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
June 2012

Statistics SS03

Unit Statistics 3

Thursday 31 May 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 each question in the space provided for that question. If you require extra space, use an AQA supplementary answer book; do **not** use the space provided for a different question.
- 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.

Answer each question in the space provided for that question.

1 After an intensive advertising campaign for a particular brand of breakfast cereal, a market research company asked 12 households who regularly purchase breakfast cereal to record the number of packets of this brand that they had purchased.

(a) The numbers of packets purchased per household during the three months following the campaign are

4 8 7 12 11 10 6 7 0 6 0 15

The median number of packets of this particular brand of breakfast cereal purchased per household over a similar time period before the campaign was 5 .

Carry out a sign test, using the 10% level of significance, to investigate whether the median number of packets purchased per household increased following the campaign. *(6 marks)*

(b) Name another distribution-free test that could be used to investigate whether the median number of packets purchased per household increased following the campaign. *(1 mark)*

QUESTION PART REFERENCE

Answer space for question 1

QUESTION PART REFERENCE

Answer space for question 1

2 Records of claims following hurricanes in the USA are kept by an insurance company. The number of injuries, x thousands, and the total approximate cost at today's prices, y million dollars, for each of 10 hurricanes that occurred between 1930 and 2010 are given in the table.

x	415	362	190	150	47	24	290	390	95	69
y	3200	1390	680	260	200	180	1050	2430	290	210

The 10 selected hurricanes may be regarded as a random sample.

(a) Calculate the value of Spearman's rank correlation coefficient between x and y. *(6 marks)*

(b) Carry out a hypothesis test, at the 1% level of significance, to determine whether your calculated value in part **(a)** indicates a positive association between x and y.

Interpret your conclusion in the context of the question. *(4 marks)*

(c) (i) Plot the data on **Figure 1**, printed on the page opposite. *(2 marks)*

(ii) With reference to your scatter diagram, give a reason why the product moment correlation coefficient is **not** an appropriate measure of association for the given data. *(1 mark)*

