Version 1.0

**General Certificate of Education (A-level)**
**January 2013**

# Statistics SS04

# (Specification 6380)

# Statistics 4

# *Report on the Examination*

Further copies of this Report on the Examination are available from: aqa.org.uk



Set and published by the Assessment and Qualifications Alliance.

The Assessment and Qualifications Alliance (AQA) is a company limited by guarantee registered in England and Wales (company number 3644723) and a registered charity (registered charity number 1073334).
Registered address: AQA, Devas Street, Manchester M15 6EX.

## Question 1

In part (a), several students incorrectly used the binomial distribution $\mathrm{B}(50,0.8)$, but most correctly used a normal approximation. Continuity corrections were often missing or wrong (usually $49.5$ instead of $50.5$). Some students lost marks because although they used their calculator functions to do all the calculations, which is to be encouraged, they did not show the necessary intermediate working. Consequently, incorrect final answers could receive no credit as no method was shown.

In part (b), the correct answer of $0.209$ in part (a) was described as 'large' or 'small' by students and either was acceptable if followed by a sensible recommendation, ie if described as a high probability, then overbooking to the level given is unwise, but if described as a low probability, then overbooking is sensible and should continue. In general, the answer required a reference to the size of the relevant probability together with a recommendation for Chao.

## Question 2

Part (a)(i) was very well answered.

In part (a)(ii), some students were not clear enough about which statistics they were comparing and should have been clear about whether they were comparing upper or lower limits and to which value they were comparing them.

In part (b)(i), most students correctly used a binomial calculation rather than an approximate distribution. The most common error was to find $\mathrm{P}(X > 4)$ rather than $\mathrm{P}(X \geq 4)$. The conclusion in context was generally worded very well in this question where $\mathrm{H}_0$ was rejected rather than accepted.

In part (b)(ii), the comments were of average quality but most realised that Rowan's sample was rather restricted. A common error was thinking that the population to be generalised to was 'all adults' rather than 'all teenagers' probably as a result of not reading the question carefully. Similar careless reading led quite a few students to refer *only* to sample size, despite the question saying 'in addition to the sample size'. Also, many answers were not given in the context of the question.

## Question 3

In part (a), most included Safeerah's weight in a calculation. A fairly common error was to calculate the percentage increase from $75$, but this was not what was required.

In part (b), it was good to see that nearly all students estimated the standard error correctly using $s_{n-1}/\sqrt{n}$ or $s_n/\sqrt{n-1}$. The vast majority used critical values rather than $p$-values. If using critical values, occasionally a two-sided one was incorrectly used. If using $p$-values, most students, unfortunately, incorrectly used a normal distribution rather than a $t$-distribution. It was pleasing to see that nearly all conclusions were made in context but it should be noted that conclusions such as 'There is evidence that the mean is $90$' are too definite and are incorrect so score no marks. In general, in a test where we are unable to reject $\mathrm{H}_0$, students should state 'there is no evidence to doubt the hypothesis that…' or 'there is insufficient evidence to support the hypothesis $\mathrm{H}_1$ ….' in context.

## Question 4

In part (a), answers were mostly correct but often clumsily worded.

In part (b), students occasionally got the $8$ and $12$ mixed up (for example finding $\mathrm{P}(\mathrm{X} \leq 12)$ from a $\mathrm{Po}(8)$ distribution and some wrongly used $\mathrm{P}(\mathrm{X} < 8)$.

In part (c), full marks were rare because of the need to work with totals rather than means. A disappointing number of students used *t*-values rather than *z*-values in constructing the interval.

Part (d)(i) was usually correct but answers to (ii) and (iii) were generally poor. Many thought that the variance measured how far apart the insects were and this lost 4 marks. The argument used was:
'Ants live close together in colonies → not spread out much within a colony → small measure of spread → variance < mean'. Conversely, 'Beetles are solitary → spread out a lot → high measure of spread → variance > mean.

Many students did not seem to understand what a variance measures. Also, there were some failures to recognise that if events 'cluster' or 'repel each other' then they are not occurring at random so a Poisson distribution is not appropriate.

## Question 5

In part (a), many realised they had to multiply $0.12$ and $0.15$ to get $0.018$ but then frequently concluded that this meant the events were **not** independent.

In part (b)(i) most used a Poisson approximation as required. Only a few used an exact binomial or a normal approximation. A fairly common error was to find $\mathrm{P}(\mathrm{X} \geq 2)$ rather than $\mathrm{P}(\mathrm{X} > 2)$.

In part (b)(ii) hypotheses were nearly always stated correctly and a normal approximation was usually used. A test statistic with or without a continuity correction was allowed regardless of whether numbers or proportions were used. There were a few wrong continuity corrections, such as $21.5$, seen.

Students should be careful not to compare a probability with a critical value. Also, when using *p*-values, students often compared a one-tailed *p*-value with $0.01$ which, for a two-tailed test, is incorrect and the comparison should be with $0.005$.

Finally, students should note that when using *p*-values from a calculator they should ensure that sufficient information is given in their solution, so that if their answer is incorrect, it is clear where errors were made so that some marks can be awarded. The conclusions were usually well worded and in context.

## Question 6

In part (a)(i), most answers demonstrated sound understanding and only a few subtracted $0.51^2$.

In part (a)(ii)There was some confusion over units (m and cm) but most students scored full marks.

In part (b)(i) many students left their answer as a probability rather than a number as required. Some students appeared confused between the fact that the standard deviation was 3 and the 'rule' that virtually all possible heights would be expected to fall within 3 standard deviations of the mean.

In part (b)(ii) 3 or 4 marks were rare here. Most students tried to add something to 500 with many correctly getting to 680. However, only a few mentioned a normal distribution although a **distribution** of income was asked for which requires name of distribution plus parameters. The correct variance of 74.4 was very rare.

In part (b)(iii), most made a sensible decision based on expected income as required.

## Mark Ranges and Award of Grades

Grade boundaries and cumulative percentage grades are available on the Results Statistics page of the AQA Website: http://www.aqa.org.uk/over/stat.html

## Converting Marks into UMS marks

Convert raw marks into Uniform Mark Scale (UMS) marks by using the link below.

**UMS conversion calculator** www.aqa.org.uk/umsconversion