WELSH JOINT EDUCATION COMMITTEE
General Certificate of Education
Advanced Subsidiary/Advanced

WJEC CBAC

CYD-BWYLLGOR ADDYSG CYMRU
Tystysgrif Addysg Gyffredinol
Uwch Gyfrannol/Uwch

503/01

MATHEMATICS S3

STATISTICS 3

A.M. WEDNESDAY, 25 January 2006

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

LEGACY SPECIFICATION

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.

INFORMATION FOR CANDIDATES

Graphical calculators may be used for this paper.

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 following ten independent observations were made on a random variable that is normally distributed with mean μ and variance 3·5.

18·4 19·7 20·3 16·0 17·7 19·5 18·9 16·9 18·0 20·6

(a) Determine a 99% confidence interval for μ. [5]

(b) Interpret your confidence interval. [1]

2. A bag contains three red balls and four blue balls. Three of these balls are selected at random, without replacement. The number of red balls selected is denoted by X and the number of blue balls selected is denoted by Y. Find the probability distribution of $X - Y$. [8]

3. A market research organisation wishes to estimate the proportion p of the population supporting a ban on smoking in public places. It therefore questions a random sample of 1200 people and finds that 840 of them support the ban.

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

(b) Calculate, approximately, the standard error of your estimate. [2]

(c) Calculate an approximate 90% confidence interval for p. [4]

4. Alan works in an office and he records the time spent, x minutes, on the telephone on each of 100 randomly chosen days. He summarises his results as follows.

$$\sum x = 6225; \quad \sum x^2 = 393475$$

(a) Calculate unbiased estimates of the mean μ and the variance σ^2 of the time spent per day on the telephone. [3]

(b) Alan had claimed that he spends, on average, 60 minutes per day on the telephone but his employer believes that the average time is greater than this.

(i) State appropriate hypotheses to test Alan's claim.

(ii) Calculate and interpret the p-value of his results. [7]

(c) Giving a reason, state whether or not your analysis requires the assumption that the times spent on the telephone each day are normally distributed. [2]

5. Fred and George are discus throwers. Their coach, Mike, believes that Fred can throw the discus further, on average, than George. In order to test this belief, they each throw five times during a training session and the distances thrown, in metres, were as follows.

Distances thrown by Fred (m)	24·4	23·9	25·0	24·1	24·9
Distances thrown by George (m)	24·2	23·9	24·5	23·6	24·8

You may assume that these are random samples from normal populations with common standard deviation 0·22.

(a) State suitable hypotheses. [1]

(b) Calculate the *p*-value of these results. [7]

(c) State, with a reason, whether or not Mike's belief is supported at a significance level of

(i) 5%,

(ii) 1%. [3]

6. Pat is investigating the relationship between the solubility in water, y g/litre, of a certain substance and the temperature, x °C, of the water. She obtains the following experimental results.

x	10	15	20	25	30	35	40
y	15·3	16·8	17·8	19·6	21·8	23·1	25·2

[You are given that $\sum y = 139{\cdot}6$; $\sum xy = 3721{\cdot}5$]

(a) Assuming a linear relationship $y = \alpha + \beta x$, calculate a and b, the least squares estimates of α and β. [7]

(b) The values of x are exact whereas the measured values of y are subject to independent normally distributed errors with mean zero and standard deviation 0·2.

(i) Find a 95% confidence interval for the value of α.

(ii) The value of β is thought to be 0·35. Find the *p*-value of your value of b, assuming a two-sided alternative. [11]

TURN OVER

7. The discrete random variable X takes the values 0, 1 and 2 with probabilities θ, 3θ and $1 - 4\theta$ respectively, where θ is a constant ($0 < \theta < 0{\cdot}25$).

(a) Show that

(i) $E(X) = 2 - 5\theta$,

(ii) $\text{Var}(X) = 7\theta - 25\theta^2$. [3]

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

(b) Consider the estimator

$$\hat{\theta}_1 = \frac{2 - \overline{X}}{5}$$

where $\overline{X}$ denotes the mean of these observations.

(i) Show that $\hat{\theta}_1$ is unbiased.

(ii) Obtain an expression for the variance of $\hat{\theta}_1$. [4]

(c) Another possible estimator is

$$\hat{\theta}_2 = \frac{N}{n}$$

where N denotes the number of times the value 0 occurs in the sample.

(i) Show that $\hat{\theta}_2$ is unbiased.

(ii) Write down the variance of $\hat{\theta}_2$ and show that

$$\text{Var}\left(\hat{\theta}_2\right) - \text{Var}\left(\hat{\theta}_1\right) = \frac{18\theta}{25n} \cdot$$

(iii) Hence, giving a reason, state which is the better estimator. [6]