Version : 1.0 0607

ASSESSMENT and QUALIFICATIONS ALLIANCE

General Certificate of Education

Mathematics 6360
Statistics 6380

MS/SS1B Statistics 1B

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 paper was found to be accessible by most candidates with only a small minority scoring very low marks. As on previous papers, able candidates performed well on parts requiring calculations but on this paper they were noticeably less successful when comments and interpretations were required. This, together with the generally unexpectedly poor responses to Question 4, resulted in a smaller proportion of candidates achieving the higher marks.

The majority of candidates used the statistical functions on their calculators to answer, usually 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 and 3, but not always Table 4, in answering Questions 3, 6 and 7.

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. A small minority of candidates quoted their answer to only 2 significant figures and/or, through carelessness or misunderstanding, omitted the minus sign. Those candidates who used one of the formulae from the supplied booklet were 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. Those candidates who extracted the necessary information from the table rather than attempting to apply formulae invariably had the greater success. Whilst almost all could answer parts (a)(i), (ii) & (iii) correctly, only the more able candidates were capable of dealing with the conditional probabilities as required in parts (a)(iv) & (v). Thus it was not unusual to see $\frac{6}{50}$ quoted for part (a)(iv) and/or $\frac{25}{50}$ or $\frac{25}{27}$ quoted for part (a)(v).

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

A large majority of candidates scored full marks on this question. In part (a), a minority of candidates used an incorrect z-value, usually 1.6449, or, more critically, omitted to divide by $\sqrt{50}$. An 'adjustment' to the value of 25.1 for the sample standard deviation was not expected and was rarely, if ever, seen. Similarly, use of the t-distribution (Statistics 2) was very rare but not penalised when applied correctly. Almost all candidates stated a valid reason in part (b), usually referring to 'selection by size of potato'. Unacceptable reasons usually referred to a small sample size.

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 although, even here, an answer of 24 was not that unusual. 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. Even those candidates who found the correct value for the mean sometimes had insurmountable difficulties in finding the standard deviation.

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. However, in answering part (a), many candidates failed to indicate clearly that the 'time taken depended on temperature' or an equivalent statement. The majority of candidates stated, for no reward, that "Bob set the temperatures" or that "It is the independent variable".

In part (b), most candidates 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 (c)(i), the majority of candidates were able to indicate that, as $b < 0$, then as temperature increased the time taken decreased, although confusion with descriptions of negative correlation was too prevalent. Most candidates failed to indicate that the magnitude of b reflected the decrease for each $1°\text{C}$ rise in temperature. In part (c)(ii), most candidates who identified $x = 0$ $(0°\text{C})$ realised that the water could be frozen. A small minority thought that a represented $y = 0$ (zero minutes) or simply indicated 'extrapolation'.

In part (d)(i), most candidates substituted correctly the value of $x = 30$ into their equations, although the use of $x = 60$ was sadly not that uncommon. A significant proportion of candidates used the mid-point of the values given in the table for $x = 28$ and $x = 32$ to score 1 of the available 2 marks. Similarly in part (d)(ii), many candidates substituted correctly the value of $x = 75$ into their equations to obtain a negative result. Most, but certainly not all, recognised

why this must be invalid. A significant minority of candidates made an inappropriate case based solely on 'extrapolation' often almost repeating their answers to part (c)(ii).

Question 6

This proved to be a good source of marks for many candidates with the more able often scoring full marks. Answers to part (a)(i) were usually correct from using Table 1 in the booklet provided. The minority using the formula often calculated $\mathrm{P}(T = 3)$. Part (a)(ii) caused some difficulties. Whilst almost all candidates attempted a subtraction of values from Table 1, many had difficulty in translating 'on more than 10 days but fewer than 20 days' to $\mathrm{P}(R \leq 19) - \mathrm{P}(R \leq 10)$ with most errors involving the use of $\mathrm{P}(R \leq 20)$ and/or $\mathrm{P}(R \leq 9)$.

Knowledge of the relevant formulae needed to answer part (b)(i) was much improved. Failure to take the square root of 1.2 was the usual reason for forfeiting marks. As expected, part (b)(ii) was answered correctly by the vast majority of candidates. Answers to part (b)(iii) frequently contained comments that were valid, consistent and appropriate.

Question 7

Many candidates scored between 6 and 10 marks on this question but only the most able could accumulate more than 10 marks. In part (b), it was often the case that the only mark scored was in part (ii). Almost all candidates recognised the need to standardise with the great majority scoring full marks. Thankfully, the use of 59 or $\sqrt{20}$ was rarer than corresponding errors on previous papers. In answering part (a)(ii), most candidates made the necessary area change before subtraction to obtain the correct answer. However, a minority of candidates evaluated 0.81594 – (answer (i)) for no reward. Given similar examples on previous papers, answers to part (a)(iii) were often disappointing. Whilst the somewhat frequent use of 1.29 (presumably from Table 3) rather than the more accurate 1.2816 from Table 4 lost only 1 mark, the all-too-frequent equating of $\frac{k - 48}{20}$ to 0.9 or $\Phi(0.9)$ lost at least 3 marks. Most candidates could not provide a valid explanation in part (b)(i). In the main answers simply stated that "the standard deviation was large when compared with the mean" or even attempted to describe the practicalities of installing gas metres!

Even the few candidates who stated "negative values likely" did not justify the statement by considering, for example, the value of $\mu - 2\sigma$. Answers to part (b)(ii) usually involved "Central Limit Theorem", "n large" or "$n > 30$" for the 1 mark available. In answering part (b)(iii), the majority of candidates standardised 40 by $\frac{40 - 37}{25}$ and so scored no marks. Of those that did use $\frac{40 - 37}{25/\sqrt{35}}$ correctly, about 50% did not then perform the necessary area change.

Mark Ranges and Award of Grades

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