Centre Number						Candidate Number				
Surname										
Other Names										
Candidate Signature										

For Examiner's Use	
Examiner's Initials	
Question	Mark
1	
2	
3	
4	
5	
6	
TOTAL	

General Certificate of Education
Advanced Subsidiary Examination
January 2012

Statistics SS03

Unit Statistics 3

Monday 23 January 2012 9.00 am to 10.30 am

For this paper you must have:
- the blue AQA booklet of formulae and statistical tables.

You may use a graphics calculator.

Time allowed
- 1 hour 30 minutes

Instructions
- Use black ink or black ball-point pen. Pencil should only be used for drawing.
- Fill in the boxes at the top of this page.
- Answer **all** questions.
- Write the question part reference (eg (a), (b)(i) etc) in the left-hand margin.
- You must answer the questions in the spaces provided. Do not write outside the box around each page.
- Show all necessary working; otherwise marks for method may be lost.
- Do all rough work in this book. Cross through any work that you do not want to be marked.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

Information
- The marks for questions are shown in brackets.
- The maximum mark for this paper is 75.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.
- You do not necessarily need to use all the space provided.

Answer **all** questions in the spaces provided.

1 For films released in the USA between 1998 and 2008, information is available on the body count, x, and on the box office total gross takings, \$$y$ million.

The table shows information for 10 of these films, selected at random from those which have a body count greater than 50.

		x	y
Film	**Titanic**	307	1849
	Return of the King	836	1133
	The Two Towers	468	926
	Troy	572	497
	Saving Private Ryan	255	481
	Gladiator	77	458
	The Last Samurai	558	457
	Bad Boys II	63	273
	Rambo	271	117
	We Were Soldiers	305	115

The box office total gross takings are given in \$million adjusted to 2008 figures.

(a) Calculate the value of Spearman's rank correlation coefficient between x and y.

(6 marks)

(b) Carry out a hypothesis test, at the 10% level of significance, to determine whether the value that you calculated in part **(a)** indicates a positive association between x and y.

Interpret your conclusion in context. *(4 marks)*

QUESTION PART REFERENCE

.....

.....

.....

.....

.....

Do not write outside the box

QUESTION PART REFERENCE

Do not write outside the box

2 The gross annual salaries for medical specialists working in the EU during 2003 were investigated.

The median gross annual salary for medical specialists working in the UK during 2003 was found to be £81 050 .

The gross annual salary, converted into £, for each of 9 randomly selected medical specialists working in France during 2003 was recorded as follows.

63 520 64 600 72 000 58 450 82 200 66 500 86 600 68 200 69 100

(a) Carry out a Wilcoxon signed-rank test, using the 5% significance level, to investigate whether, during 2003, the average gross annual salary for medical specialists working in the UK was greater than that for those working in France. *(8 marks)*

(b) State **one** assumption that is necessary for the test in part **(a)** to be valid. *(1 mark)*

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

3 A hospital clinic treats a large number of teenagers who have persistent abdominal pain for which no definite cause can be found. The doctor in charge of this clinic believes that reflexology therapy might help to reduce such pain. This therapy was suggested to the teenagers, who attended the clinic with their parents, as something that they might like to try, in addition to their usual treatment.

Ten teenagers tried the reflexology therapy and, after 6 weeks of the therapy, the teenagers were asked to record their pain level as 'Reduced', 'No change' or 'Increased'. The results are given in the table.

		Teenager									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Pain level	**Reduced**	✓		✓	✓	✓		✓	✓		✓
	No change						✓				
	Increased		✓							✓	

(a) (i) Name the test that you would use to investigate the doctor's belief. *(1 mark)*

(ii) Give a reason for your choice of test. *(1 mark)*

(b) Carry out the test that you named in part **(a)(i)** in order to investigate the doctor's belief. Use the 5% level of significance. *(7 marks)*

(c) Give one reason why the results of the test that you carried out in part **(b)** should be treated with caution. *(1 mark)*

QUESTION PART REFERENCE

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Do not write outside the box

QUESTION PART REFERENCE

4 A magazine editor wished to investigate the effectiveness of weight-loss diets aimed at men. Twenty-four overweight men volunteered for the investigation. These men were randomly assigned to follow one of three different weight-loss diets: A, B or C.

After 5 months, twenty of the men were still following their diets. Each of these men was measured for his percentage reduction in body weight and these results are given in the table.

Diet A	Diet B	Diet C
10	12	23
13	15	26
15	19	28
16	20	35
18	22	37
21	24	
27	25	
30		

(a) Carry out a Kruskal–Wallis test, using the 5% significance level, to investigate whether there is a difference between the average percentage reduction in body weight for the three diets. *(11 marks)*

(b) Of the four men who did not follow their diets for 5 months, three of them were following diet C. The fourth man, who was following diet B, became seriously ill and so had to withdraw from the diet.

(i) The editor expressed concern that three of the four men who did not follow their diets for 5 months were following diet C. Give a reason for her concern. *(1 mark)*

(ii) The editor stated that she wished that there had been more information regarding the serious illness of the man following diet B. Give a reason for her wish for more information. *(1 mark)*

QUESTION PART REFERENCE

.....

.....

.....

.....

.....

.....

.....

Do not write outside the box

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

5 A car insurance company conducted a survey on occurrence of accidents and age of driver.

(a) A sample of 1200 clients of this insurance company was investigated during 2009, and 200 of them were aged 17–18 years. There were also 300 clients who were aged 51 years and over in the sample. It was found that 86 clients in the sample were involved in a car accident. Of those clients involved in a car accident, 26 were aged 17–18 years and 48 were aged 19–50 years.

(i) Illustrate this information by completing **Table 1** below with the appropriate frequencies. *(2 marks)*

(ii) Test, using the 1% level of significance, whether involvement in a car accident is independent of age. *(10 marks)*

(iii) By comparing observed and expected frequencies, identify, in context, **two** important facts. *(2 marks)*

QUESTION PART REFERENCE

(a)(i)

Table 1

		Accident during 2009	No accident during 2009	Total
Age	**17–18 years**	26		
	19–50 years			
	51 years and over			300
	Total	86		1200

.....

.....

.....

.....

.....

.....

.....

Do not write outside the box

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

Question 5 continues on the next page

5 (b) **Table 2** summarises the information gathered during 2009 regarding the sizes of the insurance claims for the 86 accidents. In this investigation, drivers are categorised as either aged 17–30 years or 31 years and over. The claims are categorised as under £2000, £2000–£4000 or over £4000.

Table 2

		Size of claim		
		Under £2000	**£2000–£4000**	**Over £4000**
Age	**17–30 years**	26	19	9
	31 years and over	24	6	2

(i) Use the information in **Table 2** to construct a contingency table of expected frequencies that could be used to investigate whether there is an association between size of claim and age. *(2 marks)*

(ii) Carry out a test, using the 1% level of significance, to investigate whether size of claim is associated with age. You should consider pooling categories before you evaluate the test statistic. *(7 marks)*

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

6 A study of the effect of caffeine on muscle metabolism used 14 male volunteers who each underwent an arm-exercise test. Seven of the men were randomly selected to take a capsule containing pure caffeine one hour before the test. The other seven men received a placebo capsule that had no active ingredient. During each arm-exercise test, the subject's respiratory exchange ratio (RER) was measured.

The question of interest to the experimenter was whether, on average, caffeine affects RER.

The results were as follows.

RER	
Men who have taken caffeine	**Men who have not taken caffeine**
106	119
99	105
96	101
94	100
93	97
89	95
88	94

Carry out the appropriate distribution-free test, using the 5% level of significance, to determine whether there is evidence of a difference, on average, between RER for men who have taken caffeine and that for men who have not. *(10 marks)*

QUESTION PART REFERENCE

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Do not write outside the box

QUESTION PART REFERENCE

END OF QUESTIONS

Do not write outside the box

There are no questions printed on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED

Do not write outside the box

There are no questions printed on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED

Do not write outside the box

There are no questions printed on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED