Version 1.0

General Certificate of Education (A-level)
January 2011

Mathematics MS/SS1B

(Specification 6360)

Statistics 1B

Report on the Examination

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.

General

Candidates appeared to find the demands of this paper similar to those of corresponding papers in recent series. Whilst most candidates scored well on questions 1, 2 and 5, they found much of question 3 and parts of questions 4 and 6 more challenging.

Calculation responses were generally sound and accurate, usually with sufficient evidence of method, whereas explanations and comments were often too vague, with the use of ‘it’ being a particular issue. Candidates made good use of their calculators’ inbuilt statistical functions, particularly in questions 1, 3 and 5, although, from answers to question 3, far too many candidates appeared to have no idea as to how to enter mid-points and frequencies, something that clearly needs attention. Whilst most candidates used tables in questions 4 and 6, those who used their calculators’ cumulative binomial probability functions and normal distribution functions did so with greater skill than candidates in previous papers.

Question 1

Most candidates scored at least 5 of the 7 marks available. In part (a), estimates were often sufficiently accurate in terms of magnitude, but a significant minority stated ‘zero’ or omitted the minus sign in part (a)(ii).

As is now the norm, the vast majority of candidates used the appropriate statistical function on their calculators to obtain a correct value for r in part (b)(i), although there were some cases where answers were penalised for quoting to less than three significant figures. The minority of candidates who calculated r using a formula often did so with a good understanding, if not particularly accurately. It was pleasing to see a marked reduction in, for example, the use of $\left(\sum x\right)\left(\sum y\right)$ rather than $\sum xy$. In part (b)(ii), most candidates mentioned ‘positive correlation’ but an acceptable adjective describing the strength was occasionally missing. Almost all candidates included a necessary statement in context by referencing ‘circumference’ and ‘weight’, but the all too frequent use of ‘centimetres’ and ‘weight’ was not acceptable.

Question 2

Most candidates scored between 7 and 10 marks, with the marks usually being lost in parts (b) and (c). Too many candidates lost valuable marks for ignoring ‘to three decimal places’. Fractional answers only scored full marks in parts (a)(i) to (a)(iii), where almost all candidates scored the 3 marks available. In part (a)(iv), $\mathrm{P}(L \cap F)$ or $\mathrm{P}(F|L)$, rather than $\mathrm{P}(L|F)$, was often found. Similarly, the use of $\mathrm{P}(L' \cap M)$ or $\mathrm{P}(L'|M)$ was often in evidence in answers to part (a)(v).

In part (b), far too many candidates used $\left(\frac{94}{126}\right)^2$ or $\frac{94}{126}\times\frac{93}{126}$ instead of $\frac{94}{126}\times\frac{93}{125}$.

In answering part (c), many candidates had the correct numerator of $349\times193\times103$ but a majority then used a denominator of $(645)^3$, and almost all candidates omitted the permutation multiplier of $3! = 6$.

Question 3

Surprisingly, this was probably the worst answered question on the paper, with even the better candidates scoring no more than about 9 of the 13 marks available. Most candidates scored the mark for answering part (a)(i) but many made a meal of the calculation, rather than simply evaluating $\left(\frac{0.98+1.00}{2}\right)$. Answers to part (a)(ii), which was essentially GCSE grade C material, were in the main very poor. All too often, the values quoted were 0.98 and 1.00, although numerous other incorrect values were seen. Even the better candidates, who managed to obtain the correct minimum value of 0.975, often quoted a maximum value of 1.004, 1.049, 1.0049 or even 1.05 rather than 1.005 or $1.004\dot{9}$.

Again in part (b), far too many incorrect answers were seen. A significant proportion of candidates used formulae, with varying success usually due to numerical or approximation errors. Even those candidates who used their calculators' inbuilt statistical functions obtained incorrect answers or quoted answers to insufficient accuracy. The most common incorrect answers were 1.065 and 0.073, obtained by using correct mid-points but ignoring frequencies.

Answers to part (c)(i) generally showed a sound understanding with at least 3 marks available despite incorrect answers in part (b). Incorrect z-values, omission of $\sqrt{n}$ or use of $\sqrt{8}$ were errors occasionally seen. Answers to part (c)(ii) showed that knowledge of when and why the Central Limit Theorem should be used was weak, with many candidates considering only the sample size. Those candidates who referred to 'normal' were often vague or incorrect in their statements, often stating 'it is normal', 'the data is normal' or 'the sample is normal'. Other comments were made about the juice itself rather than the volume of juice. Candidates exhibited a better understanding in answering part (c)(iii), perhaps because the idea of the use of mid-points for grouped data leading to an estimate of the mean was familiar from GCSE. Again, there was some confusion when candidates tried to discuss samples and populations, whilst some thought that the confidence interval was approximate because it was 99% rather than 100%.

