Edexcel Maths S4

Past Paper Pack

2008-2013

Centre No.						Paper Reference	Surname	Initial(s)
Candidate No.						6 6 8 6 / 0 1	Signature	

Paper Reference(s)

6686/01

Edexcel GCE

Statistics S4

Advanced/Advanced Subsidiary

Wednesday 18 June 2008 – Morning

Time: 1 hour 30 minutes

Examiner's use only

Team Leader's use only

Question Number	Leave Blank
1	
2	
3	
4	
5	
6	
7	
Total	

Materials required for examination
Mathematical Formulae (Green)

Items included with question papers
Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulae stored in them.

Instructions to Candidates

In the boxes above, write your centre number, candidate number, your surname, initials and signature.
Check that you have the correct question paper.
Answer ALL the questions.
You must write your answer for each question in the space following the question.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
The marks for individual questions and the parts of questions are shown in round brackets: e.g. **(2)**.
There are 7 questions in this question paper. The total mark for this paper is 75.
There are 24 pages in this question paper. Any blank pages are indicated.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You should show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

Printer's Log. No.
H31327A
W850/R6686/57570 3/3/3

H 3 1 3 2 7 A 0 1 2 4

Turn over

edexcel
advancing learning, changing lives

Leave blank

1. A random sample $X_1, X_2, ..., X_{10}$ is taken from a population with mean μ and variance σ^2.

(a) Determine the bias, if any, of each of the following estimators of μ.

$$\theta_1 = \frac{X_3 + X_4 + X_5}{3},$$

$$\theta_2 = \frac{X_{10} - X_1}{3},$$

$$\theta_3 = \frac{3X_1 + 2X_2 + X_{10}}{6}.$$

(4)

(b) Find the variance of each of these estimators.

(5)

(c) State, giving reasons, which of these three estimators for μ is

(i) the best estimator,

(ii) the worst estimator.

(4)

Leave blank

Question 1 continued

2. A large number of students are split into two groups A and B. The students sit the same test but under different conditions. Group A has music playing in the room during the test, and group B has no music playing during the test. Small samples are then taken from each group and their marks recorded. The marks are normally distributed.

The marks are as follows:

Sample from Group A	42	40	35	37	34	43	42	44	49
Sample from Group B	40	44	38	47	38	37	33		

(a) Stating your hypotheses clearly, and using a 10% level of significance, test whether or not there is evidence of a difference between the variances of the marks of the two groups.

(8)

(b) State clearly an assumption you have made to enable you to carry out the test in part (a).

(1)

(c) Use a two tailed test, with a 5% level of significance, to determine if the playing of music during the test has made any difference in the mean marks of the two groups. State your hypotheses clearly.

(7)

(d) Write down what you can conclude about the effect of music on a student's performance during the test.

(1)

Leave blank

Question 2 continued

H31327A0724

Turn over

Leave blank

3. The weights, in grams, of mice are normally distributed. A biologist takes a random sample of 10 mice. She weighs each mouse and records its weight.

The ten mice are then fed on a special diet. They are weighed again after two weeks.

Their weights in grams are as follows:

Mouse	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
Weight before diet	50.0	48.3	47.5	54.0	38.9	42.7	50.1	46.8	40.3	41.2
Weight after diet	52.1	47.6	50.1	52.3	42.2	44.3	51.8	48.0	41.9	43.6

Stating your hypotheses clearly, and using a 1% level of significance, test whether or not the diet causes an increase in the mean weight of the mice.

(8)

H 3 1 3 2 7 A 0 1 0 2 4

Leave blank

4. A town council is concerned that the mean price of renting two bedroom flats in the town has exceeded £650 per month. A random sample of eight two bedroom flats gave the following results, £x, per month.

705, 640, 560, 680, 800, 620, 580, 760

[You may assume $\sum x = 5345$ $\sum x^2 = 3621025$]

(a) Find a 90% confidence interval for the mean price of renting a two bedroom flat. **(6)**

(b) State an assumption that is required for the validity of your interval in part (a). **(1)**

(c) Comment on whether or not the town council is justified in being concerned. Give a reason for your answer. **(2)**

Leave blank

5. A machine is filling bottles of milk. A random sample of 16 bottles was taken and the volume of milk in each bottle was measured and recorded. The volume of milk in a bottle is normally distributed and the unbiased estimate of the variance, s^2, of the volume of milk in a bottle is 0.003

(a) Find a 95% confidence interval for the variance of the population of volumes of milk from which the sample was taken.

(5)

The machine should fill bottles so that the standard deviation of the volumes is equal to 0.07

(b) Comment on this with reference to your 95% confidence interval.

(3)

Leave blank

Question 5 continued

Q5

(Total 8 marks)

Leave blank

6. A drug is claimed to produce a cure to a certain disease in 35% of people who have the disease. To test this claim a sample of 20 people having this disease is chosen at random and given the drug. If the number of people cured is between 4 and 10 inclusive the claim will be accepted. Otherwise the claim will not be accepted.

(a) Write down suitable hypotheses to carry out this test.
(2)

(b) Find the probability of making a Type I error.
(3)

The table below gives the value of the probability of the Type II error, to 4 decimal places, for different values of p where p is the probability of the drug curing a person with the disease.

P(cure)	0.2	0.3	0.4	0.5
P(Type II error)	0.5880	r	0.8565	s

(c) Calculate the value of r and the value of s.
(3)

(d) Calculate the power of the test for $p = 0.2$ and $p = 0.4$
(2)

(e) Comment, giving your reasons, on the suitability of this test procedure.
(2)

H 3 1 3 2 7 A 0 1 6 2 4

Leave blank

Question 6 continued

H31327A01724

Leave blank

7. An engineering firm buys steel rods. The steel rods from its present supplier are known to have a mean tensile strength of 230 N/mm^2.

A new supplier of steel rods offers to supply rods at a cheaper price than the present supplier. A random sample of ten rods from this new supplier gave tensile strengths, x N/mm^2, which are summarised below.

Sample size	Σx	Σx^2
10	2283	524 079

(a) Stating your hypotheses clearly, and using a 5% level of significance, test whether or not the rods from the new supplier have a tensile strength lower than the present supplier. (You may assume that the tensile strength is normally distributed).

(7)

(b) In the light of your conclusion to part (a) write down what you would recommend the engineering firm to do.

(1)

Leave blank

Question 7 continued

H31327A02124

Turn over

June 2009

Centre No.						Paper Reference	Surname	Initial(s)
Candidate No.						6 6 8 6 / 0 1	Signature	

Paper Reference(s)

6686/01

Edexcel GCE

Statistics S4

Advanced/Advanced Subsidiary

Friday 19 June 2009 – Afternoon

Time: 1 hour 30 minutes

Examiner's use only

Team Leader's use only

Question Number	Leave Blank
1	
2	
3	
4	
5	
6	
Total	

Materials required for examination	Items included with question papers
Mathematical Formulae (Orange or Green)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulae stored in them.

Instructions to Candidates

In the boxes above, write your centre number, candidate number, your surname, initials and signature.
Check that you have the correct question paper.
Answer ALL the questions. You must write your answer for each question in the space following the question.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
The marks for individual questions and the parts of questions are shown in round brackets: e.g. **(2)**.
There are 6 questions in this question paper. The total mark for this paper is 75.
There are 24 pages in this question paper. Any blank pages are indicated.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You should show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

Printer's Log. No.
M34281A
W850/R6686/57570 8/4/3/3/
M 3 4 2 8 1 A 0 1 2 4

Turn over

edexcel
advancing learning, changing lives

Leave blank

1. A company manufactures bolts with a mean diameter of 5 mm. The company wishes to check that the diameter of the bolts has not decreased. A random sample of 10 bolts is taken and the diameters, x mm, of the bolts are measured. The results are summarised below.

$$\sum x = 49.1 \quad \sum x^2 = 241.2$$

Using a 1% level of significance, test whether or not the mean diameter of the bolts is less than 5 mm.

