AS

MATHEMATICS

MS1B Statistics 1B
Report on the Examination

6360
June 2018

Version: 1.0

General

The general level of attainment was slightly lower than in previous series. This was perhaps partly due to a changed entry profile for this final series of the modular specification, but also due to some more challenging parts in some questions. These parts provided good differentiation for the highest achievers whilst there remained plenty of scope to score marks for less able students. Nevertheless, there appeared to be a significant increase in the proportion of students who were unable to score other than minimal marks on all questions.

Most students made appropriate, usually accurate, use of their calculators in questions 2 and 4. They also made use of the booklet of tables or, in an increasing proportion of cases, their calculators in questions 3, 5 and 6. On the negative side, there were some very untidy, almost unreadable, scripts. Additionally, there were many students who struggled with more basic skills in algebra and arithmetic. Typical of these were:

- an inability to solve directly $\frac{350-\mu}{\sigma}=0$ for σ without a second equation;
- calculating n from $\sqrt{n}=8.37...$ as $2.89...$ or 3.

Finally, not reading carefully the question led to students missing out on marks.

Question 1

Many students didn't follow the instruction 'For these data ...' and simply stated general reasons such as "not representative" and "an anomaly" without any reference to 5 or 95. Most students scored full marks in parts (b) and (c) with values for the mean and standard deviation quoted correctly from calculators. However, many students could not adequately explain which pair was the more appropriate and reference to the skewness of the data was almost non-existent.

Question 2

Most students gave at least one, very often two, appropriate values based on the given scatter diagrams where $–0.899$ and $–0.007$ were the actual values. The most common errors were to state a value of –0.7 or omit the negative sign for the first value. Calculations, mainly using calculators, were generally very sound in part (b)(i). Many stated an acceptable answer of 3.75×10^{-3} but a minority stated 3.75 or missed 'to three significant figures'. Despite the near-zero value for r, there was a clear reluctance amongst students to commit to "no correlation" with many adding qualifications such as "almost" or preferring "extremely weak positive". To complete the interpretation, students had to reference the two variables using at least "time" and "cost of items" but these were sometimes changed to "minutes" and "money spent".

Question 3

Almost all students were able to find correctly the two normal probabilities in part (a). When an error occurred, it was usually for finding the wrong tail in part (i) or using other than 0.5 for $P(Z<0)$ in part (ii). Part (b) proved challenging and provided good differentiation. Stronger students were usually successful in part (i) but some methods, involving two simultaneous equations, were unnecessarily lengthy. Far too many, otherwise correct solutions, missed the instruction 'to the nearest gram'. Some students often used $+2.3263$ then left or dropped the negative value for σ

whilst weaker students used the given probabilities instead of z-values. Part (b)(ii) was found to be one of the hardest calculations on the paper. Multiple attempts were often seen as students tried to find a method that gave an answer that they thought reasonable. Common errors were the use of $z = 0.8416$ or $\frac{2w}{\sigma} = z$-value, and answers including the mean value.

Question 4

There was a variety of answers to part (a), many of which scored the mark but some statements revealed an apparent lack of knowledge of the importance of correct identification. The vast majority of students found the equation of the regression line correctly using their calculators, although a small minority were too severe in rounding their directly-stated values. Even those using the formulae approach were generally successful. However, perhaps as a result of part (a), more students than in previous series found the equation of x on y and so lost almost all of the subsequent marks. As in Question 2, many students either changed or gave incomplete descriptions of the variables involved with “concentration”, “weight” and simply “yield” been very common. Some students commented on only one of their values or confused the two, whilst a small minority referred to a context from a similar question on a previous paper. The calculation of the residual for plant H was often correct except for a sign error. However, fewer students appreciated the word ‘Hence’ in part (c)(ii) and set about calculating all 11 residuals with little success. This was disappointing given that, on a previous paper, knowledge that $\sum r_i = 0$ was tested. Whilst most students appreciated that extrapolation was needed in part (d), only a minority provided numerical support and only a few of these made the required comparison. Simply stating that calculated estimates were “different from observed” or “had big residuals” were insufficient to score the mark.

Question 5

Responses to part (a) revealed that many students appeared to be unaware of the necessary conditions for a binomial model. Only a minority of students were able even to make the three correct decisions. In part (i), when a binomial model was deemed appropriate, $n = 20$ and $p = 0.33$ were the norm, but $p = 0.167$ or $p = 0.05$ were not that rare. In part (ii), when deemed inappropriate, “n not known” was a common insufficient reason. Many students decided incorrectly that a binomial model was appropriate in part (iii). For those who made the correct decision, only a minority referenced a changing probability. Answers to part (b) showed that the vast majority of students could apply the binomial model with many students scoring full marks. When marks were lost it was invariably for errors in interpreting ‘at most’, ‘at least’ and, less frequently, ‘more than’. Calculators often helped in part (iii) as there was then no requirement to change from $\mathrm{B}(50, 0.9)$ to $\mathrm{B}(50, 0.1)$.

Question 6

Students were competent in constructing a confidence interval with full marks being scored in most cases. Where this was not the case, it was usually for using 1.6449 instead of 1.96 (loss of 2 marks) or for not rounding answers to the required accuracy (loss of 1 mark). Only in the rarest of cases was an incorrect form used. In part (a)(ii), almost all students referenced the confidence interval but far too many continued to use “it” or “mean” instead of 1072. Part (b) provided a good challenge for the more able students although success was seen by low-scoring students. Most students used the correct z-value but only a minority opted for the efficient approach based on using 1070 in the equation for n. The alternative required the setting up of two simultaneous equations for 1060 and 1080. However, many only set up one equation using $\bar{x}$ or μ and then either abandoned the question as they had one equation involving two unknowns or substituted 1075 or even 1072 for the mean value. It was frustrating to again see students missing a mark for not quoting the answer to the required accuracy.

Question 7

It was somewhat rare to see three correct values stated in part (a) and even rarer to see them attached to the appropriate part. Whilst answers to part (i) were often correct, those for parts (ii) and (iii) were either numerically correct but reversed or simply numerically incorrect. This suggested that many students had a limited understanding of the terms 'independent' and 'mutually exclusive', with the latter sometimes confused with 'exhaustive'. Most students multiplied the three correct probabilities in part (b)(i). Most students also obtained the correct answer for part (b)(ii) but in part (iii), it was not unusual to see the omission of including 'all 3 attending' or, in a minority of cases, multiplying an otherwise correct answer by 3 or 6. In part (iv), the probability for part (A) was usually correct but, in part (B), the use of 0.48 instead of 0.52 was quite common.

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