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 2010

Statistics SS03

Unit Statistics 3

Friday 11 June 2010 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 Information was recorded about football teams in a competition.

(a) Part of this information included, for each team, the number of shots at goal that were on target and the number of shots at goal that were off target. The results for a random sample of 10 teams in this competition are given in **Table 1**.

Table 1

Team	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**
On target	51	72	61	35	46	53	39	34	46	30
Off target	50	44	40	31	28	52	38	35	44	36

Carry out a sign test, at the 5% significance level, to investigate whether, on average, teams have more shots on target than off target.

Interpret your conclusion in context. *(6 marks)*

(b) The information also included the number of goals scored. The results for the same 10 teams are given in **Table 2**.

Table 2

Team	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**
Number of shots at goal	101	116	101	66	74	105	77	69	90	66
Number of goals scored	19	18	14	12	11	9	8	8	8	7

It was claimed by a journalist that teams that have more shots at goal also score more goals.

(i) Calculate the value of Spearman's rank correlation coefficient between the number of shots at goal and the number of goals scored. *(6 marks)*

(ii) Carry out a hypothesis test, at the 1% significance level, to determine whether the value that you calculated in part **(b)(i)** indicates a positive association between the number of shots at goal and the number of goals scored. *(5 marks)*

(iii) With reference to your answers to parts **(b)(i)** and **(b)(ii)**, comment on the journalist's claim. *(2 marks)*

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

QUESTION PART REFERENCE

QUESTION PART REFERENCE

QUESTION PART REFERENCE

2 Research was carried out to investigate for a possible connection between weekly alcohol consumption and development of Type 2 diabetes. In the research report, it was stated that a sample of 1500 women, aged between 49 and 70, was studied and that 85 of these women went on to develop Type 2 diabetes.

The women were categorised according to their average level of weekly alcohol consumption. This was measured, in grams of alcohol per week, as 'less than 5', 'between 5 and 30' or 'more than 30'.

The results are summarised in the table.

		Type 2 diabetes developed	
		Yes	**No**
Average level of weekly alcohol consumption	**Less than 5**	38	382
	Between 5 and 30	12	653
	More than 30	35	380

(a) Test, at the 1% level of significance, whether the development of Type 2 diabetes is independent of the average level of weekly alcohol consumption.

Assume that the sample of women was random. *(10 marks)*

(b) A medical reviewer for a newspaper read the report and then he stated that people should increase their weekly alcohol consumption in order to decrease their chance of developing Type 2 diabetes.

Make **two** comments on his statement, referring to both the study and the sources of association, if any, identified when carrying out the test in part **(a)**. *(4 marks)*

(c) In fact, 15 000 women were involved in the research but the frequencies in the resulting contingency table had been divided by 10 in order to make the calculations simpler.

The test in part **(a)** was therefore repeated using the correct frequencies.

For this test, state:

(i) the critical value;

(ii) the value of the test statistic;

(iii) the conclusion. *(4 marks)*

QUESTION PART REFERENCE

QUESTION PART REFERENCE

QUESTION PART REFERENCE

3 In an educational study to investigate ways of teaching children to read, eighteen 6-year-old children were divided randomly into three groups. Every day during the study, the reading of each child was listened to by a teacher.

The teacher was instructed to respond as follows.

Group A: praise as much as possible and do not criticise
Group B: criticise errors and do not praise
Group C: comment as little as possible

The children were assessed at the start and at the end of the study. A measurement of the improvement in reading on a scale of 0 to 50, where 50 represented the highest level of improvement, was recorded for each child.

The results are given in the table.

Group A	Group B	Group C
47	39	39
46	38	36
43	34	33
40	31	32
37	28	30
35	26	25

Carry out a Kruskal–Wallis test, using the 5% significance level, to investigate whether there is any difference between the average improvement in reading for the three types of teacher response.

Interpret your conclusion in context. *(12 marks)*

QUESTION PART REFERENCE

.....

.....

.....

.....

.....

.....

.....

.....

.....

QUESTION PART REFERENCE

QUESTION PART REFERENCE

QUESTION PART REFERENCE

4 (a) The Mann–Whitney U test is to be used to compare two populations.

A random sample of size 8 is taken from each of the two populations.

The complete set of sixteen sample values is ranked in ascending order. It is known that no two sample values are equal. The sum of the ranks for one of the samples is denoted by T.

Find the minimum and the maximum possible values for T. *(4 marks)*

(b) A sample of 18 children, born during 1999, was obtained.

Six of these children, Group A, had older siblings only and so were the youngest in their family.

The other 12 children, Group B, were either an only child or not the youngest in their family.

The heights of all 18 children were measured and were then ranked from 1 to 18, with the smallest given rank 1.

The total, T_A, of the ranks for Group A was 31 and the total, T_B, of the ranks for Group B was 140.

(i) Carry out a Mann–Whitney U test, at the 5% level of significance, to determine whether there is evidence of a difference in heights between children who are the youngest in their family and those who are either an only child or not the youngest in their family. The sample of children may be regarded as random. *(6 marks)*

(ii) Interpret your conclusion in context. *(1 mark)*

QUESTION PART REFERENCE

.....

.....

.....

.....

.....

.....

.....

.....

.....

QUESTION PART REFERENCE

QUESTION PART REFERENCE

QUESTION PART REFERENCE

5 Trials were carried out on a new tablet designed to speed pain relief for patients with a bone disease.

(a) A randomly selected sample of patients, all of whom had been experiencing pain because of the bone disease, were given the new tablets. They each measured the waiting time for pain relief after taking a new tablet on waking in the morning.

The times recorded, in minutes, were

47 56 38 43 20 40 57 46 32 25 66 39 44

The average waiting time for pain relief after taking the existing tablet on waking in the morning is 56 minutes.

Carry out a Wilcoxon signed-rank test, at the 1% significance level, to investigate whether the average waiting time for pain relief after taking the new tablet is less than that after taking the existing tablet. *(9 marks)*

(b) (i) Give a reason why a Wilcoxon signed-rank test might be preferred to a sign test when carrying out the investigation in part **(a)**.

(ii) Under what conditions would it **not** be possible to carry out a Wilcoxon signed-rank test but it would be possible to carry out a sign test?

(iii) The standard deviation for the waiting time for pain relief after the new tablet is taken on waking in the morning is known to be 13.3 minutes.

It can be assumed that these times follow a normal distribution.

State an appropriate test to carry out in order to investigate whether or not the mean waiting time is lower for the new tablet. *(4 marks)*

QUESTION PART REFERENCE

.....

.....

.....

.....

.....

.....

.....

.....

QUESTION PART REFERENCE

QUESTION PART REFERENCE

END OF QUESTIONS