GCE AS/A level

985/01

MATHEMATICS S3
STATISTICS 3

A.M. THURSDAY, 23 June 2011

$1\frac{1}{2}$ hours

ADDITIONAL MATERIALS

In addition to this examination paper, you will need:

- a 12 page answer book;
- a Formula Booklet;
- a calculator;
- statistical tables (Murdoch and Barnes or RND/WJEC Publications)

INSTRUCTIONS TO CANDIDATES

Use black ink or black ball-point pen.

Answer **all** questions.

Sufficient working must be shown to demonstrate the **mathematical** method employed.

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 necessity for good English and orderly presentation in your answers.

1. Bill is a darts player and he claims that the probability p of his hitting the 'bull' when he throws a dart is 0·75. In order to test his claim, he throws 100 darts and hits the 'bull' 67 times.

(a) Calculate an unbiased estimate of p. [1]

(b) Calculate an approximate value for the standard error of your estimate. [2]

(c) Calculate an approximate 95% confidence interval for p. [3]

(d) What conclusion do you reach concerning Bill's claim? Justify your answer. [1]

2. A certain type of battery is claimed by the manufacturer to have a mean lifetime of 1500 hours. To test this claim, a consumer organisation defined the following hypotheses,

$$H_0 : \mu = 1{\cdot}5 \quad ; \quad H_1 : \mu \neq 1{\cdot}5$$

where μ denotes the mean lifetime (in thousands of hours). The consumer organisation then determined the lifetimes (x thousand hours) of a random sample of 100 batteries and obtained the following summary statistics.

$$\sum x = 149{\cdot}1, \qquad \sum x^2 = 222{\cdot}9$$

Calculate the p-value of these results and state your conclusion in context. [8]

3. A bag contains five balls numbered 1, 1, 2, 3, 4 respectively. A random sample of three of these balls is taken **without replacement**.

(a) Determine the sampling distribution of the sum of the numbers on the selected balls. [5]

(b) Determine the expected value of the largest number shown on the selected balls. [3]

4. A zoologist discovers a new species of animal on a remote island. He traps 12 males of the species and he weighs each of them with the following results (in kg).

24·1 22·9 21·2 24·7 24·9 23·6 22·9 21·6 23·5 23·0 21·9 24·7

You may assume that this is a random sample from a normal distribution with mean μ and variance σ^2.

(a) Calculate unbiased estimates of μ and σ^2. [4]

(b) Calculate a 90% confidence interval for μ. [5]

5. In a factory, two methods, A and B, are used to complete a certain task. The managing director believes that Method A takes, on average, a shorter time than Method B. A trial was therefore designed to investigate this belief.

(a) State suitable hypotheses. [1]

(b) Method A was used by each of 60 operatives and the times taken (x minutes) gave the following results.

$$\sum x = 1485, \qquad \sum x^2 = 37364$$

Method B was used by each of a different set of 60 operatives and the times taken (y minutes) gave the following results.

$$\sum y = 1560, \qquad \sum y^2 = 41221$$

Test the managing director's belief using a 5% significance level. [10]

6. The solubility y, in appropriate units, of a certain chemical in water is related to the temperature, x °C, by an equation of the form $y = \alpha + \beta x$. In order to estimate α and β, the following measurements were made.

x	10	12	14	16	18	20
y	21·7	24·4	27·3	29·6	31·7	34·5

[You are given that $\sum x = 90$, $\sum x^2 = 1420$, $\sum y = 169{\cdot}2$, $\sum xy = 2626{\cdot}2$]

(a) Calculate least squares estimates for α and β. [6]

(b) The values of x are exact but the values of y are subject to independent normally distributed measurement errors with mean zero and standard deviation 0·15. Determine a 99% confidence interval for the solubility of the chemical in water at 17 °C. [7]

TURN OVER

7. The probability density function of the continuous random variable X is given by

$$f(x) = \frac{1}{2} + \theta x \qquad -1 \leqslant x \leqslant 1,$$
$$f(x) = 0 \qquad \text{otherwise,}$$

where θ is an unknown constant whose value lies between $-\frac{1}{2}$ and $\frac{1}{2}$.

(a) (i) Obtain an expression for $E(X)$ and show that

$$\text{Var}(X) = \frac{3-4\theta^2}{9} .$$

(ii) Show that

$$P(X > 0) = \left(\frac{1+\theta}{2}\right).$$ [8]

In order to estimate θ, a random sample of n observations of X is taken.

(b) The mean of the observations in the sample is denoted by $\overline{X}$. Show that

$$U = \frac{3\overline{X}}{2}$$

is an unbiased estimator for θ and obtain an expression for the variance of U. [4]

(c) Let Y denote the number of observations in the sample that are greater than zero. Show that

$$V = \frac{2Y}{n} - 1$$

is an unbiased estimator for θ and obtain an expression for the variance of V. [5]

(d) Show that

$$\text{Var}(V) - \text{Var}(U) = \frac{1}{4n} .$$

State, with a reason, which is the better estimator, U or V. [2]