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 Level Examination
June 2010

Statistics SS04

Unit Statistics 4

Tuesday 15 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.

Do not write outside the box

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

1 The number of births announced in *The Dubworth Weekly Gleaner* may be modelled by a Poisson distribution. Last week, 24 births were announced.

(a) Calculate an approximate 99% confidence interval for the weekly mean number of births announced in *The Dubworth Weekly Gleaner*. *(4 marks)*

(b) When the current editor of *The Dubworth Weekly Gleaner* was appointed ten years ago, there was a weekly mean of 17 births announced. Comment on the editor's claim that the weekly mean number of births announced has increased since his appointment. *(2 marks)*

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

Do not write outside the box

2 Before a book is published, a proofreader is employed to check the manuscript for spelling and grammatical errors.

(a) Peter proofreads manuscripts written by Amelia. He knows, from past experience, that 0.35% of words in her manuscripts contain spelling errors. He claims that the number of words containing spelling errors on a page may be modelled by a binomial distribution with $p = 0.0035$.

(i) State one assumption which must be made for Peter's claim to be valid. *(1 mark)*

(ii) Assuming Peter's claim is valid, use a suitable distributional approximation to find the probability that a page of 200 words will include more than two words containing spelling errors. *(3 marks)*

(b) The number of grammatical errors on a page of manuscript written by Amelia may be modelled by a Poisson distribution with mean 0.75.

(i) Specify the distribution of the total number of grammatical errors in a 365-page manuscript written by Amelia. *(1 mark)*

(ii) Using a distributional approximation, find the probability that a 365-page manuscript written by Amelia contains fewer than 300 grammatical errors. *(4 marks)*

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

Do not write outside the box

3 Bags of oats, sold in a health food store, have a nominal weight of 1 kg. Jean, a trading standards inspector, instructed Howard, the store manager, to provide a random sample of ten bags. Jean obtained the weights, in grams, of these bags, with the following results.

995 1012 965 1023 1046 927 998 1012 994 983

(a) Howard claimed that there was no significant evidence that the mean weight of all bags was less than 1000 grams.

Carry out a suitable hypothesis test, at the 5% significance level, to confirm Howard's claim. Assume that the sample comes from a normal distribution.

(8 marks)

(b) Jean stated that, as the mean weight of the sample was less than 1000 grams, she still suspected that the mean weight of all bags was less than 1000 grams. State, with a reason, which type of error, Type I or Type II, Jean suspected had been made.

(2 marks)

(c) Because of her suspicions, Jean asked Howard to provide a further random sample, this time of 120 bags. The mean weight of these 120 bags was 1002.2 grams and the standard deviation was 18.1 grams.

Explain why it is unnecessary to carry out a further hypothesis test in order to conclude that there is no significant evidence that the mean weight of all bags is less than 1000 grams. *(2 marks)*

(d) How, if at all, would the reliability of the conclusion in part **(c)** be affected if it later emerged that, although the distribution was normal, the sample was not random? Justify your answer. *(2 marks)*

QUESTION PART REFERENCE

.....

.....

.....

.....

.....

.....

.....

.....

.....

Do not write outside the box

QUESTION PART REFERENCE

Do not write outside the box

4 A cooperative, set up to provide energy from local alternative sources, has 7090 members. Yoni, the secretary, is making arrangements for the annual general meeting and needs to estimate how many members will attend. She estimates that 20 per cent of members will attend and so a meeting room capable of accommodating about 1400 members will be needed.

(a) Simone, the chief executive of the cooperative, believes that fewer than 20 per cent of members will attend. She telephones a random sample of 25 members and finds that only 1 of these members plans to attend the meeting.

Use an exact binomial distribution and the 5% significance level to investigate Simone's belief. *(6 marks)*

(b) Yoni decides to take a larger sample in order to obtain a more accurate estimate of the proportion who will attend the meeting. She therefore telephones a random sample of 235 members and finds that 11 of them plan to attend the meeting.

Use a distributional approximation to calculate an approximate 90% confidence interval for the proportion of members who plan to attend the meeting. *(5 marks)*

(c) Comment on Simone's suggestion that a meeting room capable of accommodating 600 members will be adequate. *(3 marks)*

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

Do not write outside the box

5 A large hotel provides breakfast at any time between 6 am and 10 am. The manager believed that, on any day, 10 per cent of all guests arrive for breakfast before 7 am and the remaining 90 per cent arrive for breakfast after 7 am.

(a) Gofer, a member of staff, was asked to record the times at which guests arrived for breakfast on a particular day. Gofer observed that, of 234 guests, 26 arrived for breakfast before 7 am.

(i) Use a distributional approximation to examine the manager's belief. Use the 10% significance level, and assume that all guests arrive for breakfast independently and that the probability of a guest arriving for breakfast before 7 am is the same for each guest on each day. *(9 marks)*

(ii) Suggest **two** possible reasons why the assumptions made in part **(a)(i)** may not be valid. *(2 marks)*

(b) On another day, Gofer selected 9 guests at random from those staying in the hotel and recorded the number of minutes after 6 am at which they arrived for breakfast. The results were

156 240 23 118 125 0 139 114 99

(i) Calculate a 95% confidence interval for the mean number of minutes after 6 am at which guests arrive for breakfast. *(5 marks)*

(ii) State an assumption that you have had to make in order for the calculation in part **(b)(i)** to be valid. *(1 mark)*

(iii) Gofer also calculated a 95% confidence interval. He was concerned that he had made a mistake in his calculations as far fewer than 95% of the 9 values in the sample lay within the limits of the interval that he had calculated. Comment on Gofer's concern. *(2 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

Do not write outside the box

6 On the annual carnival day at a seaside town, a race is organised which requires each participant to complete a course involving canoeing, running and then cycling. Ian intends to enter this race and, after practising on the course, estimates that the time, in minutes, that he will take to complete:

- the canoeing section of the course will be an observation from a normal distribution with mean 12.5 and standard deviation 1.5 ;
- the running section of the course will be an observation from a normal distribution with mean 9.6 and standard deviation 1.3 ;
- the cycling section of the course will be an observation from a normal distribution with mean 19.0 and standard deviation 1.9 .

Assume that Ian's estimates are accurate and that his times for each section of the course are independent.

(a) (i) Specify the distribution of the total time that Ian will take to complete the course. *(3 marks)*

(ii) Find the probability that Ian will complete the course in a total time of less than 42 minutes. *(2 marks)*

(b) In the actual race, Ian and another participant, Fred, each complete the canoeing and running sections in the same total time of 22.5 minutes.

Fred's time, in minutes, for completing the cycling section will be an observation from a normal distribution with mean 19.5 and standard deviation 3.2 .

(i) Last year, the winner completed the course in 39.0 minutes. Determine which of Fred and Ian is more likely to complete the course in less time than last year's winner. *(4 marks)*

(ii) Find the probability that Ian will complete the cycling section in less time than Fred. *(4 marks)*

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

Do not write outside the box

QUESTION PART REFERENCE

END OF QUESTIONS

Copyright © 2010 AQA and its licensors. All rights reserved.