Version : 1.0 0608

ASSESSMENT and QUALIFICATIONS ALLIANCE

General Certificate of Education

Mathematics 6360
Statistics 6380

MS/SS1A Statistics 1A

Report on the Examination

2008 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.*

Written Component

General

Overall, this paper proved to be of similar accessibility to recent papers. As a result, almost 20% of candidates were able to achieve at least 48 (raw) marks and approximately 80% were able to score at least 24 (raw) marks.

The majority of candidates scored well in Questions 1, 2 and 3 (except for part 3(d)) and in parts of Questions 4 and 6. Whilst most candidates were able to score at least minimal marks in Question 5, high marks proved accessible only to the higher achievers.

Most candidates, as expected, used their calculators' statistical functions to maximum effect in parts (a) of Questions 1 and 3. The small minority of candidates who used their calculators' more advanced statistical functions to simply state answers in Questions 4 and 6 suffered severely when their answers were incorrect. Centres are reminded that examiners are not expected to try to deduce the reasoning behind a stated incorrect answer and also that multiple undeleted answers/solutions generally lose marks since the mark gained is the average for those answers/solutions offered. Many candidates completed the scatter diagram in ink rather than in pencil. Those who then tried to correct mistakes often lost marks for some points being unclear.

Question 1

For the great majority of candidates, this question proved a positive start and many scored full marks. Use of regression functions on calculators was the most common approach in part (a) with the result that most obtained correct values for a and b. A small minority of candidates confused a and b, so losing at least 5 of the 6 marks available for the question. There remained a significant number of candidates who spent valuable time using formulae to calculate values for a and b. Whilst most such candidates were successful, a small number calculated b^{-1} or ignored the sign for b and, of course, all used up valuable time. In answering part (b), the majority of candidates substituted $x = 21$ correctly into their equations, but a significant proportion obtained 31.85 from $\left(\frac{33.0+30.7}{2}\right)$ to score 1 mark.

Question 2

This was a very accessible question for candidates of all abilities. Many candidates scored full marks and the awarding of fewer than 5 marks was rare indeed. Candidates who extracted the necessary information from the table generally had more success than those attempting the use of probability formulae. Few candidates failed to score the 3 marks available for parts (a) and (b). In part (c), common errors were to quote answers of $0.4 + 0.7 = 1.1$, $0.4 \times 0.7 = 0.28$ or $0.4 + 0.7 - 0.28 = 0.82$ which appeared to indicate, as on previous papers, a lack of understanding of the addition law for two (non-mutually exclusive and dependent) events; something that needs attention in future. Many candidates coped well with the conditional probability in part (d). When marks were lost, it was usually for quoting $\frac{42}{400}$, $\frac{120}{400}$ or $\frac{42}{70}$.

Question 3

Many candidates scored the first 8 marks in this question. Those candidates who used the correlation function on their calculators almost invariably scored full marks in part (a) though a small minority lost 1 or 2 marks through quoting the answer to less than 3 significant figures, such as 0.81 or 0.8. Candidates who calculated r using a formula often did so with good understanding but sometimes less than accurately. In part (b), most candidates recognised that

there was 'positive correlation' but some did not attach an adjective to indicate the strength. Almost all candidates included in their statements a reference to the context usually by mentioning 'length' and 'width'. Whilst most scatter diagrams scored full marks, sometimes the points were not labelled or a point, particularly $\mathrm{I}(244, 128)$, was plotted incorrectly. Answers to part (d) were generally poor, often scoring no marks. Common incorrect answers were r, $\frac{r}{2}$, a value or values above 0.5, a range of values such as 0 to 0.4, or phrases such as 'no correlation' or 'strong correlation'. Marks were only gained for an appropriate value linked to a source.