(You may assume that the diameter of the bolts follows a normal distribution.)

(8)

Leave blank

2. An emission-control device is tested to see if it reduces CO_2 emissions from cars. The emissions from 6 randomly selected cars are measured with and without the device. The results are as follows.

Car	*A*	*B*	*C*	*D*	*E*	*F*
Emissions without device	151.4	164.3	168.5	148.2	139.4	151.2
Emissions with device	148.9	162.7	166.9	150.1	140.0	146.7

(a) State an assumption that needs to be made in order to carry out a *t*-test in this case.
(1)

(b) State why a paired *t*-test is suitable for use with these data.
(1)

(c) Using a 5% level of significance, test whether or not there is evidence that the device reduces CO_2 emissions from cars.
(8)

(d) Explain, in context, what a type II error would be in this case.
(2)

Leave blank

Question 2 continued

M34281A0524

Turn over

Leave blank

3. Define, in terms of H_0 and/or H_1,

(a) the size of a hypothesis test,

(1)

(b) the power of a hypothesis test.

(1)

The probability of getting a head when a coin is tossed is denoted by p.

This coin is tossed 12 times in order to test the hypotheses H_0: $p = 0.5$ against H_1: $p \neq 0.5$, using a 5% level of significance.

(c) Find the largest critical region for this test, such that the probability in each tail is less than 2.5%.

(4)

(d) Given that $p = 0.4$

(i) find the probability of a type II error when using this test,

(ii) find the power of this test.

(4)

(e) Suggest two ways in which the power of the test can be increased.

(2)

Leave blank

Question 3 continued

M34281A0924

Turn over

Leave blank

4. A farmer set up a trial to assess whether adding water to dry feed increases the milk yield of his cows. He randomly selected 22 cows. Thirteen of the cows were given dry feed and the other 9 cows were given the feed with water added. The milk yields, in litres per day, were recorded with the following results.

	Sample size	Mean	s^2
Dry feed	13	25.54	2.45
Feed with water added	9	27.94	1.02

You may assume that the milk yield from cows given the dry feed and the milk yield from cows given the feed with water added are from independent normal distributions.

(a) Test, at the 10% level of significance, whether or not the variances of the populations from which the samples are drawn are the same. State your hypotheses clearly.
(5)

(b) Calculate a 95% confidence interval for the difference between the two mean milk yields.
(7)

(c) Explain the importance of the test in part (a) to the calculation in part (b).
(2)

Leave blank

Question 4 continued

M34281A01324

Turn over

Leave blank

5. A machine fills jars with jam. The weight of jam in each jar is normally distributed. To check the machine is working properly the contents of a random sample of 15 jars are weighed in grams. Unbiased estimates of the mean and variance are obtained as

$$\hat{\mu} = 560 \quad s^2 = 25.2$$

Calculate a 95% confidence interval for,

(a) the mean weight of jam,

(4)

(b) the variance of the weight of jam.

(5)

A weight of more than 565 g is regarded as too high and suggests the machine is not working properly.

(c) Use appropriate confidence limits from parts (a) and (b) to find the highest estimate of the proportion of jars that weigh too much.

(5)

Leave blank

Question 5 continued

Leave blank

6. A continuous uniform distribution on the interval $[0, k]$ has mean $\frac{k}{2}$ and variance $\frac{k^2}{12}$.

A random sample of three independent variables X_1, X_2 and X_3 is taken from this distribution.

(a) Show that $\frac{2}{3}X_1 + \frac{1}{2}X_2 + \frac{5}{6}X_3$ is an unbiased estimator for k.

(3)

An unbiased estimator for k is given by $\hat{k} = aX_1 + bX_2$ where a and b are constants.

(b) Show that $\text{Var}(\hat{k}) = (a^2 - 2a + 2)\,\frac{k^2}{6}$

(6)

(c) Hence determine the value of a and the value of b for which $\hat{k}$ has minimum variance, and calculate this minimum variance.

(6)

Leave blank

Question 6 continued

June 2010

Centre No.						Paper Reference	Surname	Initial(s)
Candidate No.						6 6 8 6 / 0 1	Signature	

Paper Reference(s)

6686/01

Edexcel GCE

Statistics S4

Advanced/Advanced Subsidiary

Thursday 24 June 2010 – Morning

Time: 1 hour 30 minutes

Examiner's use only

Team Leader's use only

Question Number	Leave Blank
1	
2	
3	
4	
5	
6	
Total	

Materials required for examination
Mathematical Formulae (Pink)

Items included with question papers
Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulae stored in them.

Instructions to Candidates

In the boxes above, write your centre number, candidate number, your surname, initials and signature.
Check that you have the correct question paper.
Answer ALL the questions.
You must write your answer to each question in the space following the question.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
The marks for individual questions and the parts of questions are shown in round brackets: e.g. **(2)**.
There are 6 questions in this question paper. The total mark for this paper is 75.
There are 24 pages in this question paper. Any blank pages are indicated.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You should show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

Printer's Log. No.
M35398A
W850/R6686/57570 4/5/5/3

M 3 5 3 9 8 A 0 1 2 4

Turn over

edexcel
advancing learning, changing lives

Leave blank

1. A teacher wishes to test whether playing background music enables students to complete a task more quickly. The same task was completed by 15 students, divided at random into two groups. The first group had background music playing during the task and the second group had no background music playing.
The times taken, in minutes, to complete the task are summarised below.

	Sample size n	Standard deviation s	Mean $\bar{x}$
With background music	8	4.1	15.9
Without background music	7	5.2	17.9

You may assume that the times taken to complete the task by the students are two independent random samples from normal distributions.

(a) Stating your hypotheses clearly, test, at the 10% level of significance, whether or not the variances of the times taken to complete the task with and without background music are equal.

(5)

(b) Find a 99% confidence interval for the difference in the mean times taken to complete the task with and without background music.

(7)

Experiments like this are often performed using the same people in each group.

(c) Explain why this would not be appropriate in this case.

(1)

Leave blank

Question 1 continued

Leave blank

2. As part of an investigation, a random sample of 10 people had their heart rate, in beats per minute, measured whilst standing up and whilst lying down. The results are summarised below.

Person	1	2	3	4	5	6	7	8	9	10
Heart rate lying down	66	70	59	65	72	66	62	69	56	68
Heart rate standing up	75	76	63	67	80	75	65	74	63	75

(a) State one assumption that needs to be made in order to carry out a paired *t*-test.

(1)

(b) Test, at the 5% level of significance, whether or not there is any evidence that standing up increases people's mean heart rate by more than 5 beats per minute. State your hypotheses clearly.

(8)

Leave blank

Question 2 continued

Leave blank

3. A manager in a sweet factory believes that the machines are working incorrectly and the proportion p of underweight bags of sweets is more than 5%. He decides to test this by randomly selecting a sample of 5 bags and recording the number X that are underweight. The manager sets up the hypotheses H_0: $p = 0.05$ and H_1: $p > 0.05$ and rejects the null hypothesis if $x > 1$.

(a) Find the size of the test.

(2)

(b) Show that the power function of the test is

$$1 - (1 - p)^4(1 + 4p)$$

(3)

The manager goes on holiday and his deputy checks the production by randomly selecting a sample of 10 bags of sweets. He rejects the hypothesis that $p = 0.05$ if more than 2 underweight bags are found in the sample.

