GCE AS/A level

985/01

MATHEMATICS S3

STATISTICS 3

A.M. WEDNESDAY, 17 June 2009

1½ 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

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. A bag contains 8 balls, 5 of which are blue and 3 of which are red. A random sample of 5 balls is taken from the bag. Let X denote the number of blue balls in the sample.

(a) When sampling is done **without replacement**,

(i) find the sampling distribution of X,

(ii) calculate $E(X)$. [7]

(b) Suppose now that the sampling is done **with replacement**.

(i) Identify the distribution of X in this case.

(ii) Show that the value of $E(X)$ is the same as in *(a)*. [3]

2. When Sian plays a certain computer game, she has probability p of winning. In order to estimate p, she plays the game 120 times. She wins 78 of these games.

(a) Calculate

(i) an unbiased estimate of p,

(ii) the estimated standard error of this estimate,

(iii) an approximate 95% confidence interval for p. [6]

(b) Information supplied with the game stated that, for people of average intelligence, the value of p should lie between 0·3 and 0·5. What does your interval tell you about Sian? [1]

3. A dairy sells butter in large packs. The owner states that the mean weight of these packs is 1·5 kg. As a quality control check, 80 packs are chosen at random and the weight, x kg, of each pack is measured. The results are summarised below.

$$\sum x = 121{\cdot}2, \quad \sum x^2 = 184{\cdot}42$$

(a) State suitable hypotheses to carry out a two-sided test. [1]

(b) Calculate the p-value of these results and state your conclusion. [8]

(c) State **two** different approximations that you have to make in your analysis. [2]

4. A machine produces ball bearings whose diameters are normally distributed with mean μ mm and standard deviation σ mm. A random sample of 10 of these ball bearings had the following diameters (in mm).

6·12 6·05 6·09 6·16 6·14 6·04 6·08 6·09 6·15 6·18

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

(b) Calculate a 95% confidence interval for μ. [5]

5. A Consumer Organisation wishes to compare the petrol consumption of two similar cars, Model A and Model B. It therefore sets up a trial in which 60 cars of each model are each given 5 litres of petrol and they are driven around a level track at constant speed until they run out of petrol. The distances covered by each car of Model A, x miles, and by each car of Model B, y miles, are recorded. The results are summarised below.

$$\sum x = 3930, \quad \sum x^2 = 258\,000, \quad \sum y = 4020, \quad \sum y^2 = 269\,900.$$

(a) Calculate an approximate 90% confidence interval for the difference between the mean distances travelled on 5 litres of petrol for the two car models. [10]

(b) Does your result indicate that one of the car models is better than the other as regards petrol consumption? [1]

6. The independent random variables X and Y have a common mean μ and variances σ_x^2 and σ_y^2 respectively. In order to estimate μ, random samples of m values of X and n values of Y are taken. The means of these samples are denoted by $\overline{X}$ and $\overline{Y}$ respectively.

(a) Show that

$$U = \lambda \overline{X} + (1 - \lambda)\overline{Y}$$

is an unbiased estimator for μ for all values of the constant λ. [2]

(b) Find an expression for the variance of U. [3]

(c) (i) Determine the value of λ which gives the best estimator for μ.

(ii) Show that the standard error of the best estimator is

$$\frac{\sigma_x \sigma_y}{\sqrt{m\sigma_y^2 + n\sigma_x^2}}.$$

[9]

7. The variables x and y are known to be related by an equation of the form $y = \alpha + \beta x$. In order to estimate the values of α and β, the values of y were measured for six different values of x. The following results were obtained.

x	5	10	15	20	25	30
y	15·5	27·2	37·4	49·1	60·8	72·6

[You are given that $\sum x = 105$, $\sum y = 262{\cdot}6$, $\sum x^2 = 2275$, $\sum xy = 5590{\cdot}5$]

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 0·5.

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

(b) The value of β is thought to be 2·34. The following hypotheses are therefore defined:

$$H_0: \beta = 2{\cdot}34 \text{ versus } H_1: \beta < 2{\cdot}34$$

Calculate the p-value of your result and interpret it. [6]

(c) Alun is given the same data and he evaluates the least squares estimate of β as 0·52. Explain briefly why this answer is obviously incorrect. [1]