Question 4

Parts of this question proved a good source of marks for many candidates with the most able scoring all 12 marks. Answers to part (a) were usually correct and found, as was intended, from tables. A minority of candidates calculated $\mathrm{P}(M = 15)$ using the formula. Most candidates used the correct formula in part (b), but a small minority approximated 0.29 by 0.30 so as to use tables, resulting in a loss of all of the 3 marks available. Answers to part (c)(i) showed an improvement in the knowledge of the relevant formulae. Only a very small proportion of candidates apparently chose to ignore the emboldened word '**do**' and used $p = 0.29$ rather than $p = 0.71$, resulting in a loss of at least 4 marks since the comment marks in part (c)(ii) were dependent upon correct answers in part (c)(i). Answers to part (c)(ii) were somewhat disappointing. Many candidates made no reference to their (correct) answers in part (c)(i) but simply stated that "the samples could not be random as they only included women". Only better candidates were able to score at least 2 of the 3 marks available. The mark lost was usually for stating that "since the samples were not random, the claim was not justified".

Question 5

This question proved to be the most difficult on the paper. In part (a), many candidates made a complete hash of the deduction process by apparently not realising that 1 hour was equivalent to 60 minutes. Thus it was all too common to see candidates adding 1, or even 100, to both mean and standard deviation values or multiplying one or both values by 60 or 100. Future candidates clearly need to be much better prepared for similar questions involving linear scaling. In answering part (b), most candidates were clearly aware of the relevant formula for a confidence interval and identified that $z = 2.3263$ but lost at least 1 accuracy mark due to the aforementioned errors in part (a) or by simply using the given values of 1.90 and 3.32. The awarding of the mark in part (c) was rare indeed since the answer required a comparison of 1 hour or 60 minutes with a correspondingly **correct** confidence interval in part (b).

Question 6

Many candidates scored around 10 marks on this question but only the highest achievers scored anywhere near the full 17 marks. In part (a)(i), almost all candidates realised the need for standardisation and so obtained $\mathrm{P}(Z < 0.909)$. Far too many candidates then lost the final mark by quoting, from Table 3, $\mathrm{P}(Z < 0.9)$ or $\mathrm{P}(Z < 0.99)$. In part (a)(ii), most candidates recognised the need for an area change followed by a subtraction of areas but again lost the final mark by compounding an aforementioned error made in part (a)(i). In answering part (a)(iii), at least 50% of candidates undertook a calculation to give an answer other than zero. In part (b), many candidates scored 3 marks by showing a correct method but using $z = +1.2816$ instead of $z = -1.2816$. Perhaps a simple sketch would have shown that the value must be less than 69.5. Only better candidates realised that the answer required in part (c)(i) was {(a)(i) answer}^{20}. Many candidates attempted to introduce $\frac{0.55}{\sqrt{20}}$ or something similar at this point.

Answers to part (c)(ii) showed a small but welcome improvement. Whilst many candidates considered $\frac{69.25-69.5}{0.55}$ and so scored 0 marks, better candidates did attempt to find the variance or standard error of $\bar{X}$ before standardising. Sadly, many such candidates then failed to find the correct area and so lost 2 marks. Again, a simple sketch would have shown that the answer was greater than 0.5.

Coursework Component

A number of scripts had little marking on them; please annotate scripts fully indicating any calculations checked for accuracy and/or erroneous work or interpretation.

Please ensure that any packages sent to the moderator do not require a signature on delivery.

It was pleasing to note that centres had been thorough in obtaining signatures for the Candidate Record Forms from teachers and candidates. There were still some errors in the totalling of marks.

As has been mentioned in previous reports regarding 'other areas of work', candidates should quote other tasks **not specific to or modifications of their task**, which would use similar skills and techniques which **were** used in their task. For example, if doing a task on confidence intervals for the mean height, other areas might include doing confidence intervals for the mean time to run 100m.

There was some good quality statistical 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.

Candidates are reminded that there is no requirement to 'show', 'prove' or 'validate' the Central Limit Theorem using histograms. There is still some confusion amongst candidates about how the Central Limit Theorem applies to their task.

In tasks involving regression, some of the variables being linked were statistically questionable; candidates need to be very careful using some discrete bivariate data. The use of football statistics for regression can be difficult and is probably best avoided for a task. Candidates need to discuss the dependent and independent nature of their variables; some just did regression without any discussion of this at all.

Mark Ranges and Award of Grades

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