QUESTION PART REFERENCE

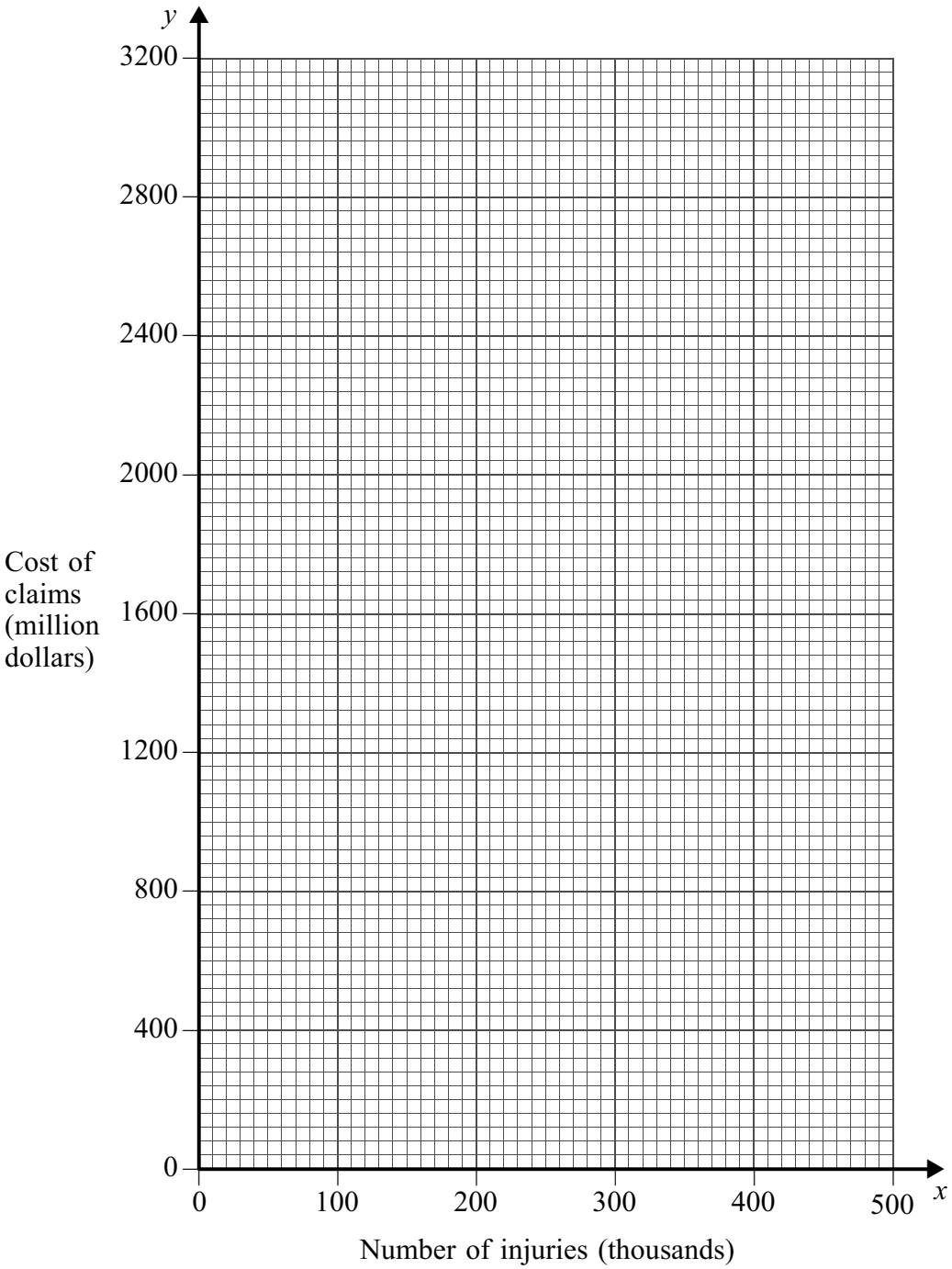
Answer space for question 2

QUESTION PART REFERENCE

(c)(i)

Answer space for question 2

Figure 1



.....

.....

.....

.....

.....

QUESTION PART REFERENCE

Answer space for question 2

QUESTION PART REFERENCE

Answer space for question 2

3 It is believed that a person's skin reaction when exposed to ultra-violet light is related to their eye colour. As part of an investigation into this belief, a sample of 120 people was exposed to a standard safe level of ultra-violet light and their skin reactions were recorded as None, Mild or Strong. Each person's eye colour was also recorded as Blue, Brown or Other.

The results of the investigation are summarised in the table.

		Eye colour		
		Blue	**Brown**	**Other**
Skin reaction	**None**	13	17	10
	Mild	25	15	5
	Strong	20	4	11
	Total	58	36	26

You may assume that the 120 people represent a random sample.

(a) Test, using the 5% level of significance, whether skin reaction when exposed to ultra-violet light is independent of eye colour. *(10 marks)*

(b) By comparing observed and expected frequencies, make **two** comments about the association, if any, between skin reaction when exposed to ultra-violet light and eye colour. *(2 marks)*

QUESTION PART REFERENCE

Answer space for question 3

QUESTION PART REFERENCE

Answer space for question 3

QUESTION PART REFERENCE

Answer space for question 3

QUESTION PART REFERENCE

Answer space for question 3

4 The coaches at a large rugby club were interested in whether some visiting teams were more aggressive than others.

For each of four home games against different visiting teams, A, B, C and D, five players were selected at random from the club's team. The total number of fresh injuries, minor or serious, suffered by each of these five players was recorded for each game.

The table below contains the results and some of the rank values for these results. A rank of 1 indicates the smallest number of injuries suffered in a game.

