Version : 1.0 0607

ASSESSMENT and QUALIFICATIONS ALLIANCE

General Certificate of Education

Mathematics 6360
Statistics 6380

MS/SS1A Statistics 1A

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

The overall performance on this paper showed a slight improvement over that seen on the June 2006 paper. As a result, it was pleasing to note that almost 25% of candidates were able to achieve 50 or more (raw) marks. On the other hand, almost 15% of candidates were unable to score at least 20 (raw) marks.

Weaker candidates tended to gain the majority of their marks from one or two questions whereas strong candidates were able to attempt most, if not all, questions satisfactorily often only losing marks in parts requiring comment or interpretation. Most candidates used the statistical functions on their calculators to answer, to sufficient accuracy, Questions 1 and 5 but were often less successful in answering Question 4(b)(iii). In the main, candidates made appropriate references to Tables 1, 3 and 4 in answering Questions 3 and 6.

Question 1

As intended, this question proved to be a 'confidence booster' for the vast majority of candidates who usually scored the first 3 marks by quoting the correct value of r to at least the required degree of accuracy (3 significant figures) from their calculators' inbuilt function. Those candidates who used one of the formulae from the supplied booklet were also usually successful, but perhaps at some cost of time.

The awarding of full marks in part (b) was rare. Although almost all candidates put their interpretation in context by making reference to the two variables involved, far too many deemed the correlation either simply 'negative' or 'strong/fairly strong negative' rather than 'moderate/some/weak negative'.

Question 2

This question was also quite well answered by a large proportion of candidates. Whilst almost all could answer part (a)(i) correctly, only the more able candidates were capable of dealing with the conditional probabilities as required in parts (a)(ii) and (iii). Thus it was not unusual to see $\frac{6}{50}$ quoted for part (a)(ii) and/or $\frac{25}{50}$ or $\frac{25}{27}$ quoted for part (a)(iii). In answering part (b), it was disappointing to see the number of candidates who apparently ignored the (obvious) implication from the context that the selection had to be '**without replacement'** but chose instead to quote $\left(\frac{22}{50}\right)^4$ or to employ a binomial distribution!

Question 3

Most candidates were able to make a worthwhile attempt at this question with many scoring full marks and only the weakest not recognising that binomial distributions were required. By far the most common approach in part (a) was use of the formula, rather than tables, usually with complete success. Thankfully however, the vast majority of candidates then switched to the use of Table 1 in part (b), although a small minority used their calculators' inbuilt binomial cumulative distribution function.

In part (b)(i), the use of $\mathrm{P}(R \leq 15)$ for 'fewer than 15' was not uncommon but most candidates correctly equated 'more than 10' to $1 - \mathrm{P}(R \leq 10)$, although perhaps by luck at times. In part (b)(iii), 'at least 12 but at most 18' was often incorrectly determined from $\mathrm{P}(R \leq 18) - \mathrm{P}(R \leq 12)$.

Question 4

The majority of candidates were apparently unprepared for this type of question. Whilst part (b) might be considered quite challenging to the average candidate, it was expected that most candidates would be able to make significant progress in part (a). This expectation was certainly not realised as only minimal marks were often awarded. Indeed it was not that rare to award a candidate only 1 mark for the whole question; this for correctly stating the modal value. Answers stated for the range were often $0 - 15$ or $24 - 4 = 20$.

In part (a)(ii), many candidates attempted interpolation, presumably on the basis that the data were continuous. Those who recognised its discrete nature rarely helped themselves by constructing a cumulative frequency table. However they were often able to identify the median as 3. Many candidates had less success with the quartiles. Even some of those who identified 2 and 4 did not take the difference to find the IQR whilst a small minority stated that the $\text{IQR} = 72 - 24 = 48$ or that this implied the value of 3.

In part (a)(iii), it was rare to see 7 and 12 identified as the two group mid-points, although from the considerable number of stated correct answers, direct from calculators, they had been used. Candidates who used formulae often failed to identify the frequencies and so took n as 8 or even 15.

Correct answers to part (b) were extremely rare. Candidates often chose to explain in detail what each statistic measured, or stated for example: “median and IQR are not affected by extreme values or outliers”; “median and IQR are whole numbers”; “mean and standard deviation are closer together than mode and range”. Centres are encouraged to refer to a copy of the published Mark Scheme for acceptable answers.

Question 5

This question on regression was answered well by a large majority of candidates. In part (a), most used their calculators’ regression functions to quote b and a correctly to at least three significant figures although $b = -0.09$ was too common. As in Question 1, a formulae approach, though often successful, did perhaps have a time cost. Thankfully, fewer candidates than on previous papers interchanged a and b.

In part (b)(i), the majority of candidates were able to indicate that, as $b < 0$, then as temperature increased the time taken decreased. However most candidates did not identify that the magnitude of b reflected the decrease for each 1°C rise in temperature. In part (b)(ii), most candidates who identified $x = 0$ (0°C) realised that the water could be frozen although a small minority simply indicated ‘extrapolation’.

Question 6

It was most gratifying to note the many candidates scoring 14 or more marks (out of 16) on this question. In part (a)(i), ‘corrections to 40’ were thankfully very rare with the result that most candidates scored full marks. Again in part (a)(ii), full marks was the norm although a small minority of candidates failed to allow for the negative sign of 1.6 and so evaluated $0.94520 - 0.65542$ (loss of all 3 marks). The most common mistake in part (a)(iii), seen even on the best scripts, was to equate $\frac{x - 38}{5}$ to $+0.6745$ instead of -0.6745 (loss of 1 mark).

Despite comments on previous Examiner Reports, there still remain candidates who are apparently unaware of Table 4.

There were many correct answers given for the confidence interval in part (b)(i). The most common error was the use of an incorrect z-value, often 2.0537 (Table 4 yet again). Thankfully

very few candidates omitted to divide the standard deviation by $\sqrt{40}$. In part (b)(i), candidates had to first identify that $2.5 \times 38 = 95$ before any marks were awarded. Most candidates did indeed do this but a few evaluated $\frac{107}{2.5}$. Having obtained the value of 95 candidates then had to compare it against their confidence interval, not 107, in order to gain marks for a (follow-through) correct conclusion. Thus it was pleasing to see the many instances of candidates scoring all 3 marks.

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 was some good work seen, but there was a tendency for the highest scoring scripts to be over-marked. This was usually caused by the discussion of the sampling being too brief along with a lack of depth in the interpretation strand. In tasks involving Confidence Intervals, a clear description of what the interval is (in context) and a sensible comparison of any intervals (looking for any overlap etc) is expected for the highest marks. There was still some confusion over the Central Limit Theorem, many candidates thinking they had to assume that their sample or population (or both) were Normally distributed. Worryingly this was not always challenged on the scripts.

Mark Ranges and Award of Grades

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