Although correct answers, by comparing $\mu = 1$ with the confidence interval, were seen to part (d)(i), too many answers compared the sample mean with 1 litre or referred to 91 of the 100 cartons containing more than 1 litre. Again, in part (d)(ii), whilst some correct answers were in evidence, many candidates simply either noted that 1 litre was within the given interval or stated that 'it must contain all values as total probability is 1'. Some acceptable calculations of a normal probability were given, and a few candidates referred to $\bar{x} \pm 3s$ but rarely supported their statements with the necessary numerical justification.

Question 4

This question on the binomial distribution was a good source of marks for many candidates, with the more able often scoring most, if not all, of the 15 marks available. Save for a small minority of candidates who calculated $\mathrm{P}(R = 5)$ using the formula, part (a)(i) was usually answered correctly from the appropriate table in the supplied booklet. Again, in part (a)(ii), answers were usually correct with only a minority of candidates giving $\mathrm{P}(R \le 10)$, $\mathrm{P}(R \le 9)$ or $1 - \mathrm{P}(R \le 9)$ as their answers. Answers to part (a)(iii) were almost always correct, with evaluation by formula by far the more common method. As in previous series, part (a)(iv) caused candidates more difficulty, with many uncertain as to how to find $\mathrm{P}(5 \le R \le 10)$. Whilst most candidates did attempt to subtract two cumulative probabilities, they often selected one, or even two, incorrect values, or attempted (i) – (1 – (ii)).

In part (b)(i), most candidates showed correctly that either $0.85+\left(0.15\times0.80\right)$ or $1-\left(0.15\times0.20\right)$ resulted in 0.97. Incorrect reasoning that also led to the given answer was to assume that that a second shot was always made. Although many correct answers were seen to part (b)(ii), the logic involved proved too much of a challenge for many candidates. Whilst many identified the correct model of $\mathrm{B}(50,\ 0.97)$, most then attempted $\mathrm{P}(\mathrm{S}=48)$ by formula. Most correct answers appeared to be obtained directly from a calculator's inbuilt cumulative binomial function rather than by $\mathrm{P}(\mathrm{S}'\le 2)$ using tables for $\mathrm{B}(50,\ 0.03)$. In answering part (b)(iii), most candidates attempted to use np but then failed to identify the correct values of n and p. As a result, $80\times0.80=64$ and $\left(80\times0.15\right)\times0.80=9.6$ were common incorrect answers.

Question 5

A large majority of candidates scored the 11 marks available for calculations but then struggled with the explanations. Candidates' statements in part (a) were often too vague with the key reference to 'dependence' rarely seen. Some statements confused 'arrival time' with 'time taken' whilst others introduced traffic conditions.

In part (b), almost all candidates found accurate values for b (gradient) and a (intercept) using the regression functions on their calculators and then, presented correctly the least squares regression line. A small minority of, usually able, candidates calculated regression coefficients by formulae. In almost all such cases they were successful, but may have penalised themselves regarding time available to answer other questions.

In part (c), most candidates made correct use of their equations with $x=15$, although a significant minority used $y_{15}=58$ from the given table of data. Apart from some candidates who mistakenly took the leaving time as 7.30 am, most proceeded, often without evidence, to a correct final answer.

As expected, it was rare to see an incorrect answer to part (d)(i). In answering part (d)(ii), candidates were much better at giving a valid statistical explanation (extrapolation) than one based on the context. Again, however, many such explanations lacked clarity, suggesting, for example, that their y-value was outside the observed range. Very few contextual reasons were sufficiently clear and detailed to be awarded a mark.

Question 6

Most candidates achieved at least 7 marks on this question, with the better candidates scoring full, or almost full, marks. Almost all candidates knew how to standardise — without introducing an unnecessary continuity correction, which was penalised — and so the majority completed part (a)(i) accurately. Again, in part (a)(ii), the majority of candidates obtained the correct answer, with failure to apply the necessary area change by far the most common error. Despite similar questions on previous papers, far too many candidates incorrectly stated 'cannot be calculated' or continued to perform a variety of lengthy worthless standardisations and subtractions of areas for the 1 mark available in part (a)(iii).

The advice to use a suitable sketch in answering part (b)(i) proved helpful, and many good attempts were presented as a result. Many candidates scored 2 of the available 3 marks for rearranging their standardisation equations but lost a mark for not explicitly stating or indicating that -1.6449 and 2.3263 were z-values. Part (b)(ii) was often answered correctly, usually by firstly eliminating μ. However, attempts that began by trying to eliminate σ were often much less successful, and those candidates who decided to solve their own, different equations gained no credit whatsoever.

Mark Ranges and Award of Grades

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