OCR
RECOGNISING ACHIEVEMENT

OXFORD CAMBRIDGE AND RSA EXAMINATIONS

Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education

MATHEMATICS 4733

Probability & Statistics 2

Wednesday **22 JUNE 2005** Afternoon 1 hour 30 minutes

Additional materials:
Answer booklet
Graph paper
List of Formulae (MF1)

TIME 1 hour 30 minutes

INSTRUCTIONS TO CANDIDATES

- Write your name, centre number and candidate number in the spaces provided on the answer booklet.
- Answer **all** the questions.
- 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.
- The total number of marks for this paper is 72.
- Questions carrying smaller numbers of marks are printed earlier in the paper, and questions carrying larger numbers of marks later in the paper.
- **You are reminded of the need for clear presentation in your answers.**

This question paper consists of 4 printed pages.

1 It is desired to obtain a random sample of 15 pupils from a large school. One pupil suggests listing all the pupils in the school in alphabetical order and choosing the first 15 names on the list.

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

(ii) Suggest a better method. [2]

2 A continuous random variable has a normal distribution with mean 25.0 and standard deviation σ. The probability that any one observation of the random variable is greater than 20.0 is 0.75. Find the value of σ. [4]

3 **(a)** The random variable X has a $B(60, 0.02)$ distribution. Use an appropriate approximation to find $P(X \leqslant 2)$. [3]

(b) The random variable Y has a $Po(30)$ distribution. Use an appropriate approximation to find $P(Y \leqslant 38)$. [5]

4 The height of sweet pea plants grown in a nursery is a random variable. A random sample of 50 plants is measured and is found to have a mean height 1.72 m and variance $0.0967\ m^2$.

(i) Calculate an unbiased estimate for the population variance of the heights of sweet pea plants. [2]

(ii) Hence test, at the 10% significance level, whether the mean height of sweet pea plants grown by the nursery is 1.8 m, stating your hypotheses clearly. [7]

5 The random variable W has the distribution $B(30, p)$.

(i) Use the exact binomial distribution to calculate $P(W = 10)$ when $p = 0.4$. [2]

(ii) Find the range of values of p for which you would expect that a normal distribution could be used as an approximation to the distribution of W. [3]

(iii) Use a normal approximation to calculate $P(W = 10)$ when $p = 0.4$. [6]

6 A factory makes chocolates of different types. The proportion of milk chocolates made on any day is denoted by p. It is desired to test the null hypothesis $H_0 : p = 0.8$ against the alternative hypothesis $H_1 : p < 0.8$. The test consists of choosing a random sample of 25 chocolates. H_0 is rejected if the number of milk chocolates is k or fewer. The test is carried out at a significance level as close to 5% as possible.

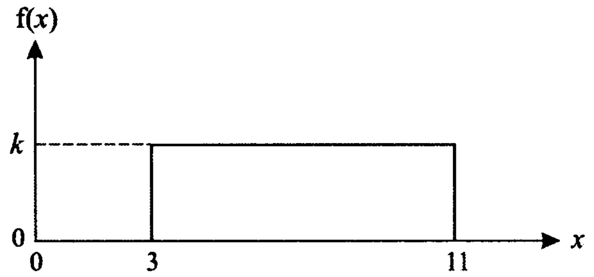
(i) Use tables to find the value of k, giving the values of any relevant probabilities. [3]

(ii) The test is carried out 20 times, and each time the value of p is 0.8. Each of the tests is independent of all the others. State the expected number of times that the test will result in rejection of the null hypothesis. [2]

(iii) The test is carried out once. If in fact the value of p is 0.6, find the probability of rejecting H_0. [2]

(iv) The test is carried out twice. Each time the value of p is equally likely to be 0.8 or 0.6. Find the probability that exactly one of the two tests results in rejection of the null hypothesis. [4]

7 The continuous random variable X has the probability density function shown in the diagram.



(i) Find the value of the constant k. [2]

(ii) Write down the mean of X, and use integration to find the variance of X. [5]

(iii) Three observations of X are made. Find the probability that $X < 9$ for all three observations. [3]

(iv) The mean of 32 observations of X is denoted by $\overline{X}$. State the approximate distribution of $\overline{X}$, giving its mean and variance. [3]

[Question 8 is printed overleaf.]

[Turn over

8 In excavating an archaeological site, Roman coins are found scattered throughout the site.

(i) State two assumptions needed to model the number of coins found per square metre of the site by a Poisson distribution. [2]

Assume now that the number of coins found per square metre of the site can be modelled by a Poisson distribution with mean λ.

(ii) Given that $\lambda = 0.75$, calculate the probability that exactly 3 coins are found in a region of the site of area $7.20\,\text{m}^2$. [3]

A test is carried out, at the 5% significance level, of the null hypothesis $\lambda = 0.75$, against the alternative hypothesis $\lambda > 0.75$, in Region LVI which has area $4\,\text{m}^2$.

