cea

Rewarding Learning

ADVANCED

General Certificate of Education

2015

Mathematics

Assessment Unit S4

assessing

Module S2: Statistics 2

[AMS41]

TUESDAY 16 JUNE, AFTERNOON

TIME

1 hour 30 minutes.

INSTRUCTIONS TO CANDIDATES

Write your Centre Number and Candidate Number on the Answer Booklet provided.
Answer **all eight** questions.
Show clearly the full development of your answers.
Answers should be given to three significant figures unless otherwise stated.
You are permitted to use a graphic or a scientific calculator in this paper.

INFORMATION FOR CANDIDATES

The total mark for this paper is 75
Figures in brackets printed down the right-hand side of pages indicate the marks awarded to each question or part question.
A copy of the **Mathematical Formulae and Tables booklet** is provided.
Throughout the paper the logarithmic notation used is $\ln z$ where it is noted that $\ln z \equiv \log_e z$

Answer all eight questions.

Show clearly the full development of your answers.

Answers should be given to three significant figures unless otherwise stated.

Normal and t-distribution values should be read from the tables provided.

1 **(a)** State the effect on the product-moment correlation coefficient between two variables x and y if:

(i) the units of x are changed; [1]

(ii) the origin of x is changed. [1]

(b) A researcher in the Department of Justice decides to investigate if there is a possible correlation between the percentage unemployed, x, and the percentage increase in juvenile court cases, y, over a period of 10 years. The information gathered is shown in **Table 1** below.

Table 1

% Unemployed x	6.6	7.2	7.3	6.7	7.6	7.1	8.6	8.2	7.7	7.6
% Increase y	2.0	1.8	2.2	2.5	2.1	2.2	3.3	2.9	2.7	3.0

The summary statistics are:

n	Σx	Σx^2	Σy	Σy^2	Σxy
10	74.6	560	24.7	63.17	186.3

(i) Calculate the product-moment correlation coefficient. [4]

(ii) Comment on the value obtained in **(i)**. [1]

2 **Table 2** below summarises the number of car thefts occurring in a city on 30 randomly chosen days.

Table 2

Number of thefts	1	2	3	4	5	6	7	8
Number of days	3	5	4	3	5	4	4	2

(i) Calculate unbiased estimates of the mean and variance of the number of thefts per day. [4]

20 more days were randomly selected. Estimates from this sample were mean 5 and variance 5

(ii) Treating the 50 days as a single sample obtain further unbiased estimates of the population mean and variance. [6]

3 A specialist car hire company rents two types of car; vintage cars and stretch limousines. The time to clean each type of car is Normally distributed. For a vintage car the mean time is 15 minutes with standard deviation 4 minutes and for a limousine the mean time is 25 minutes with standard deviation 6 minutes.
On a particular morning there are two vintage cars and one limousine to be cleaned.

Find the probability that the time taken to clean the limousine will be greater than the total time to clean the two vintage cars. [8]

4 Farmer Jones recorded the amount of fertiliser in tonnes applied in spring to eight identical plots of land. He then recorded the autumn yield of hay in tonnes per hectare for each of the plots. His results are given in **Table 3** below.

Table 3

Amount of fertiliser x	3	4.5	6	7.5	9	10.5	12	13.5
Yield of hay y	4.85	5.20	5.76	6.60	7.35	7.50	7.51	7.49

The summary statistics are:

n	Σx	Σx^2	Σy	Σy^2	Σxy
8	66	639	52.26	35.01	458.1

(i) Find the equation of the regression line of y on x in the form $y = a + bx$ [5]

(ii) Interpret the values of a and b in the equation of the regression line. [2]

(iii) Estimate the yield of hay if 7 tonnes of fertiliser were applied. [2]

(iv) By examining the data do you consider a regression line best reflects the relationship between x and y? Give a reason. [1]

5 The age at which students enrolled at a local university was thought to be a Normally distributed random variable with mean 18.1 years.
It is suspected that this mean may have increased.
A random sample of 50 students in the 2014 intake was found to have a mean of 18.5 and variance 2.54

Carry out a suitable test at 5% level to determine whether or not this result provides evidence that the mean age has increased. [7]

6 A breakfast cereal is packed in boxes and the boxes are then packed in cartons each of which holds 10 boxes. The weights of the boxes of cereal are Normally distributed with mean weight 500 grams and standard deviation 12 grams. The weights of the empty cartons are also Normally distributed with mean weight 2000 grams and standard deviation 30 grams. All the random variables are independent.

(i) Find the mean and standard deviation of the weight of a full carton. [4]

(ii) Find the probability that a full carton will weigh between 6900 grams and 7100 grams. [4]

7 Eight candidates competing for a new company post were given two test papers thought to be of the same level of difficulty. Their scores are given in **Table 4** below.

Table 4

Candidate	1	2	3	4	5	6	7	8
Paper 1	106	110	120	120	106	105	98	114
Paper 2	104	114	112	114	92	105	102	117

Carry out the appropriate *t*-test at 5% level to determine whether the papers were of the same level of difficulty. [12]

8 **(a)** Briefly explain what is meant by a '95% confidence interval for the mean'. [1]

(b) The manager of a supermarket commissioned a survey to estimate the mean expenditure by its customers on groceries each week. A random sample of 100 customers were questioned and the resulting 95% confidence interval was (92.16, 107.84).

(i) Show that the mean and standard deviation of the sample are £100 and £40 respectively. [3]

The manager considers the interval too wide and requires a confidence interval of width £10

(ii) Using the mean value from **(i)** write down the new confidence interval. [1]

(iii) Find the percentage confidence for this new interval. [4]

Instead of a reduced level of confidence the manager decides to increase the sample size so that both a 95% confidence level and an interval width of £10 would be created.

(iv) Using the results from **(i)** estimate the smallest size of sample that would satisfy both these conditions. [4]

THIS IS THE END OF THE QUESTION PAPER

9449