer, art 0.115
		Normal distribution only approx.	B1	**6**	Any equivalent comment, e.g. CLT used
8	(i)	P(≤ 4) = 0.0473	M1		P(≤ *r*) from B(10, 0.7), *r* = 3/4/5, *not* N
		Therefore CR is ≤ 4	B1		"≤ 4" stated, not just "4", nothing else
		P(Type I error) = **4.73%**	A1	**3**	Answer, art 0.0473 or 4.73%, must be stated
	(ii)	B(10, 0.4) and find P(> 4)	M1		Must be this, *not* isw, ✓on **(i)**
		1 – P(≤ 4)	M1		Allow for 0.6177 or 0.1622
		= **0.3669**	A1	**3**	Answer, art 0.367
	(iii)	0.5 × 0.3669	M1		0.5 × **(ii)**
		= **0.18345**	A1✓	**2**	Ans correct to 3 SF, e.g. 0.184 from 0.367

9	(i)	$1 - P(\leq 7) = 1 - 0.9881$ $= \mathbf{0.0119}$	M1 A1	**2**	Allow for 0.0038 or 0.0335 Answer, a.r.t. 0.0119
	(ii)	Po(12) $P(\leq 14) - P(\leq 12)$ [0.7720 – 0.5760] $= \mathbf{0.196}$	M1 M1 A1	**3**	Po(12) stated or implied Formula, 2 consecutive correct terms, or tables, e.g. .0905 or .3104 or .1629 Answer, art 0.196
	(iii)	Po(60) ≈ N(60, 60) $\Phi\left(\frac{69.5-60}{\sqrt{60}}\right) = \Phi(1.226)$ $= \mathbf{0.8899}$	M1 A1 M1 A1 A1	**5**	N(60, …) Variance or SD 60 Standardise, λ & $\sqrt{\lambda}$, allow λ or wrong or no cc $\sqrt{\lambda}$ and cc both correct Answer 0.89 or a.r.t. 0.890
	(iv)	(a) $1 - e^{-3m}(1 + 3m)$	M1 A1	**2**	M1 for one error, e.g. no "1 –", or extra term, or 0^{th} term missing; answer, aesf
		(b) $m = 1.29$, $p = 0.89842$ $m = 1.3$, $p = 0.9008$ Straddles 0.9, therefore solution between 1.29 and 1.3	M1 A1 A1 A1	**4**	Substitute 1.29 or 1.3 into appropriate fn Comp | 0.9 | 0.1 | 0 1.29 | 0.898 | 0.10158 | –.00158 1.3 | 0.901 | 0.09918 | .0008146 Explicit comparison with relevant value, & conclusion, needs both *p*s correct
	or	Method for iteration; 1.296… 1.2965or better; conclusion stated	M1A1 A1A1		Can be implied by at least 1.296… Need at least 4 dp for M1A2