Version : 1.0 0607

ASSESSMENT and QUALIFICATIONS ALLIANCE

General Certificate of Education

Mathematics 6360

MS2A Statistics 2A

Report on the Examination

2007 examination - June series

Further copies of this Report are available to download from the AQA Website: www.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
Dr Michael Cresswell Director General.

General

It was again very pleasing to see fully correct solutions to each question. In general, candidates seemed to be well prepared for the examination. There were, however, still a few candidates who did not state hypotheses correctly and/or did not give correct conclusions in context. Lack of good algebraic skills also seemed to impair some candidate's progress, especially in questions 5 and 6.

Question 1

This question, as usual on this topic, proved to be a good source of marks for all candidates. However, there were still a minority who either did not state their hypotheses in context or omitted them completely. Simply stating "H_0: not associated" and "H_1:associated" was not sufficient. Although the vast majority of candidates realised that Yates's correction had to be used, there were those who could not apply it correctly. The most common incorrect applications were $O_i - E_i - 0.5$ or $\left(O_i - E_i\right)^2 - 0.5$, whereas $\left|O_i - E_i\right| - 0.5$ was required. Although most candidates formulated a conclusion in context, these were usually too positive in nature. Also, simply stating "Reject H_0" was not sufficient to gain full marks.

Question 2

In part (a)(i), the vast majority of candidates realised that the formula or their calculator had to be used in order to obtain the answer. However, there were some who tried to use Table 2 in the booklet provided, by averaging the values that they found under $\lambda = 3.4$ and $\lambda = 3.6$, presumably under the misapprehension that this would give them the required value for $\lambda = 3.5$. However, most candidates did use Table 2 correctly in part (a)(ii) by realising that the most efficient way to obtain the result was by using $\mathrm{P}(Y \geq 5) = 1 - \mathrm{P}(Y \leq 4)$. Some candidates did not seem to be able to use these tables (or perhaps even know of their existence) as they worked out $1 - \mathrm{P}(Y = 0, 1, 2, 3 \text{ or } 4)$ by using the formula. Although they usually managed to arrive at the correct answer, this is not the most efficient way of tackling this type of question.

In part (b)(i), candidates had to indicate that the distribution was Poisson and that $\lambda = 9.5$. It was not sufficient to simply write "$\lambda = 9.5$" or just to state "Poisson". Part (b)(ii) caused most of the problems in this question. Whilst the vast majority of candidates realised correctly that $\mathrm{P}(7 \leq T \leq 10) = \mathrm{P}(T \leq 10) - \mathrm{P}(T \leq 6)$, there were some who incorrectly used $\mathrm{P}(7 \leq T \leq 10) = \mathrm{P}(T \leq 11) - \mathrm{P}(T \leq 6)$ or who either used the wrong value of λ or were simply careless and misread the tables. The vast majority of candidates realised that, in part (b)(iii), the correct answer could be obtained by simply evaluating $(\text{their part (b)(ii)})^3$. However, there were a few candidates who incorrectly thought that $3 \times (\text{their part (b)(ii)})$ was the way forward and did not seem at all deterred when this gave them an answer greater than one.

Question 3

There were still candidates who failed to write down any hypotheses. It was essential that the both hypotheses (null and alternative) were stated as it was these that were compared when the test was conducted. The outcome of the test was that one of these hypotheses would be accepted (and that the other would be rejected) and, if they were not stated in the first place, this could not be done sensibly. For the majority of candidates who tried to state their hypotheses, there was the issue of how these should be stated. Hypotheses were often incorrectly stated as H_0: mean = 36 and H_1: mean < 36, or H_0: $\bar{x} = 36$ and H_1: $\bar{x} < 36$, or even H_0: = 36 and H_1: < 36. The forms of the two hypotheses that were acceptable, and

which were expected, were either H_0: population mean = 36 and H_1: population mean < 36, or preferably H_0: $\mu = 36$ and H_1: $\mu < 36$.

The generally-considered-good practice of sketching a distribution curve would have helped ensure that the correct critical value of $z = -2.3263$, not $z = +2.3263$, was used. Conclusions should be stated in context and should not be too positive in nature. "The number of putts has reduced from 36" was both too positive and failed to mention the mean number of putts. The fact that '(test statistic) < (critical value) $\Rightarrow$ Reject H_0' did not imply that the mean number of putts had definitely reduced from 36. The evidence suggested this was the case but there could, for example, have been a Type I error.

