Edexcel Maths S3

Past Paper Pack

2007-2013

June 2007

Centre No.						Paper Reference							Surname	Initial(s)
Candidate No.						6	6	9	1	/	0	1	Signature	

Paper Reference(s)

6691/01

Edexcel GCE

Statistics S3

Advanced/Advanced Subsidiary

Wednesday 20 June 2007 – 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	
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 formulas 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.
N29602A
W850/R6691/57570 3/3/3/3/3

Turn over

edexcel
advancing learning, changing lives

Leave blank

1. During a village show, two judges, *P* and *Q*, had to award a mark out of 30 to some flower displays. The marks they awarded to a random sample of 8 displays were as follows:

Display	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*
Judge *P*	25	19	21	23	28	17	16	20
Judge *Q*	20	9	21	13	17	14	11	15

(a) Calculate Spearman's rank correlation coefficient for the marks awarded by the two judges.

(6)

After the show, one competitor complained about the judges. She claimed that there was no positive correlation between their marks.

(b) Stating your hypotheses clearly, test whether or not this sample provides support for the competitor's claim. Use a 5% level of significance.

(4)

Leave blank

Question 1 continued

Leave blank

2. The Director of Studies at a large college believed that students' grades in Mathematics were independent of their grades in English. She examined the results of a random group of candidates who had studied both subjects and she recorded the number of candidates in each of the 6 categories shown.

	Maths grade A or B	Maths grade C or D	Maths grade E or U
English grade A or B	25	25	10
English grade C to U	15	30	15

(a) Stating your hypotheses clearly, test the Director's belief using a 10% level of significance. You must show each step of your working.

(9)

The Head of English suggested that the Director was losing accuracy by combining the English grades C to U in one row. He suggested that the Director should split the English grades into two rows, grades C or D and grades E or U as for Mathematics.

(b) State why this might lead to problems in performing the test.

(1)

Leave blank

Question 2 continued

Leave blank

3. The time, in minutes, it takes Robert to complete the puzzle in his morning newspaper each day is normally distributed with mean 18 and standard deviation 3. After taking a holiday, Robert records the times taken to complete a random sample of 15 puzzles and he finds that the mean time is 16.5 minutes. You may assume that the holiday has not changed the standard deviation of times taken to complete the puzzle.

Stating your hypotheses clearly test, at the 5% level of significance, whether or not there has been a reduction in the mean time Robert takes to complete the puzzle.

(7)

Leave blank

4. A quality control manager regularly samples 20 items from a production line and records the number of defective items x. The results of 100 such samples are given in Table 1 below.

x	0	1	2	3	4	5	6	7 or more
Frequency	17	31	19	14	9	7	3	0

Table 1

(a) Estimate the proportion of defective items from the production line.

(2)

The manager claimed that the number of defective items in a sample of 20 can be modelled by a binomial distribution. He used the answer in part (a) to calculate the expected frequencies given in Table 2.

x	0	1	2	3	4	5	6	7 or more
Expected frequency	12.2	27.0	r	19.0	s	3.2	0.9	0.2

Table 2

(b) Find the value of r and the value of s giving your answers to 1 decimal place.

(3)

(c) Stating your hypotheses clearly, use a 5% level of significance to test the manager's claim.

(7)

(d) Explain what the analysis in part (c) tells the manager about the occurrence of defective items from this production line.

(1)

Leave blank

Question 4 continued

Leave blank

5. In a trial of diet *A* a random sample of 80 participants were asked to record their weight loss, *x* kg, after their first week of using the diet. The results are summarised by

$$\sum x = 361.6 \quad \text{and} \quad \sum x^2 = 1753.95$$

(a) Find unbiased estimates for the mean and variance of weight lost after the first week of using diet *A*.

(5)

The designers of diet *A* believe it can achieve a greater mean weight loss after the first week than a standard diet *B*. A random sample of 60 people used diet *B*. After the first week they had achieved a mean weight loss of 4.06 kg, with an unbiased estimate of variance of weight loss of 2.50 kg^2.

(b) Test, at the 5% level of significance, whether or not the mean weight loss after the first week using diet *A* is greater than that using diet *B*. State your hypotheses clearly.

(7)

(c) Explain the significance of the central limit theorem to the test in part (b).

(1)

(d) State an assumption you have made in carrying out the test in part (b).

(1)

Leave blank

Question 5 continued

Leave blank

6. A random sample of the daily sales (in £s) of a small company is taken and, using tables of the normal distribution, a 99% confidence interval for the mean daily sales is found to be

(123.5, 154.7)

Find a 95% confidence interval for the mean daily sales of the company.

(6)

Leave blank

7. A set of scaffolding poles come in two sizes, long and short. The length L of a long pole has the normal distribution $N(19.7, 0.5^2)$. The length S of a short pole has the normal distribution $N(4.9, 0.2^2)$. The random variables L and S are independent.

A long pole and a short pole are selected at random.

(a) Find the probability that the length of the long pole is more than 4 times the length of the short pole.

(7)

Four short poles are selected at random and placed end to end in a row. The random variable T represents the length of the row.

(b) Find the distribution of T.

(3)

(c) Find $P(|L - T| < 0.1)$.