(c) Find the probability of a Type I error using the deputy's test.

(2)

Question 3 continues on page 12

Leave blank

Question 3 continued

Question 3 continues on the next page

Leave blank

Question 3 continued

The table below gives some values, to 2 decimal places, of the power function for the deputy's test.

p	0.10	0.15	0.20	0.25
Power	0.07	s	0.32	0.47

(d) Find the value of s.

(1)

The graph of the power function for the manager's test is shown in Figure 1.

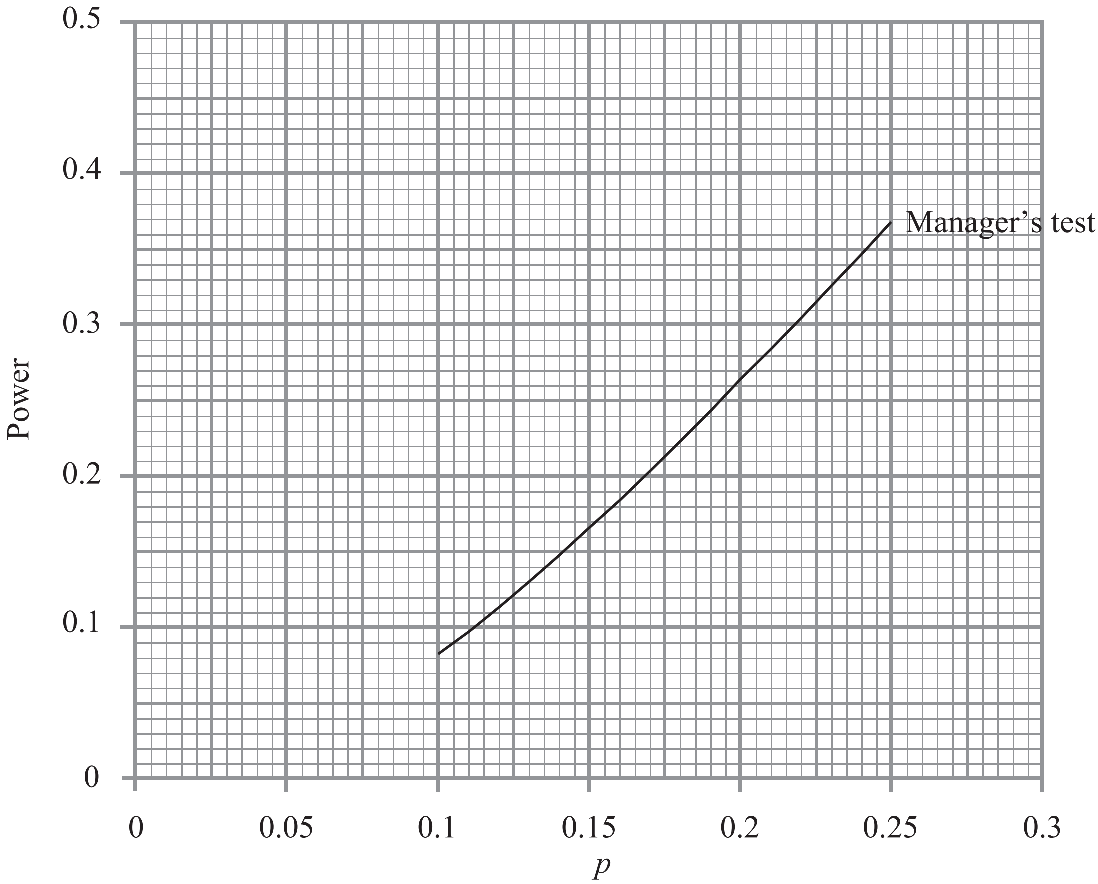


Figure 1

(e) On the same axes, draw the graph of the power function for the deputy's test.

(1)

(f) (i) State the value of p where these graphs intersect.

(ii) Compare the effectiveness of the two tests if p is greater than this value.

(2)

The deputy suggests that they should use his sampling method rather than the manager's.

(g) Give a reason why the manager might not agree to this change.

(1)

M 3 5 3 9 8 A 0 1 2 2 4

Leave blank

Question 3 continued

Q3

(Total 12 marks)

Leave blank

4. A random sample of 15 strawberries is taken from a large field and the weight x grams of each strawberry is recorded. The results are summarised below.

$$\sum x = 291 \qquad \sum x^2 = 5968$$

Assume that the weights of strawberries are normally distributed.
Calculate a 95% confidence interval for

(a) (i) the mean of the weights of the strawberries in the field,

(ii) the variance of the weights of the strawberries in the field.

(12)

Strawberries weighing more than 23 g are considered to be less tasty.

(b) Use appropriate confidence limits from part (a) to find the highest estimate of the proportion of strawberries that are considered to be less tasty.

(4)

Leave blank

Question 4 continued

Leave blank

5. A car manufacturer claims that, on a motorway, the mean number of miles per gallon for the Panther car is more than 70. To test this claim a car magazine measures the number of miles per gallon, x, of each of a random sample of 20 Panther cars and obtained the following statistics.

$$\bar{x} = 71.2 \qquad s = 3.4$$

The number of miles per gallon may be assumed to be normally distributed.

(a) Stating your hypotheses clearly and using a 5% level of significance, test the manufacturer's claim.
(5)

The standard deviation of the number of miles per gallon for the Tiger car is 4.

(b) Stating your hypotheses clearly, test, at the 5% level of significance, whether or not there is evidence that the variance of the number of miles per gallon for the Panther car is different from that of the Tiger car.
(6)

Leave blank

Question 5 continued

Leave blank

6. Faults occur in a roll of material at a rate of λ per m^2. To estimate λ, three pieces of material of sizes $3\,m^2$, $7\,m^2$ and $10\,m^2$ are selected and the number of faults X_1, X_2 and X_3 respectively are recorded.

The estimator $\hat{\lambda}$, where

$$\hat{\lambda} = k\,(X_1 + X_2 + X_3)$$

is an unbiased estimator of λ.

(a) Write down the distributions of X_1, X_2 and X_3 and find the value of k.

(4)

(b) Find Var($\hat{\lambda}$).

(3)

A random sample of n pieces of this material, each of size $4\,m^2$, was taken. The number of faults on each piece, Y, was recorded.

(c) Show that $\frac{1}{4}\overline{Y}$ is an unbiased estimator of λ.

(2)

(d) Find Var($\frac{1}{4}\overline{Y}$).

(3)

(e) Find the minimum value of n for which $\frac{1}{4}\overline{Y}$ becomes a better estimator of λ than $\hat{\lambda}$.

(2)

Leave blank

Question 6 continued

M 3 5 3 9 8 A 0 2 3 2 4

Turn over

June 2011

Centre No.						Paper Reference							Surname	Initial(s)
Candidate No.						6	6	8	6	/	0	1	Signature	

Paper Reference(s)

6686/01

Edexcel GCE

Statistics S4

Advanced/Advanced Subsidiary

Thursday 23 June 2011 – Morning

Time: 1 hour 30 minutes

Examiner's use only

Team Leader's use only

Question Number	Leave Blank
1	
2	
3	
4	
5	
6	
7	
Total	

Materials required for examination	Items included with question papers
Mathematical Formulae (Pink)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation or symbolic differentiation/integration, or have retrievable mathematical formulae stored in them.