(iii) Determine the smallest number of coins that, if found in Region LVI, would lead to rejection of the null hypothesis, stating also the values of any relevant probabilities. [4]

(iv) Given that, in fact, $\lambda = 1.2$ in Region LVI, find the probability that the test results in a Type II error. [3]

1	(i)	Method is biased because many pupils cannot be chosen	B1 B1 2	"Biased" or equivalent stated, allow "not random" Valid relevant reason
	(ii)	Allocate a number to each pupil Select using random numbers	B1 B1 2	State "list numbered" Use random numbers [*not* "hat"]
2		$\frac{20-25}{\sigma} = \Phi^{-1}(0.25) = -0.674$ $\sigma = 5 \div 0.674$ $= 7.42$	M1 B1 M1 A1 4	Standardise and equate to Φ^{-1} [*not* .7754 or .5987] cc z in range [−0.675, −0.674], allow + (±) 5 ÷ z-value [*not* $\Phi(z)$ or 0.75] Answer in range [7.41, 7.42], no sign fudges [SR: σ^2: M1B1M0A0 cc: M1B1M1A0]
3	(a)	Po(1.2) Tables or correct formula used 0.8795	B1 M1 A1 3	Po(1.2) stated or implied Correct method for Poisson probability, allow "1 –" Answer, 0.8795 or 0.879 or 0.88(0)
	(b)	N(30, 30) $\frac{38.5-30}{\sqrt{30}}$ [= 1.55] [$\Phi(1.55)$ =] 0.9396	B1 B1 M1 A1 A1 5	Normal, mean 30 stated or implied Variance 30 stated or implied, allow $\sqrt{30}$ or 30^2 Standardise using $\sigma^2 = \mu$, allow $\sqrt{}$ or cc errors $\sqrt{\mu}$ and 38.5 both correct Answer in range [0.939, 0.94(0)]
4	(i)	$\hat{\sigma}^2 = \frac{50}{49} \times 0.0967 = 0.0987$	M1 A1 2	Use $\frac{n}{n-1} \times s$ or s^2, allow $\sqrt{}$ Answer, a.r.t. 0.0987
	(ii)	H_0: $\mu = 1.8$, H_1: $\mu \neq 1.8$ where μ is the population mean	B1B1	Hypotheses correctly stated in terms of μ 1-tailed? cc SR: μ wrong/omitted: B1 both, but $\overline{X}$: B0
	α, β:	$z = \frac{(1.72-1.8)}{\hat{\sigma}/\sqrt{50}} = -1.8(006)$	M1 A1	Standardise with $\sqrt{n}$, allow +, biased σ, $\sqrt{}$ errors $z = -1.80 \pm 0.01$, don't allow +
	α:	$-1.8 < -1.645$	B1√	Compare $\pm z$ with ± 1.645, signs consistent
	β:	$\Phi(-1.8) = 1 - 0.9641 < 0.05$	B1	Explicitly compare $\Phi(z)$ with 0.05, correct tail
	γ:	CV $1.8 - k.\sigma/\sqrt{50}$ $k = 1.645$, CV = 1.727 $1.72 < 1.727$	M1 A1√ B1√	Correct expression for CV, – or ±, k from Φ^{-1} CV = 1.727, √ on their k, ignore upper limit $k = 1.645$ and compare CV with 1.72
		Reject H_0 Significant evidence that mean height is not 1.8	M1 A1√ 7	Reject H_0√, correct method, needs $\sqrt{50}$, $\mu = 1.8$; allow cc,√ or k error or biased σ estimate Conclusion stated in context [SR: 1.8, 1.72 interchanged: B0B0M1A0B1M0]
5	(i)	${}^{30}C_{10}(0.4)^{10}(0.6)^{20}$ or $0.2915 - 0.1763$ $= 0.1152$	M1 A1 2	Correct formula or use of tables Answer, a.r.t. 0.115
	(ii)	$30p > 5$ so $p > 1/6$ $30q > 5$ so $q > 1/6$ $1/6 < p < 5/6$	M1 M1 A1 3	$30p$ used, or both values from 30pq 30p or 30q $30q$ or $30pq$ used 30p and 30pq, or both from 30pq Either $1/6 < p < 5/6$ *or* $0.211 < p < 0.789$, allow ≤
	(iii)	N(12, 7.2) $\frac{10.5-np}{\sqrt{npq}}$ and $\frac{9.5-np}{\sqrt{npq}}$ $\Phi(-0.559) - \Phi(-0.9317)$ $= 0.8243 - 0.7119 = 0.1124$	B1 B1 M1 A1√ M1 A1 6	12 seen 7.2 or 2.683 seen, allow 7.2^2 Both standardised, allow wrong/no cc, npq $\sqrt{npq}$, 10.5 and 9.5 correct, √ on their np, npq Correct use of tails Answer, in range [0.112, 0.113] [SR: $\frac{1}{\sqrt{2\pi \times 7.2}} e^{-\frac{1}{2}\frac{(10-12)^2}{7.2}}$ M1A1, answer A2]