(5)

Leave blank

Question 7 continued

June 2008

Centre No.						Paper Reference							Surname	Initial(s)
Candidate No.						6	6	9	1	/	0	1	Signature	

Paper Reference(s)

6691/01

Edexcel GCE

Statistics S3

Advanced/Advanced Subsidiary

Monday 16 June 2008 – 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	
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.
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.
N30984A
W850/R6691/57570 3/3/3

N 3 0 9 8 4 A 0 1 2 4

Turn over

edexcel
advancing learning, changing lives

Leave blank

1. Some biologists were studying a large group of wading birds. A random sample of 36 were measured and the wing length, x mm of each wading bird was recorded. The results are summarised as follows

$$\sum x = 6046 \qquad \sum x^2 = 1\,016\,338$$

(a) Calculate unbiased estimates of the mean and the variance of the wing lengths of these birds.

(3)

Given that the standard deviation of the wing lengths of this particular type of bird is actually 5.1 mm,

(b) find a 99% confidence interval for the mean wing length of the birds from this group.

(5)

Leave blank

2. Students in a mixed sixth form college are classified as taking courses in either Arts, Science or Humanities. A random sample of students from the college gave the following results

		Course		
		Arts	Science	Humanities
Gender	Boy	30	50	35
	Girl	40	20	42

Showing your working clearly, test, at the 1% level of significance, whether or not there is an association between gender and the type of course taken. State your hypotheses clearly.

(11)

Leave blank

Question 2 continued

Leave blank

3. The product moment correlation coefficient is denoted by r and Spearman's rank correlation coefficient is denoted by r_s.

(a) Sketch separate scatter diagrams, with five points on each diagram, to show

(i) $r = 1$,

(ii) $r_s = -1$ but $r > -1$.

(3)

Two judges rank seven collie dogs in a competition. The collie dogs are labelled *A* to *G* and the rankings are as follows

Rank	1	2	3	4	5	6	7
Judge 1	*A*	*C*	*D*	*B*	*E*	*F*	*G*
Judge 2	*A*	*B*	*D*	*C*	*E*	*G*	*F*

(b) (i) Calculate Spearman's rank correlation coefficient for these data.

(6)

(ii) Stating your hypotheses clearly, test, at the 5% level of significance, whether or not the judges are generally in agreement.

(5)

Leave blank

Question 3 continued

Leave blank

4. The weights of adult men are normally distributed with a mean of 84 kg and a standard deviation of 11 kg.

(a) Find the probability that the total weight of 4 randomly chosen adult men is less than 350 kg.

(5)

The weights of adult women are normally distributed with a mean of 62 kg and a standard deviation of 10 kg.

(b) Find the probability that the weight of a randomly chosen adult man is less than one and a half times the weight of a randomly chosen adult woman.

(6)

Leave blank

Question 4 continued

Leave blank

5. A researcher is hired by a cleaning company to survey the opinions of employees on a proposed pension scheme. The company employs 55 managers and 495 cleaners.

To collect data the researcher decides to give a questionnaire to the first 50 cleaners to leave at the end of the day.

(a) Give 2 reasons why this method is likely to produce biased results.

(2)

(b) Explain briefly how the researcher could select a sample of 50 employees using

(i) a systematic sample,

(ii) a stratified sample.

(6)

Using the random number tables in the formulae book, and starting with the top left hand corner (8) and working across, 50 random numbers between 1 and 550 inclusive were selected. The first two suitable numbers are 384 and 100.

(c) Find the next two suitable numbers.

(2)

Leave blank

Question 5 continued

Q5

(Total 10 marks)

Leave blank

6. Ten cuttings were taken from each of 100 randomly selected garden plants. The numbers of cuttings that did not grow were recorded.

The results are as follows

No. of cuttings which did not grow	0	1	2	3	4	5	6	7	8, 9 or 10
Frequency	11	21	30	20	12	3	2	1	0

(a) Show that the probability of a randomly selected cutting, from this sample, not growing is 0.223

(2)

A gardener believes that a binomial distribution might provide a good model for the number of cuttings, out of 10, that do not grow.

He uses a binomial distribution, with the probability 0.2 of a cutting not growing. The calculated expected frequencies are as follows

No. of cuttings which did not grow	0	1	2	3	4	5 or more
Expected frequency	r	26.84	s	20.13	8.81	t

(b) Find the values of r, s and t.

(4)

(c) State clearly the hypotheses required to test whether or not this binomial distribution is a suitable model for these data.

(2)

The test statistic for the test is 4.17 and the number of degrees of freedom used is 4.

(d) Explain fully why there are 4 degrees of freedom.

(2)

(e) Stating clearly the critical value used, carry out the test using a 5% level of significance.

(3)

Leave blank

Question 6 continued

Leave blank

7. A sociologist is studying how much junk food teenagers eat. A random sample of 100 female teenagers and an independent random sample of 200 male teenagers were asked to estimate what their weekly expenditure on junk food was. The results are summarised below.

	n	mean	s.d.
Female teenagers	100	£5.48	£3.62
Male teenagers	200	£6.86	£4.51

(a) Using a 5% significance level, test whether or not there is a difference in the mean amounts spent on junk food by male teenagers and female teenagers. State your hypotheses clearly.

(7)

(b) Explain briefly the importance of the central limit theorem in this problem.