Instructions to Candidates

In the boxes above, write your centre number, candidate number, your surname, initials and signature.
Check that you have the correct question paper.
Answer ALL the questions.
You must write your answer to each question in the space following the question.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
The marks for individual questions and the parts of questions are shown in round brackets: e.g. **(2)**.
There are 7 questions in this question paper. The total mark for this paper is 75.
There are 20 pages in this question paper. Any blank pages are indicated.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You should show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

Printer's Log. No.
P35418A
W850/R6686/57570 5/3/3
P 3 5 4 1 8 A 0 1 2 0

Turn over

Leave blank

1. Find the value of the constant a such that

$$P(a < F_{8,10} < 3.07) = 0.94$$

(2)

(Total 2 marks)

Q1

P 3 5 4 1 8 A 0 2 2 0

Leave blank

2. Two independent random samples $X_1, X_2, ..., X_7$ and Y_1, Y_2, Y_3, Y_4 were taken from different normal populations with a common standard deviation σ.
The following sample statistics were calculated.

$$s_x = 14.67 \quad s_y = 12.07$$

Find the 99% confidence interval for σ^2 based on these two samples.

(5)

Q2

(Total 5 marks)

Leave blank

3. Manuel is planning to buy a new machine to squeeze oranges in his cafe and he has two models, at the same price, on trial. The manufacturers of machine *B* claim that their machine produces more juice from an orange than machine *A*. To test this claim Manuel takes a random sample of 8 oranges, cuts them in half and puts one half in machine *A* and the other half in machine *B*. The amount of juice, in ml, produced by each machine is given in the table below.

Orange	1	2	3	4	5	6	7	8
Machine *A*	60	58	55	53	52	51	54	56
Machine *B*	61	60	58	52	55	50	52	58

Stating your hypotheses clearly, test, at the 10% level of significance, whether or not the mean amount of juice produced by machine *B* is more than the mean amount produced by machine *A*.

(8)

Leave blank

4. A proportion p of letters sent by a company are incorrectly addressed and if p is thought to be greater than 0.05 then action is taken.
Using $H_0: p = 0.05$ and $H_1: p > 0.05$, a manager from the company takes a random sample of 40 letters and rejects H_0 if the number of incorrectly addressed letters is more than 3.

(a) Find the size of this test.

(2)

(b) Find the probability of a Type II error in the case where p is in fact 0.10

(2)

Table 1 below gives some values, to 2 decimal places, of the power function of this test.

p	0.075	0.100	0.125	0.150	0.175	0.200	0.225
Power	0.35	s	0.75	0.87	0.94	0.97	0.99

Table 1

(c) Write down the value of s.

(1)

A visiting consultant uses an alternative system to test the same hypotheses. A sample of 15 letters is taken. If these are all correctly addressed then H_0 is accepted. If 2 or more are found to have been incorrectly addressed then H_0 is rejected. If only one is found to be incorrectly addressed then a further random sample of 15 is taken and H_0 is rejected if 2 or more are found to have been incorrectly addressed in this second sample, otherwise H_0 is accepted.

(d) Find the size of the test used by the consultant.

(3)

Question 4 continues on page 8

Leave blank

Question 4 continued

Question 4 continues on the next page

Leave blank

For your convenience Table 1 is repeated here

p	0.075	0.100	0.125	0.150	0.175	0.200	0.225
Power	0.35	s	0.75	0.87	0.94	0.97	0.99

Table 1

Figure 1 shows the graph of the power function of the test used by the consultant.

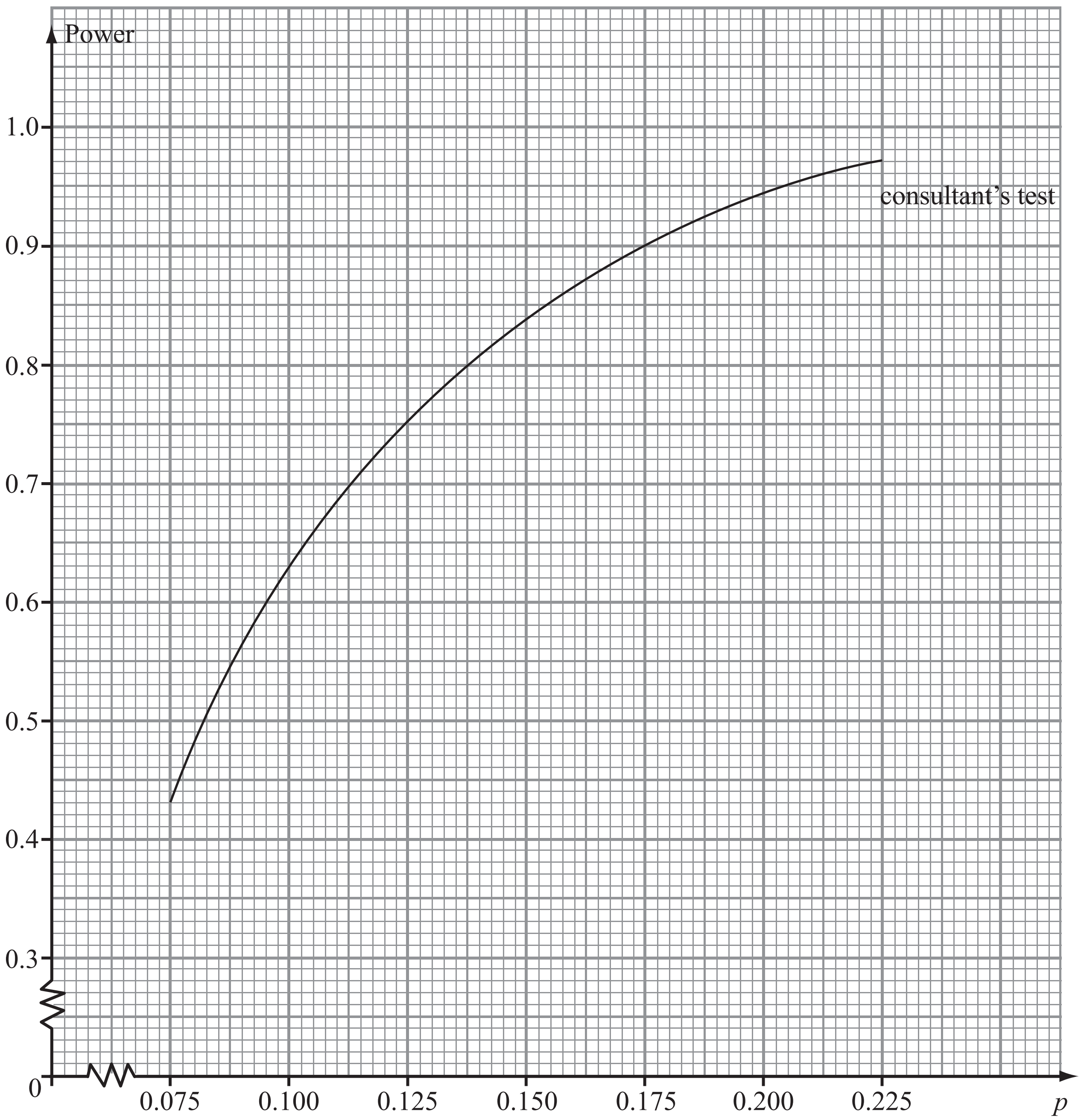


Figure 1

(e) On Figure 1 draw the graph of the power function of the manager's test. **(2)**

(f) State, giving your reasons, which test you would recommend. **(2)**

P 3 5 4 1 8 A 0 8 2 0

Leave blank

Question 4 continued

(Total 12 marks)

Q4

Leave blank

5. The weights of the contents of breakfast cereal boxes are normally distributed.
A manufacturer changes the style of the boxes but claims that the weight of the contents remains the same.
A random sample of 6 old style boxes had contents with the following weights (in grams).

