A-LEVEL
STATISTICS

SSIB – Statistics 1B
Report on the Examination

6380
June 2014

Version: 1.0

General

In general, students appeared well-prepared for most of the topics examined, though a notable exception, in many cases, was that of probability (Question 4) where the necessary knowledge often appeared to be almost non-existent. Most students appeared confident in answering the numerical parts of questions, usually by making the expected use of their calculator's in-built basic statistical functions, though woeful levels of manipulation were in evidence in far too many cases. Even the best students often struggled with parts of questions that required comment or interpretation — perhaps an area for more practice in the future.

Question1

As intended, this question proved to be a confidence-booster first question for virtually all students. In part (a), after correctly ordering, or sometimes ranking, the given data, most students then used $\frac{(n+1)}{2}$ etc to identify correct values for the median, the two quartiles and then the interquartile range. A very small minority lost some marks for using $\frac{n}{2}$ etc, whilst even fewer lost all the marks available for not ordering the given data. It was rare to see an incorrect answer to part (b). Most students gave a valid answer in part (c), although some students lost the mark for stating a general disadvantage of the mode instead of referencing their statements to the given data.

Question 2

There were many completely correct answers to this question, though part (b) proved too challenging for weaker students. In part (a), those students who simply wrote down answers from their calculator's normal distribution functions usually scored full marks, though a very small minority scored no marks for three incorrect answers. Students sometimes lost marks, particularly in parts (a)(ii) and (iii), for treating time as discrete, thus finding $P(X > 7)$ or $P(4 < X < 9)$ and/or the complementary area. Many students had problems with part (b). Whilst some of these made no attempt, others standardised using 2.4^2 and/or equated an expression to a z-value other than $0.84(16)$, or, even worse, 0.80. Even those students who had obtained $\frac{15-\mu}{2.4} = 0.84(16)$, far too often were unable to find the correct value of μ; the use of inconsistent signs often resulted in a fudged 13 or a non-fudged 17 or -13.

Question 3

It was rare to award other than full marks for answers to part (a), with almost all students making excellent use of their calculator's regression function. Again, in part (b), almost all students scored full marks, though a small minority substituted 38 instead of 30. Answers to part (c) showed that students were aware of the dangers of extrapolation but attempted explanations in context were often woolly or involved confusing phraseology. In part (b), the better students found $\hat{y}_{20}$ and then either $(117 - \hat{y}_{20})$ or $(\hat{y}_{20} - 117)$, the former been correct. Of the many other students, about half made no attempt and the other half made worthless confused attempts such as evaluating $\hat{y}_{117}$, $117 - 20$ or $\frac{117}{20}$. It was very rare to see an answer worth the one mark in part (e), as almost all answers failed to compare the size of 2.71 with the given values of y. When attempts were made, they often simply involved the words 'accurate' or 'inaccurate' with either no or a totally incorrect justification.

Question 4

For most students, this question proved to be the most demanding on the paper. Those who started in part (a) with a Venn diagram, or less frequently a 2-way table, scored the 3 marks available. The remaining students, who formed a large majority, started down a totally incorrect route, often ($0.70 \times 0.45 + 0.30 \times 0.55 + 0.45$). Following such incorrect attempts, there was often a correct start to part (b) using ($0.70 \times 0.55 = 0.385$), but then this was followed by an incorrect comparison. Those students who ignored the word 'numerical' and so attempted often convoluted qualitative or practical explanations spent time for no reward. Many students scored at least one mark in each part of (c) where, unlike part (a), multiplication of probabilities was required. The marks lost were invariably for using 0.385 instead of 0.45 for $\mathrm{P}(A \cap M)$ in part (c)(i) and, similarly, in part (c)(ii), for using (0.30×0.45) instead of 0.2 for $\mathrm{P}(A' \cap M')$. Responses to this question suggested that, whilst students were competent in working with independent events, they lacked the necessary skills to work with non-mutually exclusive and/or dependent events.

Question 5

Almost all students, as expected, used their calculator's correlation function, rather than a time-consuming formula approach, to obtain a correct value for r. Interpretations in part (a)(ii) were in context and in the main sound, though the use of 'very' attached to 'strong' or the omission of 'positive' lost a mark. In part (b), points were almost always labelled but one clearly incorrectly plotted point was by no means unusual. Full marks were the norm in part (c). Perhaps through careless reading, some students gave no reasons whilst others stated "D and C", presumably through considering only the 8 days A to H. Again, in part (d)(i), most students were aware of the correct formula or found it in the booklet, and then, in part (d)(ii), correctly revised their previous interpretations.

Question 6

Students recognised that this question involved binomial distributions and so most attempted to use the formula for $\mathrm{B}(26, 0.06)$ in part (a). However, a small, but nevertheless significant, proportion were unable to correctly write down the expression for $\mathrm{P}(M = 2)$, whilst others could not correctly evaluate a correct expression. Almost all students changed correctly to $\mathrm{B}(50, 0.15)$ in part (b), often using tables rather than their calculator's' cumulative binomial function. However, the usual confusions in interpreting the stated phrases were often in evidence: (i) $\mathrm{P}(I \le 10)$; (ii) $1 - \mathrm{P}(I \le 4)$; (iii) $\mathrm{P}(I \le 12) - \mathrm{P}(I \le 5)$ etc. In part (c), most students started by attempting to find the mean and the variance of $\mathrm{B}(50, 0.15)$ though a minority used $\mathrm{B}(100, 0.15)$ or even $\mathrm{B}(5000, 0.15)$. In all cases, there was sometimes confusion between variance and standard deviation, with some students even considering 3.94 to be a value of the sample's standard deviation. Such confusions, together with a default perception of 'the larger, the better', severely restricted the number of students scoring more than 3 marks.

Question 7

A majority of students obtained correct answers in part (a), usually directly from their calculators. However many students made elementary and often catastrophic errors: incorrect midpoints (minimum loss of 2 marks); ignoring f-values (minimum loss of 3 marks); ignoring x-values (loss of all 4 marks). Needless to say, such errors often had a major impact on available marks in part (b). Most students knew the basics of confidence-interval construction. Thus, those with correct answers in part (a) often scored full marks in part (b)(i). However, errors common to many students included the use of $z = 2.0537$ instead of 2.3263, and/or $\sqrt{13}$ instead of $\sqrt{160 \text{ or } 159}$. In answering part (b)(ii), most students who had scored well in earlier parts compared 61.7 with their confidence intervals. Sadly, too many then went on to state that "Since my confidence interval includes 61.7, the claim is supported".

Mark Ranges and Award of Grades

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

Converting Marks into UMS marks

Convert raw marks into Uniform Mark Scale (UMS) marks by using the link below.

UMS conversion calculator www.aqa.org.uk/umsconversion