(1)

Leave blank

Question 7 continued

June 2009

Centre No.						Paper Reference							Surname	Initial(s)
Candidate No.						6	6	9	1	/	0	1	Signature	

Paper Reference(s)

6691/01

Edexcel GCE

Statistics S3

Advanced/Advanced Subsidiary

Wednesday 17 June 2009 – 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
Mathematical Formulae (Orange or 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 8 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.
M34278A
W850/R6691/57570 8/4/4/3

M 3 4 2 7 8 A 0 1 2 0

Turn over

edexcel
advancing learning, changing lives

Leave blank

1. A telephone directory contains 50 000 names. A researcher wishes to select a systematic sample of 100 names from the directory.

(a) Explain in detail how the researcher should obtain such a sample.

(2)

(b) Give one advantage and one disadvantage of

(i) quota sampling,

(ii) systematic sampling.

(4)

Leave blank

2. The heights of a random sample of 10 imported orchids are measured. The mean height of the sample is found to be 20.1 cm. The heights of the orchids are normally distributed.

Given that the population standard deviation is 0.5 cm,

(a) estimate limits between which 95% of the heights of the orchids lie,

(3)

(b) find a 98% confidence interval for the mean height of the orchids.

(4)

A grower claims that the mean height of this type of orchid is 19.5 cm.

(c) Comment on the grower's claim. Give a reason for your answer.

(2)

Leave blank

3. A doctor is interested in the relationship between a person's Body Mass Index (BMI) and their level of fitness. She believes that a lower BMI leads to a greater level of fitness. She randomly selects 10 female 18 year-olds and calculates each individual's BMI. The females then run a race and the doctor records their finishing positions. The results are shown in the table.

Individual	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
BMI	17.4	21.4	18.9	24.4	19.4	20.1	22.6	18.4	25.8	28.1
Finishing position	3	5	1	9	6	4	10	2	7	8

(a) Calculate Spearman's rank correlation coefficient for these data.

(5)

(b) Stating your hypotheses clearly and using a one tailed test with a 5% level of significance, interpret your rank correlation coefficient.

(5)

(c) Give a reason to support the use of the rank correlation coefficient rather than the product moment correlation coefficient with these data.

(1)

Leave blank

4. A sample of size 8 is to be taken from a population that is normally distributed with mean 55 and standard deviation 3. Find the probability that the sample mean will be greater than 57.

(5)

Leave blank

5. The number of goals scored by a football team is recorded for 100 games. The results are summarised in Table 1 below.

Number of goals	Frequency
0	40
1	33
2	14
3	8
4	5

Table 1

(a) Calculate the mean number of goals scored per game.

(2)

The manager claimed that the number of goals scored per match follows a Poisson distribution. He used the answer in part (a) to calculate the expected frequencies given in Table 2.

Number of goals	Expected Frequency
0	34.994
1	r
2	s
3	6.752
$\geqslant 4$	2.221

Table 2

(b) Find the value of r and the value of s giving your answers to 3 decimal places.

(3)

(c) Stating your hypotheses clearly, use a 5% level of significance to test the manager's claim.

(7)

Leave blank

Question 5 continued

Leave blank

6. The lengths of a random sample of 120 limpets taken from the upper shore of a beach had a mean of 4.97 cm and a standard deviation of 0.42 cm. The lengths of a second random sample of 150 limpets taken from the lower shore of the same beach had a mean of 5.05 cm and a standard deviation of 0.67 cm.

(a) Test, using a 5% level of significance, whether or not the mean length of limpets from the upper shore is less than the mean length of limpets from the lower shore. State your hypotheses clearly.

(8)

(b) State two assumptions you made in carrying out the test in part (a).

(2)

Leave blank

7. A company produces climbing ropes. The lengths of the climbing ropes are normally distributed. A random sample of 5 ropes is taken and the length, in metres, of each rope is measured. The results are given below.

120.3 120.1 120.4 120.2 119.9

(a) Calculate unbiased estimates for the mean and the variance of the lengths of the climbing ropes produced by the company.

(5)

The lengths of climbing rope are known to have a standard deviation of 0.2 m. The company wants to make sure that there is a probability of at least 0.90 that the estimate of the population mean, based on a random sample size of n, lies within 0.05 m of its true value.

(b) Find the minimum sample size required.

(6)

Leave blank

Question 7 continued

Q7

(Total 11 marks)

Leave blank

8. The random variable A is defined as

$$A = 4X - 3Y$$

where $X \sim N(30, 3^2)$, $Y \sim N(20, 2^2)$ and X and Y are independent.

Find

(a) E(A),

(2)

(b) Var(A).

(3)

The random variables Y_1, Y_2, Y_3 and Y_4 are independent and each has the same distribution as Y. The random variable B is defined as

$$B = \sum_{i=1}^{4} Y_i$$

(c) Find P($B > A$).

(6)

Leave blank

Question 8 continued

June 2010

Centre No.						Paper Reference							Surname	Initial(s)
Candidate No.						6	6	9	1	/	0	1	Signature	

Paper Reference(s)

6691/01

Edexcel GCE

Statistics S3

Advanced/Advanced Subsidiary

Friday 18 June 2010 – 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	
7	
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 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.
N35397A
W850/R6691/57570 4/5/5/

N 3 5 3 9 7 A 0 1 2 4

Turn over

edexcel
advancing learning, changing lives

Leave blank

1. A report states that employees spend, on average, 80 minutes every working day on personal use of the Internet. A company takes a random sample of 100 employees and finds their mean personal Internet use is 83 minutes with a standard deviation of 15 minutes. The company's managing director claims that his employees spend more time on average on personal use of the Internet than the report states.

Test, at the 5% level of significance, the managing director's claim. State your hypotheses clearly.

(7)

Leave blank

2. Philip and James are racing car drivers. Philip's lap times, in seconds, are normally distributed with mean 90 and variance 9. James' lap times, in seconds, are normally distributed with mean 91 and variance 12. The lap times of Philip and James are independent. Before a race, they each take a qualifying lap.

(a) Find the probability that James' time for the qualifying lap is less than Philip's.

(4)

The race is made up of 60 laps. Assuming that they both start from the same starting line and lap times are independent,

(b) find the probability that Philip beats James in the race by more than 2 minutes.

(5)

Leave blank

3. A woodwork teacher measures the width, w mm, of a board. The measured width, X mm, is normally distributed with mean w mm and standard deviation 0.5 mm.

(a) Find the probability that X is within 0.6 mm of w.

(2)

The same board is measured 16 times and the results are recorded.

(b) Find the probability that the mean of these results is within 0.3 mm of w.

(4)

Given that the mean of these 16 measurements is 35.6 mm,

(c) find a 98% confidence interval for w.

(4)

Leave blank

Question 3 continued

Leave blank

4. A researcher claims that, at a river bend, the water gradually gets deeper as the distance from the inner bank increases. He measures the distance from the inner bank, b cm, and the depth of a river, s cm, at seven positions. The results are shown in the table below.

Position	*A*	*B*	*C*	*D*	*E*	*F*	*G*
Distance from inner bank b cm	100	200	300	400	500	600	700
Depth s cm	60	75	85	76	110	120	104

(a) Calculate Spearman's rank correlation coefficient between b and s.

(6)

(b) Stating your hypotheses clearly, test whether or not the data provides support for the researcher's claim. Use a 1% level of significance.

(4)

Leave blank

Question 4 continued

Q4

(Total 10 marks)

Leave blank

5. A random sample of 100 people were asked if their finances were worse, the same or better than this time last year. The sample was split according to their annual income and the results are shown in the table below.

Finances / Annual income	Worse	Same	Better
Under £15 000	14	11	9
£15 000 and above	17	20	29

Test, at the 5% level of significance, whether or not the relative state of their finances is independent of their income range. State your hypotheses and show your working clearly.

(10)

Leave blank

Question 5 continued

Leave blank

6. A total of 228 items are collected from an archaeological site. The distance from the centre of the site is recorded for each item. The results are summarised in the table below.

Distance from the centre of the site (m)	0–1	1–2	2–4	4–6	6–9	9–12
Number of items	22	15	44	37	52	58

Test, at the 5% level of significance, whether or not the data can be modelled by a continuous uniform distribution. State your hypotheses clearly.

(12)

Leave blank

Question 6 continued

Leave blank

7. A large company surveyed its staff to investigate the awareness of company policy. The company employs 6000 full time staff and 4000 part time staff.

(a) Describe how a stratified sample of 200 staff could be taken.

(3)

(b) Explain an advantage of using a stratified sample rather than a simple random sample.

(1)

A random sample of 80 full time staff and an independent random sample of 80 part time staff were given a test of policy awareness. The results are summarised in the table below.

	Mean score ($\bar{x}$)	Variance of scores (s^2)
Full time staff	52	21
Part time staff	50	19

(c) Stating your hypotheses clearly, test, at the 1% level of significance, whether or not the mean policy awareness scores for full time and part time staff are different.

(7)

(d) Explain the significance of the Central Limit Theorem to the test in part (c).

(2)

(e) State an assumption you have made in carrying out the test in part (c).

(1)

After all the staff had completed a training course the 80 full time staff and the 80 part time staff were given another test of policy awareness. The value of the test statistic z was 2.53

(f) Comment on the awareness of company policy for the full time and part time staff in light of this result. Use a 1% level of significance.

(2)

(g) Interpret your answers to part (c) and part (f).

(1)

Leave blank

Question 7 continued

June 2011

Centre No.						Paper Reference							Surname	Initial(s)
Candidate No.						6	6	9	1	/	0	1	Signature	

Paper Reference(s)

6691/01

Edexcel GCE

Statistics S3

Advanced/Advanced Subsidiary

Monday 20 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
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 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.
P38359A
W850/R6691/57570 5/5/3/3/2

P 3 8 3 5 9 A 0 1 2 4

Turn over

edexcel
advancing learning, changing lives

Leave blank

1. Explain what you understand by the Central Limit Theorem.

(3)

Q1

(Total 3 marks)

Leave blank

2. A county councillor is investigating the level of hardship, *h*, of a town and the number of calls per 100 people to the emergency services, *c*. He collects data for 7 randomly selected towns in the county. The results are shown in the table below.