512 503 514 506 509 515

The weights, y grams, of the contents of an independent random sample of 5 new style boxes gave

$$\bar{y} = 504.8 \text{ and } s_y = 3.420$$

(a) Use a two-tail test to show, at the 10% level of significance, that the variances of the weights of the contents of the old and new style boxes can be assumed to be equal. State your hypotheses clearly.
(5)

(b) Showing your working clearly, find a 90% confidence interval for $\mu_x - \mu_y$, where μ_x and μ_y are the mean weights of the contents of old and new style boxes respectively.
(7)

(c) With reference to your confidence interval comment on the manufacturer's claim.
(2)

Leave blank

Question 5 continued

Leave blank

6. A random sample $X_1, X_2, ..., X_n$ is taken from a population where each of the X_i have a continuous uniform distribution over the interval $[0, \beta]$.
The random variable $Y = \max\{X_1, X_2, ..., X_n\}$.
The probability density function of Y is given by

$$\mathrm{f}(y) = \begin{cases} \dfrac{n}{\beta^n} y^{n-1} & 0 \leqslant y \leqslant \beta \\ 0 & \text{otherwise} \end{cases}$$

(a) Show that $\mathrm{E}(Y^m) = \dfrac{n}{n+m}\beta^m$.

(3)

(b) Write down $\mathrm{E}(Y)$.

(1)

(c) Using your answers to parts (a) and (b), or otherwise, show that

$$\mathrm{Var}(Y) = \frac{n}{(n+1)^2(n+2)}\beta^2$$

(3)

(d) State, giving your reasons, whether or not Y is a consistent estimator of β.

(3)

The random variables $M = 2\bar{X}$, where $\bar{X} = \frac{1}{n}(X_1 + X_2 + ... + X_n)$, and $S = kY$, where k is a constant, are both unbiased estimators of β.

(e) Find the value of k in terms of n.

(1)

(f) State, giving your reasons, which of M and S is the better estimator of β in this case.

(3)

Five observations of X are: 8.5 6.3 5.4 9.1 7.6

(g) Calculate the better estimate of β.

(2)

Leave blank

Question 6 continued

Leave blank

7. A machine produces components whose lengths are normally distributed with mean 102.3 mm and standard deviation 2.8 mm. After the machine had been serviced, a random sample of 20 components were tested to see if the mean and standard deviation had changed. The lengths, x mm, of each of these 20 components are summarised as

$$\sum x = 2072 \quad \sum x^2 = 214\,856$$

(a) Stating your hypotheses clearly, test, at the 5% level of significance, whether or not there is evidence of a change in standard deviation.

(7)

(b) Stating your hypotheses clearly, test, at the 5% level of significance, whether or not the mean length of the components has changed from the original value of 102.3 mm using

(i) a normal distribution,
(ii) a t distribution.

(9)

(c) Comment on the mean length of components produced after the service in the light of the tests from part (a) and part (b). Give a reason for your answer.

(2)

Leave blank

Question 7 continued

Centre No.						Paper Reference	Surname	Initial(s)
Candidate No.						6 6 8 6 / 0 1	Signature	

Paper Reference(s)

6686/01

Edexcel GCE

Statistics S4

Advanced/Advanced Subsidiary

Friday 22 June 2012 – Afternoon

Time: 1 hour 30 minutes

Examiner's use only

Team Leader's use only

Question Number	Leave Blank
1	
2	
3	
4	
5	
6	
Total	

Materials required for examination
Mathematical Formulae (Pink)

Items included with question papers
Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation or symbolic differentiation/integration, or have retrievable mathematical formulae stored in them.

Instructions to Candidates

In the boxes above, write your centre number, candidate number, your surname, initials and signature.
Check that you have the correct question paper.
Answer ALL the questions.
You must write your answer to each question in the space following the question.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
The marks for individual questions and the parts of questions are shown in round brackets: e.g. **(2)**.
There are 6 questions in this question paper. The total mark for this paper is 75.
There are 24 pages in this question paper. Any blank pages are indicated.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You should show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

Printer's Log. No.
P40473A
W850/R6684/57570 5/4/5/3

P 4 0 4 7 3 A 0 1 2 4

Turn over

PEARSON

Leave blank

1. A medical student is investigating whether there is a difference in a person's blood pressure when sitting down and after standing up. She takes a random sample of 12 people and measures their blood pressure, in mmHg, when sitting down and after standing up.

The results are shown below.

Person	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*
Sitting down	135	146	138	146	141	158	136	135	146	161	119	151
Standing up	131	147	132	140	138	160	127	136	142	154	130	144

The student decides to carry out a paired t-test to investigate whether, on average, the blood pressure of a person when sitting down is more than their blood pressure after standing up.

(a) State clearly the hypotheses that should be used and any necessary assumption that needs to be made.

(2)

(b) Carry out the test at the 1% level of significance.

(7)

Leave blank

Question 1 continued

(Total 9 marks)

Q1

Leave blank

2. A biologist investigating the shell size of turtles takes random samples of adult female and adult male turtles and records the length, x cm, of the shell. The results are summarised below.

	Number in sample	Sample mean $\bar{x}$	$\sum x^2$
Female	6	19.6	2308.01
Male	12	13.7	2262.57

You may assume that the samples come from independent normal distributions with the same variance.

The biologist claims that the mean shell length of adult female turtles is 5 cm longer than the mean shell length of adult male turtles.

(a) Test the biologist's claim at the 5% level of significance.

(10)

(b) Given that the true values for the variance of the population of adult male turtles and adult female turtles are both 0.9 cm^2,

(i) show that when samples of size 6 and 12 are used with a 5% level of significance, the biologist's claim will be accepted if $4.07 < \bar{X}_F - \bar{X}_M < 5.93$ where $\bar{X}_F$ and $\bar{X}_M$ are the mean shell lengths of females and males respectively.

(ii) Hence find the probability of a type II error for this test if in fact the true mean shell length of adult female turtles is 6 cm more than the mean shell length of adult male turtles.

(6)

Leave blank

Question 2 continued

Leave blank

3. The sample variance of the lengths of a random sample of 9 paving slabs sold by a builders' merchant is 36 mm^2. The sample variance of the lengths of a random sample of 11 paving slabs sold by a second builders' merchant is 225 mm^2. Test at the 10% significance level whether or not there is evidence that the lengths of paving slabs sold by these builders' merchants differ in variability. State your hypotheses clearly.

(You may assume the lengths of paving slabs are normally distributed.)

(5)

Leave blank

4. A newspaper runs a daily Sudoku. A random sample of 10 people took the following times, in minutes, to complete the Sudoku.

5.0 4.5 4.7 5.3 5.2 4.1 5.3 4.8 5.5 4.6

Given that the times to complete the Sudoku follow a normal distribution,

(a) calculate a 95% confidence interval for

(i) the mean,

(ii) the variance,

of the times taken by people to complete the Sudoku.

(13)

The newspaper requires the average time needed to complete the Sudoku to be 5 minutes with a standard deviation of 0.7 minutes.

(b) Comment on whether or not the Sudoku meets this requirement. Give a reason for your answer.

(3)

Leave blank

Question 4 continued

Leave blank

5. Boxes of chocolates manufactured by Philippe have a mean weight of μ grams and a standard deviation of σ grams. A random sample of 25 of these boxes are weighed. Using this sample, the unbiased estimate of μ is 455 and the unbiased estimate of σ^2 is 55.

(a) Test, at the 5% level of significance, whether or not σ is greater than 6. State your hypotheses clearly.
(6)

(b) Test, at the 5% level of significance, whether or not μ is more than 450.
(6)

(c) State an assumption you have made in order to carry out the above tests.
(1)

Leave blank

Question 5 continued

Leave blank

6. When a tree seed is planted the probability of it germinating is p.
A random sample of size n is taken and the number of tree seeds, X, which germinate is recorded.

