GCE AS/A Level

0985/01

S17-0985-01-R1

MATHEMATICS – S3
Statistics

FRIDAY, 23 JUNE 2017 – MORNING

1 hour 30 minutes

ADDITIONAL MATERIALS

In addition to this examination paper, you will need:
- a WJEC pink 16-page answer booklet;
- 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. The weights, X grams, of the eggs sold in a certain farm shop have mean μ grams. To estimate μ, a random sample of 100 eggs was weighed, in grams, and the following sample statistics were calculated.

$$\sum x = 5910,\ \sum x^2 = 349\,425$$

Calculate an approximate 99% confidence interval for μ. [6]

2. Each of three fair dice has its six faces numbered 1, 2, 3, 4, 5, 6 respectively. The three dice are thrown simultaneously and the score on each dice is defined as the number on the uppermost face. Let X denote the highest score on these three dice.

(a) Show that

$$P(X \leqslant x) = \left(\frac{x}{6}\right)^3 \qquad \text{for } x = 1, 2, 3, 4, 5, 6.$$ [2]

(b) Deduce an expression in terms of x for $P(X = x)$, valid for $x = 1, 2, 3, 4, 5, 6$. [2]

(c) Determine the most likely value of X. [2]

3. A zoologist claims that the mean weight of male dogs of a certain breed is 5 kg more than the mean weight of female dogs of the breed. Mair believes that the difference in mean weights is greater than 5 kg. She therefore collects and weighs random samples of 50 male and 50 female dogs of the breed. She defines the following hypotheses,

$$H_0 : \mu_x - \mu_y = 5\ ; \qquad H_1 : \mu_x - \mu_y > 5$$

where μ_x, μ_y denote respectively the mean weights, in kg, of the male dogs and female dogs of the breed. The results are summarised below, where x, y denote respectively the weights, in kg, of the male dogs and the female dogs.

$$\sum x = 2055,\ \sum x^2 = 84\,773,\ \sum y = 1745,\ \sum y^2 = 61121$$

Determine an approximate p-value for these results and state your conclusion in context. [11]

4. A mathematics teacher takes a biased dice to his class, wishing to estimate p, the probability of throwing a 'six'. He throws it 75 times and obtains 24 'sixes'.

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

(b) The teacher calculates this interval and he asks Tom to interpret it. Tom states that 'There is, approximately, a 0·95 probability that the interval that the teacher has calculated contains the unknown value of p'. Explain why this statement is incorrect and give a correct interpretation. [2]

5. When Dawn throws the javelin, the distance thrown (in metres) can be assumed to be normally distributed with mean μ and variance σ^2. She throws the javelin 9 times with the following results.

33·5 34·6 33·3 34·3 34·6 34·0 33·1 35·0 33·6

(a) Calculate unbiased estimates of μ and σ^2. [5]

(b) Calculate a 95% confidence interval for μ. [4]

6. The length, y cm, of a spring subjected to a tension of x Newtons satisfies the relationship $y = \alpha + \beta x$, where α and β are unknown constants. In order to estimate α and β, the following measurements were made.

x	10	15	20	25	30	40
y	12·4	14·3	16·4	18·9	20·7	24·6

You are given that $\sum x = 140,\ \sum y = 107{\cdot}3,\ \sum x^2 = 3850,\ \sum xy = 2744.$

(a) Calculate least squares estimates for α and β, giving your answers correct to three significant figures. [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 0·2. Before the measurements were made, Emlyn believed that the value of β was 0·4.

(i) State suitable hypotheses to carry out a two-sided test of Emlyn's belief.

(ii) Calculate the *p*-value of the above results.

(iii) State whether or not the data support Emlyn's belief. [9]

TURN OVER

7. An electronic device generates random digits from the set {1, 2, 3, 4}. The probability distribution of the digit generated, X, is given by

$$P(X = x) = \begin{cases} p & \text{for } x = 1 \\ \dfrac{(1-p)}{3} & \text{for } x = 2, 3, 4 \end{cases}$$

where p is an unknown constant, $0 < p < 1$.

(a) (i) Determine an expression for $E(X)$ in terms of p.

(ii) Show that

$$\text{Var}(X) = \frac{2}{3}(1-p)(1+6p).$$ [7]

(b) In order to estimate p, a random sample of n digits is generated using the device and $\overline{X}$ denotes the sample mean.

(i) Show that

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

is an unbiased estimator for p.

(ii) Determine an expression for Var(U) in terms of n and p. [4]

(c) The number of digits in the random sample equal to 1 is denoted by Y.

(i) Write down the distribution of Y.

(ii) Show that

$$V = \frac{Y}{n}$$

is an unbiased estimator for p.

(iii) Determine an expression for Var(V) in terms of n and p. [5]

(d) By considering $\dfrac{\text{Var}(U)}{\text{Var}(V)}$, determine which is the better estimator, U or V. [4]

END OF PAPER