Town	*A*	*B*	*C*	*D*	*E*	*F*	*G*
h	14	20	16	18	37	19	24
c	52	45	43	42	61	82	55

(a) Calculate the Spearman's rank correlation coefficient between *h* and *c*.

(6)

After collecting the data, the councillor thinks there is no correlation between hardship and the number of calls to the emergency services.

(b) Test, at the 5% level of significance, the councillor's claim. State your hypotheses clearly.

(4)

Leave blank

3. A factory manufactures batches of an electronic component. Each component is manufactured in one of three shifts. A component may have one of two types of defect, D_1 or D_2, at the end of the manufacturing process. A production manager believes that the type of defect is dependent upon the shift that manufactured the component. He examines 200 randomly selected defective components and classifies them by defect type and shift. The results are shown in the table below.

Defect type / Shift	D_1	D_2
First shift	45	18
Second shift	55	20
Third shift	50	12

Stating your hypotheses, test, at the 10% level of significance, whether or not there is evidence to support the manager's belief. Show your working clearly.

(10)

Leave blank

Question 3 continued

Leave blank

4. A shop manager wants to find out if customers spend more money when music is playing in the shop. The amount of money spent by a customer in the shop is £x. A random sample of 80 customers, who were shopping without music playing, and an independent random sample of 60 customers, who were shopping with music playing, were surveyed. The results of both samples are summarised in the table below.

	$\sum x$	$\sum x^2$	Unbiased estimate of mean	Unbiased estimate of variance
Customers shopping **without** music	5 320	392 000	$\overline{x}$	s^2
Customers shopping **with** music	4 140	312 000	69.0	446.44

(a) Find the values of $\overline{x}$ and s^2.

(5)

(b) Test, at the 5% level of significance, whether or not the mean money spent is greater when music is playing in the shop. State your hypotheses clearly.

(8)

Leave blank

Question 4 continued

Leave blank

5. The number of hurricanes per year in a particular region was recorded over 80 years. The results are summarised in Table 1 below.

No of hurricanes, h	0	1	2	3	4	5	6	7
Frequency	0	2	5	17	20	12	12	12

Table 1

(a) Write down two assumptions that will support modelling the number of hurricanes per year by a Poisson distribution.

(2)

(b) Show that the mean number of hurricanes per year from Table 1 is 4.4875

(2)

(c) Use the answer in part (b) to calculate the expected frequencies r and s given in Table 2 below to 2 decimal places.

(3)

h	0	1	2	3	4	5	6	7 or more
Expected frequency	0.90	4.04	r	13.55	s	13.65	10.21	13.39

Table 2

(d) Test, at the 5% level of significance, whether or not the data can be modelled by a Poisson distribution. State your hypotheses clearly.

(6)

Leave blank

Question 5 continued

Leave blank

6. The lifetimes of batteries from manufacturer *A* are normally distributed with mean 20 hours and standard deviation 5 hours when used in a camera.

(a) Find the mean and standard deviation of the total lifetime of a pack of 6 batteries from manufacturer *A*.

(2)

Judy uses a camera that takes one battery at a time. She takes a pack of 6 batteries from manufacturer *A* to use in her camera on holiday.

(b) Find the probability that the batteries will last for more than 110 hours on her holiday.

(2)

The lifetimes of batteries from manufacturer *B* are normally distributed with mean 35 hours and standard deviation 8 hours when used in a camera.

(c) Find the probability that the total lifetime of a pack of 6 batteries from manufacturer *A* is more than 4 times the lifetime of a single battery from manufacturer *B* when used in a camera.

(6)

Leave blank

Question 6 continued

Q6

(Total 10 marks)

Leave blank

7. Roastie's Coffee is sold in packets with a stated weight of 250 g. A supermarket manager claims that the mean weight of the packets is less than the stated weight. She weighs a random sample of 90 packets from their stock and finds that their weights have a mean of 248 g and a standard deviation of 5.4 g.

(a) Using a 5% level of significance, test whether or not the manager's claim is justified. State your hypotheses clearly.

(5)

(b) Find the 98% confidence interval for the mean weight of a packet of coffee in the supermarket's stock.

(4)

(c) State, with a reason, the action you would recommend the manager to take over the weight of a packet of Roastie's Coffee.

(2)

Roastie's Coffee company increase the mean weight of their packets to μ g and reduce the standard deviation to 3 g. The manager takes a sample of size n from these new packets. She uses the sample mean $\bar{X}$ as an estimator of μ.

(d) Find the minimum value of n such that $\mathrm{P}(|\bar{X} - \mu| < 1) \geqslant 0.98$

(5)

Leave blank

Question 7 continued

June 2012

Centre No.						Paper Reference							Surname	Initial(s)
Candidate No.						6	6	9	1	/	0	1	Signature	

Paper Reference(s)

6691/01

Edexcel GCE

Statistics S3

Advanced/Advanced Subsidiary

Thursday 21 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	
7	
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 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.
P40472A
W850/R6684/57570 5/5

P 4 0 4 7 2 A 0 1 2 4

Turn over

PEARSON

Leave blank

1. Interviews for a job are carried out by two managers. Candidates are given a score by each manager and the results for a random sample of 8 candidates are shown in the table below.

Candidate	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*
Manager *X*	62	56	87	54	65	15	12	10
Manager *Y*	54	47	71	50	49	25	30	44

(a) Calculate Spearman's rank correlation coefficient for these data.