(a) (i) Show that $\hat{p}_1 = \frac{X}{n}$ is an unbiased estimator of p.

(ii) Find the variance of $\hat{p}_1$.

(4)

A second sample of size m is taken and the number of tree seeds, Y, which germinate is recorded.

Given that $\hat{p}_2 = \frac{Y}{m}$ and that $\hat{p}_3 = a(3\hat{p}_1 + 2\hat{p}_2)$ is an unbiased estimator of p,

(b) show that

(i) $a = \frac{1}{5}$,

(ii) $\text{Var}(\hat{p}_3) = \frac{p(1-p)}{25}\left(\frac{9}{n} + \frac{4}{m}\right)$.

(6)

(c) Find the range of values of $\frac{n}{m}$ for which

$$\text{Var}(\hat{p}_3) < \text{Var}(\hat{p}_1) \text{ and } \text{Var}(\hat{p}_3) < \text{Var}(\hat{p}_2)$$

(3)

(d) Given that $n = 20$ and $m = 60$, explain which of $\hat{p}_1$, $\hat{p}_2$ or $\hat{p}_3$ is the best estimator.

(3)

Leave blank

Question 6 continued

Centre No.						Paper Reference	Surname	Initial(s)
Candidate No.						6 6 8 6 / 0 1 R	Signature	

Paper Reference(s)

6686/01R

Edexcel GCE

Statistics S4

Advanced/Advanced Subsidiary

Friday 21 June 2013 – Morning

Time: 1 hour 30 minutes

Examiner's use only

Team Leader's use only

Question Number	Leave Blank
1	
2	
3	
4	
5	
6	
7	
8	
Total	

Materials required for examination	Items included with question papers
Mathematical Formulae (Pink)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation or symbolic differentiation/integration, or have retrievable mathematical formulae stored in them.

Instructions to Candidates

In the boxes above, write your centre number, candidate number, your surname, initials and signature.
Check that you have the correct question paper.
Answer ALL the questions.
You must write your answer to each question in the space following the question.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
The marks for individual questions and the parts of questions are shown in round brackets: e.g. **(2)**.
There are 8 questions in this question paper. The total mark for this paper is 75.
There are 24 pages in this question paper. Any blank pages are indicated.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You should show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

Printer's Log. No.
P42960A
W850/R6684/57570 5/5/5/

P 4 2 9 6 0 A 0 1 2 4

Turn over

PEARSON

Leave blank

1. (a) Find the value of the constant a such that

$$P(1.690 < \chi_7^2 < a) = 0.95$$

(2)

The random variable Y follows an F-distribution with 6 and 4 degrees of freedom.

(b) (i) Find the upper 1% critical value for Y.

(ii) Find the lower 1% critical value for Y.

(2)

Q1

(Total 4 marks)

Leave blank

2. The time, t hours, that a typist can sit before incurring back pain is modelled by $N(\mu, \sigma^2)$. A random sample of 30 typists gave unbiased estimates for μ and σ^2 as shown below.

$$\hat{\mu} = 2.5 \quad s^2 = 0.36$$

(a) Find a 95% confidence interval for σ^2

(5)

(b) State with a reason whether or not the confidence interval supports the assertion that $\sigma^2 = 0.495$

(2)

Q2

(Total 7 marks)

P 4 2 9 6 0 A 0 3 2 4

Turn over

Leave blank

3. The number of houses sold per week by a firm of estate agents follows a Poisson distribution with mean 2. The firm believes that the appointment of a new salesman will increase the number of houses sold. The firm tests its belief by recording the number of houses sold, x, in the week following the appointment. The firm sets up the hypotheses $H_0 : \lambda = 2$ and $H_1 : \lambda > 2$, where λ is the mean number of houses sold per week, and rejects the null hypothesis if $x \geqslant 3$

(a) Find the size of the test.

(2)

(b) Show that the power function for this test is

$$1 - \frac{1}{2}e^{-\lambda}(2 + 2\lambda + \lambda^2)$$

(3)

The table below gives the values of the power function to 2 decimal places.

λ	2.5	3.0	3.5	4.0	5.0	7.0
Power	0.46	r	0.68	s	0.88	0.97

Table 1

(c) Calculate the values of r and s.

(2)

(d) Draw a graph of the power function on the graph paper provided on page 6

(2)

(e) Find the range of values of λ for which the power of this test is greater than 0.6

(1)

Leave blank

Question 3 continued

P 4 2 9 6 0 A 0 5 2 4

Turn over

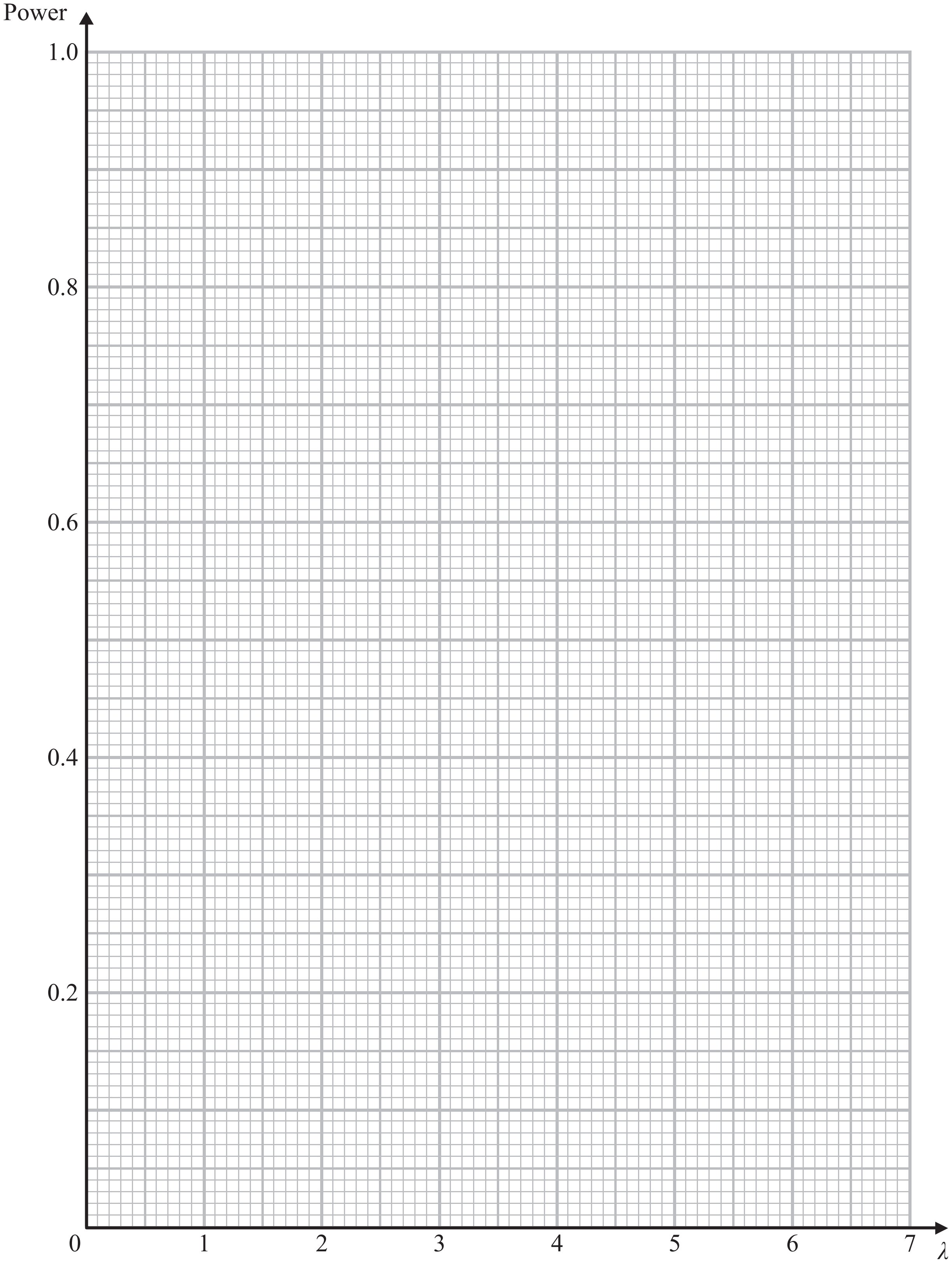
Leave blank

Question 3 continued

For your convenience Table 1 is repeated here.

