Version : 1.0 0609

General Certificate of Education

Mathematics 6360

MS03 Statistics 3

Report on the Examination

2009 examination - June series

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.*

General
Whilst this year's average level of performance was down on the impressive standard of last year, a large proportion of candidates were able to achieve high marks.

In general, most candidates found all questions accessible. However, weaker candidates had significant problems with question 2(b) and only the strongest candidates could make any real headway beyond part (a) of question 7.

Question 1
This question was expected to cause few candidates problems, but a considerable proportion produced incorrect solutions in part (a). Although a small minority found two separate confidence intervals, each for a single proportion, the main error was to find a pooled estimate of a common proportion. This was quite simply incorrect: a pooling of proportions is only valid when it can be assumed that $p_1 = p_2$, something that was clearly not the case here in calculating a confidence interval for $p_1 - p_2$. Pooling can be used in testing $\mathrm{H}_0: p_1 - p_2 = a$ but even then only when $a = 0$. Those candidates who did not attempt a pooled estimate, but worked with the approximate distribution of $\hat{p}_1 - \hat{p}_2$, invariably scored full marks. In part (b), all candidates who had attempted a confidence interval for $p_1 - p_2$ realised that the key was whether or not zero was included in their confidence interval and so scored both marks.

Question 2
In answering part (a), most candidates gained full or almost full marks. They used a correct expression to deduce the given answer of 0.34 in part (a)(i) and, in part (a)(ii), again correctly used the multiplication law for two dependent events. Similarly, in part (a)(iii), most candidates realised that Bayes' Theorem was required, and so the vast majority scored at least 3 of the 4 marks available.

Part (b) unexpectedly proved to be beyond all but the strongest candidates. The most common solution was to simply evaluate $0.30 \times 0.55 \times 0.15$; candidates should have realised that more was expected for 5 marks. It was expected that part (a)(iii) would lead candidates to evaluate $\mathrm{P(City \mid HB)}$ and $\mathrm{P(Country \mid HB)}$ prior to multiplication. Of the minority who went down this route, very few indeed saw the need to introduce the permutation multiplier of $3! = 6$.

Question 3
There were many fully correct solutions to this question. Most candidates realised that $2z\sqrt{\dfrac{p(1-p)}{n}}$ had to be equated to 0.1, followed by the substitutions of $z = 2.3263$ and $p = 0.35$ before solving correctly for n. A minority of candidates appeared to confuse the situation with that of a required accuracy for μ, as they began with $\hat{p} \pm \dfrac{z\sigma}{\sqrt{n}}$ and then substituted $\sqrt{0.35}$ for σ.

Question 4
This question also produced many very good, even excellent, solutions. When marks were lost, it was usually in quoting the two hypotheses. Common errors were the use of sample means, rather than population means, and/or the omission of the value 15. Somewhat surprisingly, this value was often then introduced correctly when evaluating a test statistic.

Most candidates correctly evaluated a z-statistic without pooling the variances. Given the comments with regard to incorrect pooling in question 1, pooling of the two sample variances here would not have been penalised since they were similar in size, nor of course would the associated use of t_{138}. It was gratifying to see the large proportion of correct conclusions in context and not phrased definitively.

Question 5

There were a few unconvincing, even contrived, answers to part (a) but, in the main, candidates were well-prepared for this proof. Answers to part (b)(i) were surprisingly good; indeed it was rare to find a candidate not gaining full marks. Answers to part (b)(ii) usually continued to be impressive. Most candidates correctly made the two necessary continuity corrections and then found the correct answer, either from making the appropriate area changes when using tables or directly from their calculators.

Question 6

Candidates are now generally much better prepared for questions involving linear combinations of dependent or independent random variables. Most candidates found correct values for the mean and variance of X, though a minority considered that $\text{Var}(X) = \text{E}(X^2) = 6.8$. These candidates, due to follow-through marks, and almost all others scored the 5 marks available in part (b). In part (c), the usual but rare errors were adding standard deviations rather than variances or finding the complement of the correct answer. The minority of candidates who, after correctly stating the mean and variance of $(L+M)$, simply stated, without any working, an incorrect probability from their calculators' normal statistical function lost 3 of the 5 marks available.

Question 7

As stated earlier, much of this question proved too much of a challenge for most candidates. Most candidates were able to make a worthwhile attempt at part (a) with many scoring all 5 marks. When marks were lost, it was invariably for standardising 149.5 or 150 rather than 150.5. In part (b)(i), many candidates failed to state hypotheses in terms of λ. Also, despite the wording of the question, about 50% of candidates attempted an approximate normal test based on either $\lambda=2$ or $\lambda=10$. The first is simply incorrect and the second is at best borderline. In general candidates should not opt for approximate tests for binomial proportions or Poisson means when tables in the supplied booklets are available for exact tests.

Only candidates opting for an exact test in part (b)(i) stood any realistic chance of scoring marks in part (b)(ii). Most such candidates had a basic idea as to what was needed but only the very best obtained the correct value of 15. The usual alternative was 14, but 13 and 16 were seen. In part (b)(iii), most candidates defined power correctly, usually as $1 - \text{P}(\text{accept H}_0 \mid \text{H}_0 \text{ false})$. They then attempted to calculate this probability using $\lambda=15$ and so often scored at least 2 marks. A very small proportion of candidates are to be congratulated on scoring all 17 marks in correctly answering this final question.

Mark Ranges and Award of Grades

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