(5)

(b) Test, at the 5% level of significance, whether there is agreement between the rankings awarded by each manager. State your hypotheses clearly.

(5)

Manager *Y* later discovered he had miscopied his score for candidate *D* and it should be 54.

(c) Without carrying out any further calculations, explain how you would calculate Spearman's rank correlation in this case.

(2)

Leave blank

2. A lake contains 3 species of fish. There are estimated to be 1400 trout, 600 bass and 450 pike in the lake. A survey of the health of the fish in the lake is carried out and a sample of 30 fish is chosen.

(a) Give a reason why stratified random sampling cannot be used.

(1)

(b) State an appropriate sampling method for the survey.

(1)

(c) Give one advantage and one disadvantage of this sampling method.

(2)

(d) Explain how this sampling method could be used to select the sample of 30 fish. You must show your working.

(4)

Leave blank

3. (a) Explain what you understand by the Central Limit Theorem.

(2)

A garage services hire cars on behalf of a hire company. The garage knows that the lifetime of the brake pads has a standard deviation of 5000 miles. The garage records the lifetimes, x miles, of the brake pads it has replaced. The garage takes a random sample of 100 brake pads and finds that $\sum x = 1\,740\,000$

(b) Find a 95% confidence interval for the mean lifetime of a brake pad.

(5)

(c) Explain the relevance of the Central Limit Theorem in part (b).

(2)

Brake pads are made to be changed every 20 000 miles on average.
The hire car company complain that the garage is changing the brake pads too soon.

(d) Comment on the hire company's complaint. Give a reason for your answer.

(2)

Leave blank

Question 3 continued

Leave blank

4. Two breeds of chicken are surveyed to measure their egg yield. The results are shown in the table below.

Egg yield / Breed	Low	Medium	High
Leghorn	22	52	26
Cornish	14	32	4

Showing each stage of your working clearly, test, at the 5% significance level, whether or not there is an association between egg yield and breed of chicken. State your hypotheses clearly.

(10)

Leave blank

Question 4 continued

Leave blank

5. Mr Alan and Ms Burns are two Mathematics teachers teaching mixed ability groups of students in a large college. At the end of the college year all students took the same examination. A random sample of 29 of Mr Alan's students and a random sample of 26 of Ms Burns' students are chosen. The results are summarised in the table below.

	Sample Size, n	Mean, $\bar{x}$	Standard Deviation, s
Mr Alan	29	80	10
Ms Burns	26	74	15

(a) Stating your hypotheses clearly, test, at the 10% level of significance whether there is evidence that there is a difference in the mean scores of their students.

(6)

Ms Burns thinks the comparison was unfair as the examination was set by Mr Alan. She looks up a different set of examination results for these students and, although Mr Alan's sample has a higher mean, she calculates the test statistic for this new set of results to be 1.6

However, Mr Alan now claims that the mean marks of his students are higher than the mean marks of Ms Burns' students.

(b) Test Mr Alan's claim, stating the hypotheses and critical values you would use. Use a 10% level of significance.

(3)

Leave blank

Question 5 continued

Leave blank

6. A total of 100 random samples of 6 items are selected from a production line in a factory and the number of defective items in each sample is recorded. The results are summarised in the table below.

Number of defective items	0	1	2	3	4	5	6
Number of samples	6	16	20	23	17	10	8

(a) Show that the mean number of defective items per sample is 2.91

(2)

A factory manager suggests that the data can be modelled by a binomial distribution with $n = 6$. He uses the mean from the sample above and calculates expected frequencies as shown in the table below.

Number of defective items	0	1	2	3	4	5	6
Expected frequency	1.87	10.54	24.82	a	22.01	8.29	b

(b) Calculate the value of a and the value of b giving your answers to 2 decimal places.

(4)

(c) Test, at the 5% level, whether or not the binomial distribution is a suitable model for the number of defective items in samples of 6 items.
State your hypotheses clearly.

(8)

Leave blank

Question 6 continued

Leave blank

7. The heights, in cm, of the male employees in a large company follow a normal distribution with mean 177 and standard deviation 5
The heights, in cm, of the female employees follow a normal distribution with mean 163 and standard deviation 4

A male employee and a female employee are chosen at random.

(a) Find the probability that the male employee is taller than the female employee.

(5)

Six male employees and four female employees are chosen at random.

(b) Find the probability that their total height is less than 17 m.

(6)

Leave blank

Question 7 continued

Centre No.						Paper Reference							Surname	Initial(s)
Candidate No.						6	6	9	1	/	0	1 R	Signature	

Paper Reference(s)

6691/01R

Edexcel GCE

Statistics S3

Advanced/Advanced Subsidiary

Thursday 13 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
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 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.
P42965A
W850/R6684/57570 5/5/5/

P 4 2 9 6 5 A 0 1 2 4

Turn over

PEARSON

Leave blank

1. A gym club has 400 members of which 300 are males.

Explain clearly how a stratified sample of size 60 could be taken.

(3)

Q1

(Total 3 marks)

Leave blank

2. A random sample of size n is to be taken from a population that is normally distributed with mean 40 and standard deviation 3. Find the minimum sample size such that the probability of the sample mean being greater than 42 is less than 5%.

