Version 1.0

General Certificate of Education (A-level)
January 2011

Mathematics MS/SS1A

(Specification 6360)

Statistics 1A

Report on the Examination

Further copies of this Report on the Examination are available from: 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.

Written Component

General

Candidates appeared to find the demands of this paper somewhat easier than those of corresponding papers in recent series. As a result, there was a welcome increase in the proportion of candidates gaining high marks and, at the same time, a reduction in those gaining very low marks. Whilst most candidates scored well on questions 1, 2 and 5, only the better candidates appeared capable of making worthwhile progress in certain parts of questions 3, 4 and 6.

Calculation responses were generally sound and accurate, usually with sufficient evidence of method, whereas the explanations and comments in question 6 were often too vague. Candidates made good use of their calculators' inbuilt statistical functions, particularly in questions 1 and 5. Whilst most candidates used tables in questions 3 and 4, those who used their calculators' normal distribution functions and cumulative binomial probability functions did so with greater skill than candidates in previous papers.

Question 1

Most candidates scored at least 4 of the 5 marks available. As is now the norm, a large majority of candidates used the appropriate statistical function on their calculators to obtain a correct value for r in part (a), 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 $\left(\sum x\right)\left(\sum y\right)$ and $\left(\sum x\right)^2$ rather than $\sum xy$ and $\sum x^2$ respectively. In part (b), most candidates mentioned 'strong correlation' but the necessary additional word of 'positive' was occasionally missing. Almost all candidates included a necessary statement in context by referencing 'circumference' and 'weight' but the use of 'centimetres' and 'weight' was not acceptable.

Question 2

Many candidates scored between 5 and 9 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), the use of $\mathrm{P}(L \cap F)$ or $\mathrm{P}(F \mid L)$, rather than $\mathrm{P}(L \mid F)$, was often in evidence. 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

Most candidates achieved the 5 marks available in part (a) but very few were able to make significant progress in part (b). This was disappointing, particularly as the distribution of the sample mean has been examined on several previous papers. 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. In part (a)(ii), the majority of candidates realised that a difference of two areas was required, with many obtaining the correct answer. Failure to apply the necessary area change for $\mathrm{P}(Z < -0.3)$ was by far the most common error. The majority of candidates simply worked with the distribution of W, rather than $\overline{W}$, and so scored no marks. Candidates who worked with the distribution of $\overline{W}$, and so obtained correctly $\mathrm{P}(Z > -1.47)$, sometimes then failed to make the correct area change and so lost 2 marks.

Question 4

This question on the binomial distribution was a good source of marks for many candidates, with the more able often scoring most 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 $\mathrm{P}(R \leq 10)$, $\mathrm{P}(R \leq 9)$ or $1 - \mathrm{P}(R \leq 9)$ 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 $\mathrm{P}(5 \leq R \leq 10)$. Whilst most candidates did attempt to subtract two cumulative probabilities, they often selected one, or even two, incorrect values. In part (b)(i), most candidates showed correctly that either $0.85 + (0.15 \times 0.80)$ or $1 - (0.15 \times 0.20)$ resulted in 0.97. Incorrect reasoning that also led to the given answer often assumed 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 $\mathrm{P}(S = 48)$ by formula. Most correct answers appeared to be obtained directly from a calculator's inbuilt cumulative binomial function rather than by $\mathrm{P}(S' \leq 2)$ 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, $80 \times 0.80 = 64$ and $(80 \times 0.15) \times 0.80 = 9.6$ were common incorrect answers.

Question 5

A large majority of candidates scored all of the 10 marks available. In part (a), almost all candidates found accurate values for b (gradient) and a (intercept), usually using the regression functions on their calculators, and then presented correctly the least squares regression line. A small minority of candidates calculated regression coefficients by formulae. In most such cases they were successful, though a minority were penalised for lack of accuracy. In part (b), 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 to a correct final answer.

Question 6

As usual, candidates found the calculations in parts (a) and (b)(i) easier than the explanations in parts (b)(ii) and (c), with the result that about 6 marks was the norm. Almost all candidates scored the 1 mark available in part (a)(i). However, more difficulties arose in answering part (a)(ii). All too often, candidates evaluated $\sqrt{\frac{400.24}{65}}$ and then attempted to make it unbiased by incorrectly multiplying by $\frac{65}{64}$ rather than $\sqrt{\frac{65}{64}}$. The best and simplest correct approach was $\sqrt{\frac{400.24}{64}}$. Answers to part (b)(i) showed a sound understanding and usually scored full marks. Answers to part (b)(ii) showed that knowledge of when and why the Central Limit Theorem should be used was weak, with many candidates stating 'large sample size' as the reason. Those candidates who referred to 'normal' were often vague or incorrect in their statements by often stating 'it is not normal', 'the data is not normal' or 'the sample is not normal'. Whilst many candidates stated that their confidence interval included 20, they then made the mistake of deducing that this supported the claim.

Coursework Component

The work submitted in this series was sent in the appropriate stationery. However, there were still some centres who sent work which required a signature, contrary to the AQA instructions.

As mentioned in previous reports, it is important that all centres read the advice offered on the feedback forms carefully, in particular if the form indicates that the centre is close to the tolerance limits, as further drifting from the standard could lead to an adjustment of the centre's marks.

The work seen was of a good standard and generally appropriately assessed. There were some examples of centres marking a little leniently, and some quite short pieces of work lacked the depth of discussion and analysis for the very high marks that they were awarded.

It is important that candidates' scripts read as a continuous piece of prose. Although some titles used during the script are appropriate, some pieces of work became very formulaic in their approach and seemed to read as the answers to a series of questions rather than a continuous piece of writing.

The candidates in this series displayed a much better understanding of the Central Limit Theorem and used the idea of overlapping confidence intervals well. The discussion of sampling continues to lack some detail and could be improved further.

Mark Ranges and Award of Grades

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