λ	2.5	3.0	3.5	4.0	5.0	7.0
Power	0.46	r	0.68	s	0.88	0.97

Table 1



Leave blank

Question 3 continued

Q3

(Total 10 marks)

P 4 2 9 6 0 A 0 7 2 4

Turn over

4. A company carries out an investigation into the strengths of rods from two different suppliers, Ardo and Bards. Independent random samples of rods were taken from each supplier and the force, x kN, needed to break each rod was recorded. The company wrote the results on a piece of paper but unfortunately spilt ink on it so some of the results can not be seen.
The paper with the results on is shown below.

Ardo: 13.1 13.6 13.2 13.8 12.8 13.5 13.8

Bards: 15.3 15.5 14.1 15.4 14.2 15.4

Ardo $n_A = 7$ $\bar{x}_A = 13.4$ s

Bards $n_B = 9$ $\bar{x}_B = 14.8$ s

Pooled estimate of variance = 0.261

(a) (i) Use the data from Ardo to calculate an unbiased estimate, s_A^2, of the variance.

(ii) Hence find an unbiased estimate, s_B^2, of the variance for the sample of 9 values from Bards.

(4)

(b) Stating your hypotheses clearly, test at the 10% level of significance whether or not there is a difference in variability of strength between the rods from Ardo and the rods from Bards.
(You may assume the two samples come from independent normal distributions.)

(5)

(c) Use a 5% level of significance to test whether the mean strength of rods from Bards is more than 0.9 kN greater than the mean strength of rods from Ardo.

(6)

Leave blank

Leave blank

Question 4 continued

P 4 2 9 6 0 A 0 9 2 4

Turn over

Leave blank

5. Students studying for their Mathematics GCSE are assessed by two examination papers. A teacher believes that on average the score on paper I is more than 1 mark higher than the score on paper II. To test this belief the scores of 8 randomly selected students are recorded. The results are given in the table below.

Student	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*
Score on paper I	57	63	68	81	43	65	52	31
Score on paper II	53	62	61	78	44	64	43	29

Assuming that the scores are normally distributed and stating your hypotheses clearly, test at the 5% level of significance whether or not there is evidence to support the teacher's belief.

(8)

Leave blank

Question 5 continued

Q5

(Total 8 marks)

P 4 2 9 6 0 A 0 1 3 2 4

Turn over

Leave blank

6. A machine fills bottles with water. The amount of water in each bottle is normally distributed. To check the machine is working properly, a random sample of 12 bottles is selected and the amount of water, in ml, in each bottle is recorded. Unbiased estimates for the mean and variance are

$$\hat{\mu} = 502 \qquad s^2 = 5.6$$

Stating your hypotheses clearly, test at the 1% level of significance

(a) whether or not the mean amount of water in a bottle is more than 500 ml,

(5)

(b) whether or not the standard deviation of the amount of water in a bottle is less than 3 ml.

(5)

Leave blank

7. A machine produces bricks. The lengths, x mm, of the bricks are distributed $N(\mu, 2^2)$. At the start of each week a random sample of n bricks is taken to check the machine is working correctly.
A test is then carried out at the 1% level of significance with

$$H_0 : \mu = 202 \quad \text{and} \quad H_1 : \mu < 202$$

(a) Find, in terms of n, the critical region of the test.

(3)

The probability of a type II error, when $\mu = 200$, is less than 0.05

(b) Find the minimum value of n.

(6)

Leave blank

Question 7 continued

P 4 2 9 6 0 A 0 1 7 2 4

Turn over

Leave blank

8. A random sample $W_1, W_2 \ldots, W_n$ is taken from a distribution with mean μ and variance σ^2

(a) Write down $\mathrm{E}\left(\sum_{i=1}^{n} W_i\right)$ and show that $\mathrm{E}\left(\sum_{i=1}^{n} W_i^2\right) = n(\sigma^2 + \mu^2)$

(4)

An estimator for μ is

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} W_i$$

(b) Show that $\bar{X}$ is a consistent estimator for μ.

(3)

An estimator of σ^2 is

$$U = \frac{1}{n}\sum_{i=1}^{n} W_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} W_i\right)^2$$

(c) Find the bias of U.

(4)

(d) Write down an unbiased estimator of σ^2 in the form kU, where k is in terms of n.

(1)

Leave blank

Question 8 continued

P 4 2 9 6 0 A 0 2 1 2 4

Centre No.						Paper Reference	Surname	Initial(s)
Candidate No.						6 6 8 6 / 0 1	Signature	

Paper Reference(s)

6686/01

Edexcel GCE

Statistics S4

Advanced/Advanced Subsidiary

Friday 21 June 2013 – Morning

Time: 1 hour 30 minutes

Examiner's use only

Team Leader's use only

Question Number	Leave Blank
1	
2	
3	
4	
5	
6	
Total	

Materials required for examination	**Items included with question papers**
Mathematical Formulae (Pink)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation or symbolic differentiation/integration, or have retrievable mathematical formulae stored in them.

Instructions to Candidates

In the boxes above, write your centre number, candidate number, your surname, initials and signature.
Check that you have the correct question paper.
Answer ALL the questions.
You must write your answer to each question in the space following the question.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
The marks for individual questions and the parts of questions are shown in round brackets: e.g. **(2)**.
There are 6 questions in this question paper. The total mark for this paper is 75.
There are 20 pages in this question paper. Any blank pages are indicated.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You should show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

Printer's Log. No.
P41806A
W850/R6686/57570 5/5/5/5/4/5/5/

P 4 1 8 0 6 A 0 1 2 0

Turn over

PEARSON

Leave blank

1. George owns a garage and he records the mileage of cars, x thousands of miles, between services. The results from a random sample of 10 cars are summarised below.

$$\sum x = 113.4 \qquad \sum x^2 = 1414.08$$

The mileage of cars between services is normally distributed and George believes that the standard deviation is 2.4 thousand miles.

Stating your hypotheses clearly, test, at the 5% level of significance, whether or not these data support George's belief.

(7)

Leave blank

2. Every 6 months some engineers are tested to see if their times, in minutes, to assemble a particular component have changed. The times taken to assemble the component are normally distributed. A random sample of 8 engineers was chosen and their times to assemble the component were recorded in January and in July. The data are given in the table below.

Engineer	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*
January	17	19	22	26	15	28	18	21
July	19	18	25	24	17	25	16	19

(a) Calculate a 95% confidence interval for the mean difference in times.

(7)

(b) Use your confidence interval to state, giving a reason, whether or not there is evidence of a change in the mean time to assemble a component. State your hypotheses clearly.

(3)

P 4 1 8 0 6 A 0 4 2 0

Leave blank

3. An archaeologist is studying the compression strength of bricks at some ancient European sites. He took random samples from two sites A and B and recorded the compression strength of these bricks in appropriate units. The results are summarised below.