(5)

Q2

(Total 5 marks)

Leave blank

3. The table below shows the population and the number of council employees for different towns and villages.

Town or village	Population	Number of council employees
A	211	10
B	356	2
C	1047	12
D	2463	21
E	4892	16
F	6479	25
G	6571	67
H	6573	45
I	9845	48
J	14 784	34

(a) Find, to 3 decimal places, Spearman's rank correlation coefficient between the population and the number of council employees.

(5)

(b) Use your value of Spearman's rank correlation coefficient to test for evidence of a positive correlation between the population and the number of council employees. Use a 2.5% significance level. State your hypotheses clearly.

(4)

It is suggested that a product moment correlation coefficient would be a more suitable calculation in this case. The product moment correlation coefficient for these data is 0.627 to 3 decimal places.

(c) Use the value of the product moment correlation coefficient to test for evidence of a positive correlation between the population and the number of council employees. Use a 2.5% significance level.

(2)

(d) Interpret and comment on your results from part (b) and part (c).

(2)

Leave blank

Question 3 continued

Leave blank

4. John thinks that a person's eye colour is related to their hair colour. He takes a random sample of 600 people and records their eye and hair colours. The results are shown in Table 1.

		Hair colour				
		Black	Brown	Red	Blonde	Total
Eye colour	Brown	45	125	15	58	243
	Blue	34	90	10	58	192
	Hazel	20	38	16	26	100
	Green	6	29	7	23	65
	Total	105	282	48	165	600

Table 1

John carries out a χ^2 test in order to test whether eye colour and hair colour are related. He calculates the expected frequencies shown in Table 2.

		Hair colour			
		Black	Brown	Red	Blonde
Eye colour	Brown	42.5	114.2	19.4	66.8
	Blue	33.6	90.2	15.4	52.8
	Hazel	17.5	47	8	27.5
	Green	11.4	30.6	5.2	17.9

Table 2

(a) Show how the value 47 in Table 2 has been calculated.

(1)

(b) Write down the number of degrees of freedom John should use in this χ^2 test.

(1)

Given that the value of the χ^2 statistic is 20.6, to 3 significant figures,

(c) find the smallest value of α for which the null hypothesis will be rejected at the α% level of significance.

(1)

(d) Use the data from Table 1 to test at the 5% level of significance whether or not the proportions of people in the population with black, brown, red and blonde hair are in the ratio 2:6:1:3
State your hypotheses clearly.

(9)

Leave blank

Question 4 continued

Leave blank

5. A manufacturer produces circular discs with diameter D mm, such that $D \sim N(\mu, \sigma^2)$.
A random sample of discs is taken and, using tables of the normal distribution, a 90% confidence interval for μ is found to be

(118.8, 121.2)

(a) Find a 98% confidence interval for μ.

(6)

(b) Hence write down a 98% confidence interval for the circumference of the discs.

(1)

Using three different random samples, three 98% confidence intervals for μ are to be found.

(c) Calculate the probability that all the intervals will contain μ.

(2)

Leave blank

Question 5 continued

Leave blank

6. The continuous random variable X is uniformly distributed over the interval

$$[a - 1, a + 5]$$

where a is a constant.

Fifty observations of X are taken, giving a sample mean of 17.2

(a) Use the Central Limit Theorem to find an approximate distribution for $\overline{X}$.

(3)

(b) Hence find a 95% confidence interval for a.

(4)

Leave blank

7. A farmer monitored the amount of lead in soil in a field next to a factory.
He took 100 samples of soil, randomly selected from different parts of the field, and found the mean weight of lead to be 67 mg/kg with standard deviation 25 mg/kg.
After the factory closed, the farmer took 150 samples of soil, randomly selected from different parts of the field, and found the mean weight of lead to be 60 mg/kg with standard deviation 10 mg/kg.

(a) Test at the 5% level of significance whether or not the mean weight of lead in the soil decreased after the factory closed. State your hypotheses clearly.

(7)

(b) Explain the significance of the Central Limit Theorem to the test in part (a).

(1)

(c) State an assumption you have made to carry out this test.

(1)

Leave blank

8. A farmer supplies both duck eggs and chicken eggs. The weights of duck eggs, D grams, and chicken eggs, C grams, are such that

$$D \sim \mathrm{N}(54, 1.2^2) \text{ and } C \sim \mathrm{N}(44, 0.8^2).$$

(a) Find the probability that the weights of 2 randomly selected duck eggs will differ by more than 3 g.

(6)

(b) Find the probability that the weight of a randomly selected chicken egg is less than $\frac{4}{5}$ of the weight of a randomly selected duck egg.

(5)

Eggs are packed in boxes which contain either 6 randomly selected duck eggs or 6 randomly selected chicken eggs. The weight of an empty box has distribution $\mathrm{N}\left(28, \sqrt{5}^2\right)$.

(c) Find the probability that a full box of duck eggs weighs at least 50 g more than a full box of chicken eggs.

