Centre Number						Candidate Number				
Surname										
Other Names										
Candidate Signature										

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

General Certificate of Education
Advanced Subsidiary Examination
June 2011

Statistics SS03

Unit Statistics 3

Monday 13 June 2011 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.

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

1 Two drugs, A and B, can each be used to reduce levels of a particular substance in the blood of adults.

Fifteen adults all had a similar high level of the substance in their blood. Of these adults, 8 took drug A for six weeks and 7 took drug B for six weeks.

At the end of the six weeks, the level of the substance in a fixed volume of each adult's blood was measured.

The results are given in the table.

Drug A	Drug B
0.43	0.44
0.51	0.49
0.53	0.59
0.55	0.62
0.60	0.65
0.65	0.67
0.69	0.68
0.71	

(a) Carry out a Mann–Whitney *U* test, at the 10% level of significance, to investigate whether there is any difference in the average level of the substance in the blood after six weeks of taking either drug A or drug B. *(10 marks)*

(b) Explain, in the context of this question, the meaning of a Type II error. *(2 marks)*

QUESTION PART REFERENCE

QUESTION PART REFERENCE

2 The student union of a large sixth form college believed that the average time per week of term that A-level students spent on individual study at home had increased because students were becoming more aware of the need to achieve high grades.

The average time per week of term that students from the college spent on individual study at home during 2009 was 10.8 hours.

A random sample of 10 students from the college was asked to record the time spent on individual study at home for three weeks during November 2010. The average time per week, in hours, for each student was recorded as follows:

17.45 14.65 12.30 11.60 7.10 15.15 16.20 7.60 6.75 8.60

(a) (i) Carry out a Wilcoxon signed-rank test to investigate whether there had been an increase in the average time per week of term spent on individual study at home. Use the 5% level of significance. *(9 marks)*

(ii) Explain why your conclusion in part **(a)(i)** might not apply to the average time per week of term spent on individual study at home for all A-level students. *(1 mark)*

(b) (i) Give **one** reason why a Wilcoxon signed-rank test might be preferred to a sign test in carrying out a test similar to the one carried out in part **(a)(i)**. *(1 mark)*

(ii) Give an example, in the context of an investigation of the time students spend on individual study at home, where a Wilcoxon signed-rank test would **not** be valid but a sign test would be valid. *(2 marks)*

QUESTION PART REFERENCE	

QUESTION PART REFERENCE

3 An investigation was carried out during 2008 into falls by the elderly. A randomly selected sample of 106 females aged 70–89 years, all of whom were on at least one medication, was obtained and, for each female, the number of falls and the number of medications taken were both recorded.

The results are summarised in **Table 1**.

Table 1

		Number of falls		
		None	**1 or 2**	**More than 2**
Number of medications taken	**1**	10	5	1
	2–4	21	26	5
	More than 4	7	18	13

(a) The investigator wished to examine whether the number of falls was associated with the number of medications taken.

(i) Find the corresponding expected frequencies and add them to the table **on the page opposite**. *(3 marks)*

(ii) Give a reason why it is necessary to pool two categories before a test for association can be carried out. *(1 mark)*

(iii) Give a reason for your choice of categories to pool. *(1 mark)*

(iv) Using the 1% significance level, examine whether the number of falls is associated with the number of medications taken. *(7 marks)*

QUESTION PART REFERENCE

.....

.....

.....

.....

.....

.....

.....

.....

.....

QUESTION PART REFERENCE

3(a)(i)

Expected Frequencies

		Number of falls		
		None	1 or 2	More than 2
Number of medications taken	1			
	2–4			
	More than 4			

Question 3 continues on the next page

3 (b) The investigation also considered the possible association between the number of falls and the number of chronic diseases suffered. Each female was identified as suffering from 0, 1, 2, or 3 or more chronic diseases.

A total of 4000 women aged 70–89 years, selected at random from general practices in 23 towns in Great Britain, were involved in the investigation. Of these 4000 women, 1200 had at least one fall during the previous twelve months.