(a) Enter the missing ranks into the table. *(2 marks)*

(b) Carry out a Kruskal–Wallis test, using the 1% level of significance, to determine whether there is evidence of a difference, on average, between the number of injuries suffered by the club's team in a home game when playing each of the four visiting teams. *(8 marks)*

(c) Explain the meaning of a Type II error in the context of the test that you carried out in part **(b)**. *(2 marks)*

QUESTION PART REFERENCE

Answer space for question 4

(a)

Visiting team							
A		**B**		**C**		**D**	
Injuries	**Rank**	**Injuries**	**Rank**	**Injuries**	**Rank**	**Injuries**	**Rank**
2	3	0	1	8	8	5	5
4	4	1	2	13		14	
7	7	6	6	15		16	
9	9	12		17		20	
11	10	13		18		21	

.....

.....

.....

.....

.....

.....

.....

.....

QUESTION PART REFERENCE

Answer space for question 4

QUESTION PART REFERENCE

Answer space for question 4

QUESTION PART REFERENCE

Answer space for question 4

5 Two different skills tests, Test A and Test B, are used regularly to evaluate the suitability of applicants for employment as machine operators at a large engineering company. The company's director wishes to compare Test A and Test B.

Each test comprises 10 different tasks, and it is decided that the 10 tasks from each test will be combined in a random order to produce one combined test comprising all 20 tasks.

This combined test is given to nine applicants. For each applicant, the scores for the two separate sets of 10 tasks are totalled to give a score for those tasks originally from Test A and a score for those tasks originally from Test B.

These scores, as percentages, are given in the table.

Applicant / Score	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Test A	48	73	65	53	23	78	47	91	49
Test B	44	74	82	64	38	81	49	81	62

The nine applicants may be regarded as a random sample.

(a) Calculate values for the mean and the standard deviation of the scores for:

(i) Test A;

(ii) Test B. *(3 marks)*

(b) Find the value of the product moment correlation coefficient between the scores for the two tests. *(3 marks)*

(c) (i) Carry out a Wilcoxon signed-rank test, at the 5% level of significance, to investigate whether there is any difference in the mean scores for the two tests. *(8 marks)*

(ii) State the assumption, regarding the distribution of scores, that was necessary for the test in part **(c)(i)** to be valid. *(1 mark)*

(d) With reference to your findings in parts **(a)**, **(b)** and **(c)(i)**, compare Test A and Test B. *(3 marks)*

(e) Dexter, the Head of Recruitment for the engineering company, felt that the suggested combined skills test was too long for applicants. He suggested instead that Test A and Test B should be given to two separate groups of applicants and then the scores of the group taking Test A could be compared with the scores of the group taking Test B.

Give **two** reasons why using the combined skills test was preferable to Dexter's suggestion when comparing Test A and Test B. *(2 marks)*

QUESTION PART REFERENCE

Answer space for question 5

QUESTION PART REFERENCE

Answer space for question 5

QUESTION PART REFERENCE

Answer space for question 5

6 Jemma believed that the time taken for her to travel to work by bus differed according to the time of day. She worked different shifts on different days of the week, and travelled to work each day either for the morning shift, M, starting at 07.00 or for the afternoon shift, A, starting at 15.00 .

She recorded the bus journey times for her travel to work for seven randomly selected morning journeys and seven randomly selected afternoon journeys during a four-month period.

The results in the table are the times taken, x minutes, for the 14 journeys.

x	19.2	21.3	22.4	26.8	22.3	19.6	20.2	22.5	24.8	21.7	24.6	21.7	28.4	26.2
Shift	M	A	M	M	A	A	A	M	M	A	M	A	M	A

Carry out a non-parametric test, using the 5% level of significance, in order to investigate for a difference between Jemma's average journey time to work for the morning shift and that for the afternoon shift. *(11 marks)*

QUESTION PART REFERENCE

Answer space for question 6

QUESTION PART REFERENCE

Answer space for question 6

QUESTION PART REFERENCE

Answer space for question 6

QUESTION PART REFERENCE

Answer space for question 6

END OF QUESTIONS

There are no questions printed on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED