OXFORD CAMBRIDGE AND RSA EXAMINATIONS

Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education

MATHEMATICS 2642

Probability & Statistics 2

Friday **17 JANUARY 2003** Afternoon 1 hour 20 minutes

Additional materials:
Answer booklet
Graph paper
List of Formulae (MF8)

TIME 1 hour 20 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 graphic 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 60.
- 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.

Registered Charity Number: 1066969
[Turn over

1 (i) Explain briefly how you would use random numbers to obtain a random sample of size 20 from those eligible to vote in a parliamentary constituency. [2]

(ii) In the constituency it is known that 30% of those eligible to vote are in social classes A or B.

(a) Find the probability that a random sample of size 20 contains at least 9 people in social classes A or B. [2]

Suppose that in your random sample there are 9 people in social classes A or B.

(b) What answer would you give to a commentator who claimed that your sampling method was biased? (You are not expected to carry out a significance test.) [1]

2 My expenditure at a supermarket each week is modelled by a normal distribution with mean £65.00 and standard deviation £σ. My expenditure exceeds £80.00 in a week with probability $\frac{1}{12}$.

(i) Calculate the value of σ, giving your answer correct to 4 significant figures. [4]

(ii) Calculate the probability that, in one randomly chosen week, my expenditure is less than £55.00. [2]

3 (i) State conditions under which the normal distribution can be used as an approximation to the distribution B(n, p). [2]

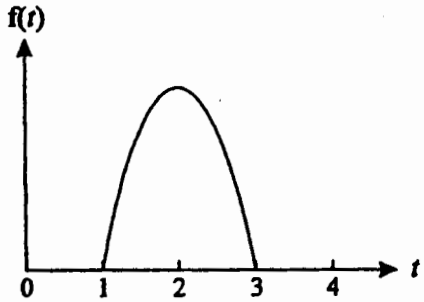
A series of digital messages, each consisting of 8000 bits, is transmitted. For each bit, the probability that it is wrongly received is 0.005. A message is unreadable if more than 50 bits are wrongly received.

(ii) Using a suitable approximation, find the probability that a message is unreadable. [6]

4 A student models the time, T hours, required for a certain journey by a continuous random variable with probability density function given by

$$f(t) = \begin{cases} k(t-1)(3-t) & 1 \leqslant t \leqslant 3, \\ 0 & \text{otherwise,} \end{cases}$$

where k is a constant. A sketch of this density function is shown in the diagram.



(i) Show that $k = \frac{3}{4}$. [2]

(ii) Write down the value of $E(T)$, and find the variance of T. [5]

(iii) Suggest one feature of the model which may not be realistic, and sketch the probability density function of a more realistic model. [2]

5 A random variable X has mean μ. It is desired to test the null hypothesis $H_0 : \mu = 48$ as opposed to the alternative hypothesis $H_1 : \mu \neq 48$. A random sample of size 36 can be summarised by $\Sigma x = 1665$, $\Sigma x^2 = 77\,890$.

(i) Carry out the test, using a 2% significance level. [8]

(ii) State with a reason whether, in carrying out this test, it is necessary to assume that X has a normal distribution. [1]

6 The number of misprints in a randomly chosen page of the Barchester *Gazette* is denoted by the random variable M. The mean value of M is 2.

(i) State two conditions needed in order to model M by a Poisson distribution. [2]

(ii) Find the probability that, in five randomly chosen pages of the *Gazette*, in the edition of 31 January 2003, there will be a total of more than 12 misprints. [3]

(iii) Find the smallest number of complete pages such that the probability that they contain a total of more than 12 misprints is greater than 0.7. You should show the values of any relevant probabilities obtained from tables. [4]

[Question 7 is printed overleaf.]

1 The time, T minutes, ... by a normal distribution with mean μ and standard deviation 7.0. It is known that, for experienced employees, the value of μ is 50.0. The times taken for a random sample of n newly qualified employees to complete the task are found. These times will be used in a test, at the 5% significance level, of whether newly qualified employees take longer than experienced employees.