Question 4

The confidence interval in part (a) was usually found correctly. However, on the whole, part (b) was not answered in the way that was expected. Most candidates attempted to work out the confidence interval again using $\bar{x} = 9.6$ or 9.8 and $s = 0.0294$ by repeating the work already tested in part (a). Candidates were expected to write down their answer to part (b) from what they had already found in part (a) together with the information given in this part of the question.

Question 5

This was the worst-answered question on the paper with the vast majority of candidates gaining fewer than half marks. The fact that probability values must fall within the range $0 \leq p \leq 1$ implied that $0 \leq p < 1$ by inspection of the first three values in the table. Inspection of the fourth value gave $1 - 3p \geq 0$, implying that $p \leq \frac{1}{3}$. This was not usually explained well by the weaker candidates. In part (b), $\mathrm{E}(X)$ and $\mathrm{E}(X^2)$ were usually attempted correctly with many fully correct expressions in terms of p seen. The majority of candidates then attempted to use the correct method by stating and then using $\mathrm{Var}(X) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2$. Unfortunately, a lack of good algebraic skills let some candidates down. They were either unable to cope with $(4 - 6p)^2$, where $16 - 36p^2$ was a common wrong expression, or they could not cope with the negative signs when attempting subtractions. Thus, it was hardly surprising that the correct answer was usually seen even though the working did not really lead to the correct expression given on the question paper for $\mathrm{Var}(X)$.

Part (c)(i) asked for the value of p which gave a maximum value of $\mathrm{Var}(X)$. Differentiation methods were employed by candidates when tackling this part. In such cases it should be remembered that it is then expected that both $\frac{\mathrm{d}V}{\mathrm{d}p} = 0$ and $\frac{\mathrm{d}^2V}{\mathrm{d}p^2} < 0$ should be considered. This was not always seen. No candidate considered properties of a quadratic in order to gain the correct answer of $p = \frac{7}{36}$. In part (c)(ii), in most cases where candidates managed to find a value for the variance, they rightly went on to attempt to deduce a value for the standard deviation.

Question 6

This was the second-worst-answered question on the paper with many major errors of technique and understanding. It was only answered well by the more able candidates. It was a

major concern that so many candidates thought incorrectly that $\mathrm{E}\left(\frac{1}{X}\right)=\frac{1}{\mathrm{E}(X)}$ and $\mathrm{Var}\left(\frac{1}{X}\right)=\frac{1}{\mathrm{Var}(X)}$. These candidates wasted a lot of time evaluating $\mathrm{E}(X)$ and $\mathrm{E}(X^2)$ to no avail. There were also many examples of poor integration techniques. When $\mathrm{E}\left(\frac{1}{X}\right)=\int_0^1 \frac{1}{x}\times 3x^2\,\mathrm{d}x$ was correctly stated, some candidates then seemed unable to integrate correctly, with $\ln x \times \frac{3x^3}{3}$ or $\frac{3x^3/3}{x^2/2}$ often seen.

The fact that the variance cannot be less than zero seemed to have escaped the attention of some candidates who were apparently quite comfortable, not only in obtaining negative values for their variance, but also in using such values in part (b), where it was also evident that many candidates were let down by poor algebraic skills. It was expected that candidates would first express $\frac{5+2X}{X}$ as $\frac{5}{X}+2$ and then use $\mathrm{E}\left(\frac{5}{X}+2\right)=5\mathrm{E}\left(\frac{1}{X}\right)+2$ and $\mathrm{Var}\left(\frac{5}{X}+2\right)=5^2\,\mathrm{Var}\left(\frac{1}{X}\right)$ to obtain answers of 9.5 and 18.75 respectively. Unfortunately many candidates could not even cope with the first step. On the other hand, there were some good attempts from evaluating $\mathrm{E}\left(\frac{5+2X}{X}\right)$ and $\mathrm{E}\left(\left(\frac{5+2X}{X}\right)^2\right)$ by integration.

Coursework Component

There was still a tendency to make transcription and addition errors when totalling the scripts. The final marks should be carefully checked prior to submission to AQA and for moderation. A number of scripts had either no marking on them or were marked in pencil. Scripts should be marked in red pen and calculations should be checked for accuracy (and indicated as such on the scripts). All Candidate Record Forms must be signed by the candidate and the teacher responsible for the assessment of the script.

There were very few Centres who submitted scripts this session. Some candidates did not recognise that a *t*-test could be used for their task and some of those who did failed to discuss the necessary requirements for the *t*-distribution to be valid in the context of their task.

Mark Ranges and Award of Grades

Grade boundaries and cumulative percentage grades are available on the **Results statistics** page of the AQA Website.