6	(i)	$R \sim B(25, 0.8)$ $P(R \le 16) = 0.0468$, $P(R \le 17) = 0.1091$ $k = 16$	B1 M1 A1	3	B(25, 0.8) stated or implied, e.g. from N(20, 4) One relevant probability seen [Normal: M0] Answer $k = 16$ only
	(ii)	$20p$ $= 0.936$	M1 A1	2	$20 \times$ their p *or* 20×0.05 Answer, a.r.t. 0.936, i.s.w.
	(iii)	$P(R \le 16 \mid p = 0.6)$ $= 0.7265$	M1 A1	2	Find $P(R \le k \mid p = 0.6)$ Answer 0.7265 or 0.727
	(iv) α:	$p' = 0.5 \times 0.0468 + 0.5 \times 0.7265$ $= 0.38665$ $2 \times p' \times (1 - p')$ $= 0.474$	M1 A1 M1 A1	4	"Tree diagram" probability Value in range [0.38, 0.39] Correct formula, including 2, any p' Answer in range [0.47, 0.48]
	or β:	0.8 A 0.8 R $.5^2 \times .9532 \times .0468 = .0112$ 0.8 R 0.8 A $.5^2 \times .0468 \times .9532 = .0112$ 0.6 A 0.8 R $.5^2 \times .2735 \times .0468 = .0032$ 0.6 R 0.8 A $.5^2 \times .7265 \times .9532 = .1731$ 0.8 A 0.6 R $.5^2 \times .9532 \times .7265 = .1731$ 0.8 R 0.6 A $.5^2 \times .0468 \times .2735 = .0032$ 0.6 A 0.6 R $.5^2 \times .2735 \times .7265 = .0497$ 0.6 R 0.6 A $.5^2 \times .7265 \times .2735 = .0497$	M1 A1 M1 A1		$p_1q_2 + p_2q_1$ etc (0.5 not needed) 4 cases, √ on their ps and qs, 0.5 not needed e.g. $2(p_1q_2 + p_2q_1)$ Completely correct list of cases and probabilities, including 0.5 Answer in range [0.47, 0.48]
7	(i)	$(11 - 3)k = 1$ $k = 1/8$	M1 A1	2	Use area = 1 [e.g. $\int kx\,dx = 1$ with limits 3, 11] Answer 1/8 or 0.125 only
	(ii)	$\mu = \frac{1}{2}(3 + 11) = 7$ $\int_3^{11} \frac{1}{8}x^2 dx = \left[\frac{x^3}{24}\right]_3^{11}$ $[= 54\frac{1}{3}]$ $\sigma^2 = 54\frac{1}{3} - 7^2$ $= 5\frac{1}{3}$	B1 M1 A1 M1 A1	5	Mean 7, cwd Attempt $\int x^2 f(x)dx$, correct limits Indefinite integral $\frac{x^3}{3k}$, their k Subtract their μ^2 Correct answer, $5\frac{1}{3}$ or a.r.t. 5.33
	(iii)	$P(X < 9) = 6k$ $[= \frac{3}{4}]$ $(\frac{3}{4})^3$ $= 27/64$ or 0.421875	B1√ M1 A1	3	Correct p for their k Work out their p^3, $0 < p < 1$ Answer 27/64 or a.r.t. 0.422
	(iv)	Normal Mean is 7 Variance is $5\frac{1}{3} \div 32$ $(= \frac{1}{6})$	B1 B1√ B1√	3	"Normal" distribution stated Mean same as in (ii) √ Variance is [(iii) ÷ 32] √ [*not* √ errors]
8	(i)	Coins occur at constant average rate and independently of one another	B1 B1	2	One contextualised condition, e.g. independent A different one, e.g. constant average rate, or "not in hoards" ["singly" not enough]. Allow "They…"
	(ii)	$R \sim Po(5.4)$ $e^{-5.4}\frac{5.4^3}{3!} = 0.1185$	B1 M1 A1	3	Poisson (5.4) stated or implied Correct formula, any λ Answer, in range [0.118, 0.119]
	(iii)	$R \sim Po(3)$ Tables, looking for 0.05 or 0.95 $P(R \ge 7) = 0.0335$ Therefore smallest number is 7	B1 M1 A1√ A1	4	Poisson (3) stated or implied Evidence of correct use of tables One relevant correct probability seen $r = 7$ only, ignore inequalities
	(iv)	$R \sim Po(4.8)$ Type II error is $R < 7$ when $\mu = 4.8$ $P(< 7) = 0.7908$	B1 M1 A1	3	Poisson (4.8) used Correct context for Type II error, √ on their r $P(< 7)$, a.r.t. 0.791, c.w.o. [$P(\ge 7)$: M0]