Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Advanced

Tuesday 15 January 2008 – Morning

Time: 1 hour 30 minutes

Materials required for examination	**Items included with question papers**
Mathematical Formulae (Green)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulae stored in them.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
The marks for individual questions and the parts of questions are shown in round brackets: e.g. **(2)**.
There are 8 questions on this paper. The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner. Answers without working may gain no credit.

N29278A

1. (*a*) Explain what you understand by a census.

(1)

Each cooker produced at GT Engineering is stamped with a unique serial number. GT Engineering produces cookers in batches of 2000. Before selling them, they test a random sample of 5 to see what electric current overload they will take before breaking down.

(*b*) Give one reason, other than to save time and cost, why a sample is taken rather than a census.

(1)

(*c*) Suggest a suitable sampling frame from which to obtain this sample.

(1)

(*d*) Identify the sampling units.

(1)

---

2. The probability of a bolt being faulty is 0.3. Find the probability that in a random sample of 20 bolts there are

(*a*) exactly 2 faulty bolts,

(2)

(*b*) more than 3 faulty bolts.

(2)

These bolts are sold in bags of 20. John buys 10 bags.

(*c*) Find the probability that exactly 6 of these bags contain more than 3 faulty bolts.

(3)

---

3. (*a*) State two conditions under which a Poisson distribution is a suitable model to use in statistical work.

(2)

The number of cars passing an observation point in a 10 minute interval is modelled by a Poisson distribution with mean 1.

(*b*) Find the probability that in a randomly chosen 60 minute period there will be

(i) exactly 4 cars passing the observation point,

(ii) at least 5 cars passing the observation point.

(5)

The number of other vehicles, other than cars, passing the observation point in a 60 minute interval is modelled by a Poisson distribution with mean 12.

(*c*) Find the probability that exactly 1 vehicle, of any type, passes the observation point in a 10 minute period.

(4)

---

4. The continuous random variable Y has cumulative distribution function $\mathrm{F}(y)$ given by

$$\mathrm{F}(y) = \begin{cases} 0 & y < 1 \\ k(y^4 + y^2 - 2) & 1 \le y \le 2 \\ 1 & y > 2 \end{cases}$$

(*a*) Show that $k = \frac{1}{18}$.

(2)

(*b*) Find $\mathrm{P}(Y > 1.5)$.

(2)

(*c*) Specify fully the probability density function $\mathrm{f}(y)$.

(3)

---

5. Dhriti grows tomatoes. Over a period of time, she has found that there is a probability 0.3 of a ripe tomato having a diameter greater than 4 cm. She decides to try a new fertiliser. In a random sample of 40 ripe tomatoes, 18 have a diameter greater than 4 cm. Dhriti claims that the new fertiliser has increased the probability of a ripe tomato being greater than 4 cm in diameter.

Test Dhriti's claim at the 5% level of significance. State your hypotheses clearly.

(7)

---

6. The probability that a sunflower plant grows over 1.5 metres high is 0.25. A random sample of 40 sunflower plants is taken and each sunflower plant is measured and its height recorded.

(*a*) Find the probability that the number of sunflower plants over 1.5 m high is between 8 and 13 (inclusive) using

(i) a Poisson approximation,

(ii) a Normal approximation.

(10)

(*b*) Write down which of the approximations used in part (*a*) is the most accurate estimate of the probability. You must give a reason for your answer.

(2)

---

7. (*a*) Explain what you understand by

(i) a hypothesis test,

(ii) a critical region.

(3)

During term time, incoming calls to a school are thought to occur at a rate of 0.45 per minute. To test this, the number of calls during a random 20 minute interval, is recorded.

(*b*) Find the critical region for a two-tailed test of the hypothesis that the number of incoming calls occurs at a rate of 0.45 per 1 minute interval. The probability in each tail should be as close to 2.5% as possible.

(5)

(*c*) Write down the actual significance level of the above test.

(1)

In the school holidays, 1 call occurs in a 10 minute interval.

(*d*) Test, at the 5% level of significance, whether or not there is evidence that the rate of incoming calls is less during the school holidays than in term time.

(5)

---

8. The continuous random variable X has probability density function $\mathrm{f}(x)$ given by

$$\mathrm{f}(x) = \begin{cases} 2(x-2) & 2 \le x \le 3 \\ 0 & \text{otherwise} \end{cases}$$

(*a*) Sketch $\mathrm{f}(x)$ for all values of x.

(3)

(*b*) Write down the mode of X.

(1)

Find

(*c*) $\mathrm{E}(X)$,

(3)

(*d*) the median of X.

(4)

(*e*) Comment on the skewness of this distribution. Give a reason for your answer.

(2)

TOTAL FOR PAPER: 75 MARKS

END