Site	Sample size (n)	Sample mean ($\bar{x}$)	Standard deviation (s)
A	7	8.43	4.24
B	13	14.31	4.37

It can be assumed that the compression strength of bricks is normally distributed.

(a) Test, at the 2% level of significance, whether or not there is evidence of a difference in the variances of compression strength of the bricks between these two sites. State your hypotheses clearly.

(5)

Site A is older than site B and the archaeologist claims that the mean compression strength of the bricks was greater at the younger site.

(b) Stating your hypotheses clearly and using a 1% level of significance, test the archaeologist's claim.

(6)

(c) Explain briefly the importance of the test in part (a) to the test in part (b).

(1)

Leave blank

Question 3 continued

P 4 1 8 0 6 A 0 7 2 0

Turn over

Leave blank

4. A random sample of size 2, X_1 and X_2, is taken from the random variable X which has a continuous uniform distribution over the interval $[-a, 2a]$, $a > 0$

(a) Show that $\overline{X} = \dfrac{X_1 + X_2}{2}$ is a biased estimator of a and find the bias. **(3)**

The random variable $Y = k\overline{X}$ is an unbiased estimator of a.

(b) Write down the value of the constant k. **(1)**

(c) Find Var(Y). **(4)**

The random variable M is the maximum of X_1 and X_2

The probability density function, $m(x)$, of M is given by

$$m(x) = \begin{cases} \dfrac{2(x+a)}{9a^2} & -a \leqslant x \leqslant 2a \\ 0 & \text{otherwise} \end{cases}$$

(d) Show that M is an unbiased estimator of a. **(4)**

Given that $E(M^2) = \dfrac{3}{2}a^2$

(e) find Var(M). **(1)**

(f) State, giving a reason, whether you would use Y or M as an estimator of a. **(2)**

A random sample of two values of X are 5 and –1

(g) Use your answer to part (f) to estimate a. **(1)**

Leave blank

Question 4 continued

Leave blank

5. Water is tested at various stages during a purification process by an environmental scientist. A certain organism occurs randomly in the water at a rate of λ every 10 ml. The scientist selects a random sample of 20 ml of water to check whether there is evidence that λ is greater than 1. The criterion the scientist uses for rejecting the hypothesis that $\lambda = 1$ is that there are 4 or more organisms in the sample of 20 ml.

(a) Find the size of the test.

(2)

(b) When $\lambda = 2.5$ find P(Type II error).

(2)

A statistician suggests using an alternative test. The statistician's test involves taking a random sample of 10 ml and rejecting the hypothesis that $\lambda = 1$ if 2 or more organisms are present but accepting the hypothesis if no organisms are in the sample. If only 1 organism is found then a second random sample of 10 ml is taken and the hypothesis is rejected if 2 or more organisms are present, otherwise the hypothesis is accepted.

(c) Show that the power of the statistician's test is given by

$$1 - e^{-\lambda} - \lambda(1 + \lambda)e^{-2\lambda}$$

(4)

Table 1 below gives some values, to 2 decimal places, of the power function of the statistician's test.

λ	1.5	2	2.5	3	3.5	4
Power	0.59	0.75	0.86	r	0.96	0.97

Table 1

(d) Find the value of r.

(1)

Question 5 continues on page 16

P41806A01420

Leave blank

Question 5 continued

Question 5 continues on the next page

P 4 1 8 0 6 A 0 1 5 2 0

Turn over

Leave blank

Question 5 continued

For your convenience Table 1 is repeated here.

λ	1.5	2	2.5	3	3.5	4
Power	0.59	0.75	0.86	r	0.96	0.97

Table 1

Figure 1 shows a graph of the power function for the scientist's test.

(e) On the same axes draw the graph of the power function for the statistician's test.

(2)

Given that it takes 20 minutes to collect and test a 20 ml sample and 15 minutes to collect and test a 10 ml sample

(f) show that the expected time of the statistician's test is slower than the scientist's test for $\lambda e^{-\lambda} > \frac{1}{3}$

(4)

(g) By considering the times when $\lambda = 1$ and $\lambda = 2$ together with the power curves in part (e) suggest, giving a reason, which test you would use.

(2)

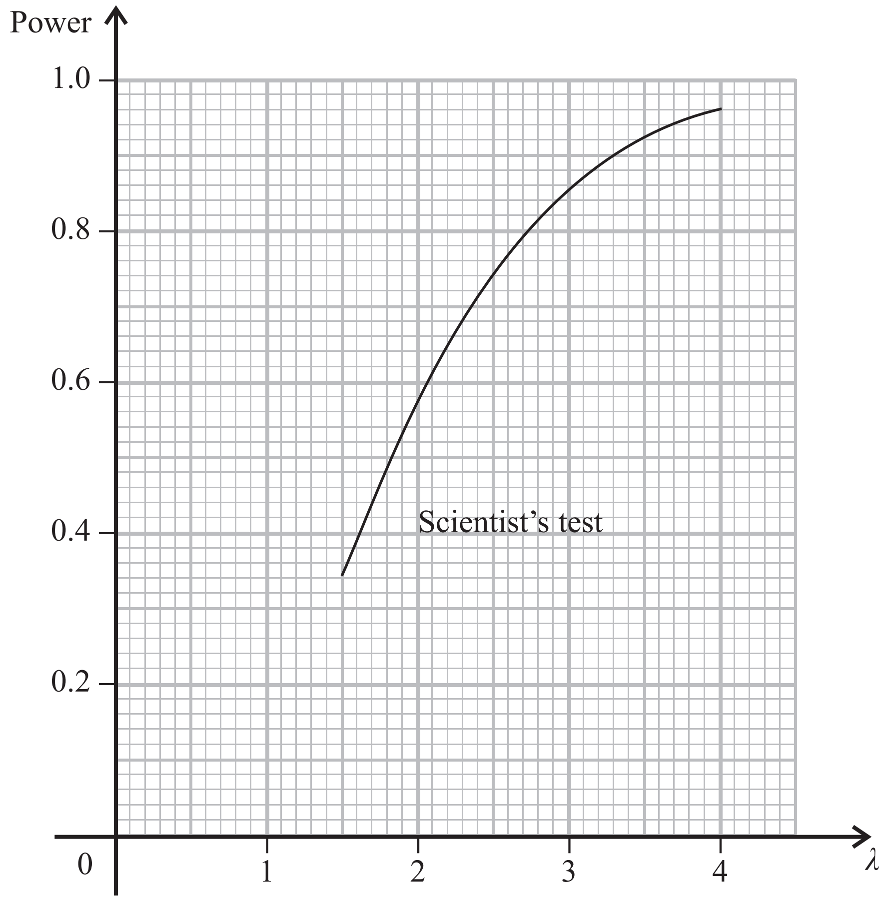


Figure 1

Leave blank

Question 5 continued

Q5

(Total 17 marks)

P 4 1 8 0 6 A 0 1 7 2 0

Turn over

Leave blank

6. The carbon content, measured in suitable units, of steel is normally distributed. Two independent random samples of steel were taken from a refining plant at different times and their carbon content recorded. The results are given below.

Sample A:	1.5	0.9	1.3	1.2		
Sample B:	0.4	0.6	0.8	0.3	0.5	0.4

(a) Stating your hypotheses clearly, carry out a suitable test, at the 10% level of significance, to show that both samples can be assumed to have come from populations with a common variance σ^2.

(7)

(b) Showing your working clearly, find the 99% confidence interval for σ^2 based on both samples.

(6)

Leave blank

Question 6 continued

Turn over