A-LEVEL
STATISTICS

SS05 Statistics 5
Report on the Examination

6380
June 17

Version: 1.0

Further copies of this Report are available from aqa.org.uk

Copyright © 2017 AQA and its licensors. All rights reserved.
AQA retains the copyright on all its publications. However, registered schools/colleges for AQA are permitted to copy material from this booklet for their own internal use, with the following important exception: AQA cannot give permission to schools/colleges to photocopy any material that is acknowledged to a third party even for internal use within the centre.

General

Most students were able to complete the paper in the time allowed. Generally, the working was clear and students seemed well-prepared. Questions on the continuous distributions were successfully completed by most students and standard tests and calculations were performed competently and accurately. Question 4 on the goodness of fit test to a Poisson distribution posed the greatest problem for the majority of students. The interpretation, especially in questions 4(b), 6(a)(i) and 6(b)(ii) showed a lack of comprehension of the situation being modelled.

It was noticeable that conclusions to the hypothesis tests were poorly done. Students often failed to express any element of doubt and, in the case where H_0 is accepted, gave significant evidence in accepting the null hypothesis rather than insufficient evidence to reject it.

Students still need to understand how to complete a test using a p-value rather than a critical value when using the statistical functions on their calculator. Intermediate values such as the mean, standard deviation, test statistic and degrees of freedom given in the calculator display can provide evidence where there is miscalculation. Nevertheless, students should continue to be encouraged to increase their facility with their calculators.

There were fewer scripts lacking clarity this year; hopefully students will continue to use black ink and to write as clearly as possible.

Question 1

This was an accessible question modelling a rectangular distribution. Students, in the main, gained full marks on part (a). Part (b), requiring the application of the Central Limit Theorem, proved to be more of a challenge.

Question 2

This question required knowledge of the test for a normal population standard deviation. Many students failed to appreciate that 15.3 minutes was the population standard deviation and had not been calculated from a sample. A fair number of students failed to calculate the standard deviation of the sample correctly; perhaps the negative signs were to blame. Where the appropriate χ^2 test was chosen, students could generally calculate the test statistic accurately. Finding the correct lower critical value of 3.94 was more difficult. Many students only gave the upper critical value of 18.307. This inevitably led to the incorrect conclusion of accepting H_0, that the standard deviation was equal to 15.3 minutes, when it should have been rejected. Part (b) required an appreciation that the result of the hypothesis test provided evidence that the standard deviation had decreased. Thus the variability of the differences between the actual and estimated delivery times had been reduced and Gilbert had achieved his aim. .

Question 3

This was a question modelling the exponential distribution. A minority of students struggled with the units. Students should learn to consider their answers in the light of the question and judge whether their response is reasonable. Errors in part (b) could have been easily corrected with more thought about the context. Part(c) was more challenging, but many students accurately calculated the number of hours for one security light as 365×24; calculated the probability that one security light would be working 365 days later and were then able to raise their probability to the power of 15 for the final answer.

Question 4

Many students found this question to be challenging. Some were unable to calculate the mean, 4, of the frequency distribution accurately. This severely affected their solution to the rest of the question. A few students erroneously specified the distribution to be modelled in their null hypothesis as $\mathrm{Po}(4)$. The majority of the students who correctly obtained the mean as 4 could calculate the probabilities and most of the expected frequencies correctly. Many students correctly pooled the classes for 0 and 1 live births per litter. There were a considerable number of students who failed to appreciate that the final class was 8 live births or more and hence did not arrive at the correct expected frequency of 6.14. This led to an additional pooling of the final class to 7 or more. $\frac{(O-E)^2}{E}$ was generally calculated and summed in a correct manner but the final test statistic of 7.23 was not often seen. Lack of accuracy in these questions often arises from rounding errors. If the graphical calculator is used, then as much evidence as possible from the test display should be recorded. Most students realised that the degrees of freedom were 2 less than their number of classes and those with the correct test statistic usually went on to accept the null hypothesis correctly. A sizeable proportion of these students, however, made an error in their conclusion.

Question 5

This was a question on two-sample hypothesis tests. Part (a) tested equality of variance and part (b) tested equality of means. Part (a) was the least successful of the tests as some students failed to recognise the application of the F-distribution. Common errors were in the order of the degrees of freedom, and in the final conclusion. Part (b) was successfully attempted by the majority of students. A common error was the use of the z-distribution rather than the t-distribution. It is worth noting that an initial error in the standard deviations had a cumulative effect throughout the question and students must be encouraged to check their data entry when using their calculators. Most students knew that a type II error might have been made in their answer to part (c) but greater attention to detail in the explanation was needed to gain full marks on this section.

Question 6

This was a question on calculating and interpreting confidence intervals for both the population mean and variance based on a random sample. Part (a)(i) confused many students; a comparison between the range, 24, and 6 standard deviations, 15, was the reasoning required. Parts (a)(ii) and (iii) were generally successfully negotiated. Errors in part (b)(i) were caused predictably by the use of the z-distribution rather than the t-distribution but less predictably by the use of 2.5 for the standard deviation rather than the sample standard deviation found earlier.

Part (b)(ii) required students to realise that while their confidence interval showed that the pots of paint contained more than $200\ \mathrm{ml}$ on average, inspection of the sample data showed that 2 pots were below $200\ \mathrm{ml}$.

Use of statistics

Statistics used in this report may be taken from incomplete processing data. However, this data still gives a true account on how students have performed for each question.

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