(6)

Leave blank

Question 8 continued

Centre No.						Paper Reference							Surname	Initial(s)
Candidate No.						6	6	9	1	/	0	1	Signature	

Paper Reference(s)

6691/01

Edexcel GCE

Statistics S3

Advanced/Advanced Subsidiary

Thursday 13 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	
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 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.
P41817A
W850/R6691/57570 5/5/5/5/

P 4 1 8 1 7 A 0 1 2 0

Turn over

PEARSON

Leave blank

1. A doctor takes a random sample of 100 patients and measures their intake of saturated fats in their food and the level of cholesterol in their blood. The results are summarised in the table below.

Cholesterol level / Intake of saturated fats	High	Low
High	12	8
Low	26	54

Using a 5% level of significance, test whether or not there is an association between cholesterol level and intake of saturated fats. State your hypotheses and show your working clearly.

(10)

Leave blank

2. The table below shows the number of students per member of staff and the student satisfaction scores for 7 universities.

University	*A*	*B*	*C*	*D*	*E*	*F*	*G*
Number of students per member of staff	14.2	13.1	13.3	11.7	10.5	15.9	10.8
Student satisfaction score	4.1	4.2	3.8	4.0	3.9	4.3	3.7

(a) Calculate Spearman's rank correlation coefficient for these data.

(5)

(b) Stating your hypotheses clearly test, at the 5% level of significance, whether or not there is evidence of a correlation between the number of students per member of staff and the student satisfaction score.

(3)

Leave blank

3. A college manager wants to survey students' opinions of enrichment activities. She decides to survey the students on the courses summarised in the table below.

Course	Number of students enrolled
Leisure and Sport	420
Information Technology	337
Health and Social Care	200
Media Studies	43

Each student takes only one course.

The manager has access to the college's information system that holds full details of each of the enrolled students including name, address, telephone number and their course of study. She wants to compare the opinions of students on each course and has a generous budget to pay for the cost of the survey.

(a) Give one advantage and one disadvantage of carrying out this survey using

(i) quota sampling,

(ii) stratified sampling.

(2)

The manager decides to take a stratified sample of 100 students.

(b) Calculate the number of students to be sampled from each course.

(3)

(c) Describe how to choose students for the stratified sample.

(2)

Leave blank

4. Customers at a post office are timed to see how long they wait until being served at the counter. A random sample of 50 customers is chosen and their waiting times, x minutes, are summarised in Table 1.

Waiting time in minutes (x)	Frequency
0–3	8
3–5	12
5–6	13
6–8	9
8–12	8

Table 1

(a) Show that an estimate of $\bar{x} = 5.49$ and an estimate of $s_x^2 = 6.88$

(3)

The post office manager believes that the customers' waiting times can be modelled by a normal distribution.
Assuming the data is normally distributed, she calculates the expected frequencies for these data and some of these frequencies are shown in Table 2.

Waiting Time	$x < 3$	3–5	5–6	6–8	$x > 8$
Expected Frequency	8.56	12.73	7.56	a	b

Table 2

(b) Find the value of a and the value of b.

(3)

(c) Test, at the 5% level of significance, the manager's belief. State your hypotheses clearly.

(8)

Leave blank

Question 4 continued

Leave blank

5. Blumen is a perfume sold in bottles. The amount of perfume in each bottle is normally distributed. The amount of perfume in a large bottle has mean 50 ml and standard deviation 5 ml. The amount of perfume in a small bottle has mean 15 ml and standard deviation 3 ml.

One large and 3 small bottles of Blumen are chosen at random.

(a) Find the probability that the amount in the large bottle is less than the total amount in the 3 small bottles.

(6)

A large bottle and a small bottle of Blumen are chosen at random.

(b) Find the probability that the large bottle contains more than 3 times the amount in the small bottle.

(6)

Leave blank

Question 5 continued

Leave blank

6. Fruit-n-Veg4U Market Gardens grow tomatoes. They want to improve their yield of tomatoes by at least 1 kg per plant by buying a new variety. The variance of the yield of the old variety of plant is 0.5 kg^2 and the variance of the yield for the new variety of plant is 0.75 kg^2. A random sample of 60 plants of the old variety has a mean yield of 5.5 kg. A random sample of 70 of the new variety has a mean yield of 7 kg.

(a) Stating your hypotheses clearly test, at the 5% level of significance, whether or not there is evidence that the mean yield of the new variety is more than 1 kg greater than the mean yield of the old variety.

(9)

(b) Explain the relevance of the Central Limit Theorem to the test in part (a).

(2)

Leave blank

Question 6 continued

Q6

(Total 11 marks)

Leave blank

7. Lambs are born in a shed on Mill Farm. The birth weights, x kg, of a random sample of 8 newborn lambs are given below.

4.12 5.12 4.84 4.65 3.55 3.65 3.96 3.40

(a) Calculate unbiased estimates of the mean and variance of the birth weight of lambs born on Mill Farm.

(3)

A further random sample of 32 lambs is chosen and the unbiased estimates of the mean and variance of the birth weight of lambs from this sample are 4.55 and 0.25 respectively.

(b) Treating the combined sample of 40 lambs as a single sample, estimate the standard error of the mean.

(7)

The owner of Mill Farm researches the breed of lamb and discovers that the population of birth weights is normally distributed with standard deviation 0.67 kg.

(c) Calculate a 95% confidence interval for the mean birth weight of this breed of lamb using your combined sample mean.

(3)

Leave blank

Question 7 continued