(i) State appropriate hypotheses for the test. [1]

(ii) Given that $n = 40$ and that the mean time taken by the sample of newly qualified employees is 52.0 minutes, carry out the test, stating your conclusions clearly. [5]

(iii) The critical region for the test is $\{\bar{T} > c\}$. The test is to be modified so that the probability that it results in a Type I error is 0.05.

(a) Show that c and n would have to satisfy the equation

$$c - 50.0 = \frac{11.515}{\sqrt{n}},$$

approximately. [3]

It is also required that the probability that the test results in a Type II error when $\mu = 52.0$ is 0.05.

(b) Find a second (approximate) equation that would have to be satisfied by c and n. [2]

(c) Find the value of c from the equations in parts (a) and (b). [1]

(d) Hence find a suitable value for n. [2]

---

1 Explain how to obtain a random sample .....

[2]

No. in class A/B $X \sim \mathrm{B}\left(20,\, 0.3\right)$

$$\mathrm{p}(X \geq 9) = 1 - \mathrm{p}(X \leq 8) = 1 - 0.8867 = \mathbf{0.113}$$

[2]

sampling method is unbiased but produces an unrepresentative sample with quite a high probability.

[1]

●●

2●

Weekly expenditure $E \sim \mathrm{N}\left(65\right.$

$\mathrm{p}(E > 80)$

$\mathrm{p}(E < 55)$

●

[4]

[2]

---

3 B(n, p) can be approximated by a normal distribution when both np, $nq > 5$

[2]

Number of transmission errors $X \sim \mathrm{B}\left(8000,\ 0.005\right) \approx \mathrm{N}\left(40,\ 39.8\right)$

$$\mathrm{p}(\text{unreadable message}) = \mathrm{p}(X > 50) = 1 - \Phi\left(\frac{50.5 - 40}{\sqrt{39.8}}\right) = 1 - \Phi(1.664...) = \mathbf{0.0480}$$

[6]

---

4 $1 = k\int_1^3 \left(4t - 3 - t^2\right)\mathrm{d}t = k\left[2t^2 - 3t - \frac{1}{3}t^3\right]_1^3 = k\left(0 + \frac{4}{3}\right)$ $\qquad \boldsymbol{k = \frac{3}{4}}$ (show)

[2]

$\mathbf{E}[\boldsymbol{T}\,] = \mathbf{2}$ (by symmetry)

$$\mathrm{Var}[X] = k\int_1^3 t^2\,(t-1)(3-t)\,\mathrm{d}t \quad -2^2 = \frac{3}{4}\left[t^4 - t^{\,3} - \frac{1}{5}t^5\right]_1^3 - 4 = 4\tfrac{1}{5} - 4 = \tfrac{1}{5}$$

[5]

---

5 $\mathrm{H}_0:\ \mu = 48$ $\qquad$ $\mathrm{H}_1:\ \mu \neq 48$

$$\text{On } H_0 \qquad Z = \frac{\overline{X} - 48}{\sqrt{S^2/36}} \sim \mathrm{N}\left(0,\ 1\right)$$

with a significance level of 2%, reject H_0 if $|Z| > 2.326$

for the given sample …

$\overline{x} = \frac{1665}{36} = 46.25$ $\qquad$ $s^2 = \frac{36}{35}\left(\frac{77890}{36} - 46.25^2\right) = \frac{36}{35}(24.5486...) = 25.25$

$$z = \frac{46.25 - 48}{\sqrt{25.25/36}} = -2.090$$

no evidence on which to reject H_0 : can accept that $\mu = 48$.

[8]

accept either position:

"assumption of normality not needed due to Central Limit Theorem"

"n not very large so the assumption of normal X is needed"

[1]

---

6 two conditions for Poisson: independence, constant rate, ….

[2]

Po(10) $\qquad \mathrm{p}(>12) = 1 - \mathrm{p}(\leq 12) = 1 - 0.7916 = \mathbf{0.208}$

[3]

For n pages, no. of misprints is Po($2n$)

7

[10]

8

[3]