The results are summarised in **Table 2**.

Table 2

		No falls	At least one fall
Number of chronic diseases suffered	**0**	60%	55%
	1	16%	15%
	2	14%	16%
	3 or more	10%	14%

(i) Use the information in **Table 2** to complete the contingency table **on the page opposite** with frequencies that could be analysed to investigate whether there is an association between the number of falls and the number of chronic diseases suffered. *(3 marks)*

(ii) For the contingency table in part **(b)(i)**, the expected frequencies are given in **Table 3** and the value of $\sum \frac{(O-E)^2}{E}$ is 18.4, correct to three significant figures.

Table 3

		No falls	At least one fall
Number of chronic diseases suffered	**0**	1638.0	702.0
	1	439.6	188.4
	2	408.8	175.2
	3 or more	313.6	134.4

Investigate, using the 1% level of significance, whether there is an association between number of falls and number of chronic diseases suffered. *(4 marks)*

(iii) By comparing the observed frequencies in part **(b)(i)** with the expected frequencies in part **(b)(ii)**, interpret, in context, the association, if any, between the number of falls and the number of chronic diseases suffered. *(2 marks)*

QUESTION PART REFERENCE
3(b)(i)

		No falls	At least one fall
Number of chronic diseases suffered	**0**		
	1		
	2		
	3 or more		

4 A research study was carried out during 2008 to investigate the basal metabolic rate, BMR, of men in Europe.

A sample of men, all aged between 30 years and 60 years, was obtained for this study.

Each man had his BMR and his body mass index, BMI, evaluated and his level of daily physical activity assessed.

The BMR, in megajoules per day, the BMI and the **rank** for the level of daily physical activity, DPA, for each of the 10 men involved in this study are shown in the table.

Rank 1 was assigned to the man with the highest level of daily physical activity.

Man	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**
BMR	5.44	5.74	5.81	6.03	6.33	6.94	7.50	7.78	7.78	8.70
BMI	17.2	17.6	18.2	19.4	21.6	19.8	20.9	23.1	21.3	23.5
DPA rank	8	10	6	7	9	5	1	2	3	4

The 10 men may be regarded as a random sample.

(a) Calculate the value of the product moment correlation coefficient between BMR and BMI. *(3 marks)*

(b) Carry out a hypothesis test, at the 5% level of significance, to determine whether your value of the correlation coefficient evaluated in part **(a)** indicates a positive correlation between BMR and BMI. *(5 marks)*

(c) Calculate the value of Spearman's rank correlation coefficient between BMR and DPA. *(5 marks)*

(d) Interpret your answers to parts **(a)**, **(b)** and **(c)** in context. *(2 marks)*

(e) (i) What assumption must be made if the product moment correlation coefficient is to be preferred to the Spearman's rank correlation coefficient, for the investigation of an association between BMR and BMI? *(1 mark)*

(ii) Give a reason why Spearman's rank correlation coefficient is the appropriate coefficient to use in part **(c)**. *(1 mark)*

QUESTION PART REFERENCE

.....

.....

.....

.....

QUESTION PART REFERENCE

QUESTION PART REFERENCE

QUESTION PART REFERENCE

5 A study compared three different methods for teaching a statistics course. The three methods were 'Programmed', 'Computer' and 'Control'.

The Programmed Method involved programmed instruction, the Computer Method involved using computer delivery of instructional materials and the Control Method was the normal lecture method.

Seventeen students agreed to be randomly assigned to one of the three methods. Their scores in a standard test at the end of the course are given in the table.

Programmed	Computer	Control
16	32	29
21	30	23
17	27	19
10	24	14
25	18	12
20		11

(a) Explain the purpose of randomly assigning students to one of the three methods. *(2 marks)*

(b) Carry out a distribution-free test, using the 5% significance level, to investigate whether there is a difference between the average scores in the standard test on statistics for the three methods of teaching. *(10 marks)*

QUESTION PART REFERENCE

QUESTION PART REFERENCE

END OF QUESTIONS

There are no questions printed on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED