AQA

AS
STATISTICS

SS03 Statistics 3
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 made a good attempt at all questions. In hypothesis testing, many students seemed unable to state appropriate hypotheses that demonstrate knowledge of the difference between 'matched pairs' and 'independent samples' testing.

In situations such as question 3(b), many students presented their answers in a prose form that lacked structure and usually failed to include sufficient, clearly stated points.

Students often omitted to include a reference to significant evidence of a change observed 'on average', or a reference to means or medians in their conclusions to hypothesis tests on a population average.

Students generally used the given context when stating conclusions and assumptions.

Question 1

In part (a) many students did not appreciate that in a matched pairs test the hypotheses should refer to the 'population median **difference**' or the 'population mean **difference**'. Those who did not appreciate this tended to quote hypotheses for carrying out a test on population averages based on independent samples. For correctly stated hypotheses, involving a mean or median difference, it should be noted that the differences found should be consistent with the alternative hypothesis.

Almost all students calculated differences correctly and most went on to rank them correctly, assigning the rank '1' to the smallest non-zero absolute difference. Many students achieved the correct total of ranks and identified clearly which total was their test statistic, found the correct critical value and made the appropriate comparison.

Few students could state the correct conclusion in context, and in particular they often neglected to say precisely enough that there is significant evidence that the **average** blood pressure is lower after three hours of not smoking.

In part (b) most students identified that there may be a difference between male and female physiology, but few students appreciated that by only considering males, if a difference in average blood pressure after three hours of smoking did exist, it was more likely to be detected since smoking was the only variable involved.

Question 2

In part (a) many students produced excellent, fully correct solutions. A considerable number of students made slips in their rankings and failed to deal with all the ties correctly. Very few students incorrectly found the product moment correlation coefficient from the raw data.

In part (b)(i) many correct hypotheses were seen but a common error was to omit the **positive** association in the alternative hypothesis. Most students found the correct Spearman's critical value, compared this correctly with their test statistic and stated an appropriate conclusion in context.

In part (b)(ii) many students stated the meaning of a Type I error correctly but fewer stated this in context.

Question 3

In part (a) most students stated their hypotheses correctly in terms of two populations having identical/non-identical distributions. The samples were generally ranked correctly as one group and students went on to find the correct totals and U test statistic, find the correct critical value and identify the appropriate value of U to compare with their critical value. A few students did not make it clear which of their U-values was the selected test statistic and hence lost marks.

Many correct conclusions, in context, were seen, but a common error was failing to appreciate that a child who has a sibling less than two years younger can successfully read, **on average**, fewer of the 60 words.

Many students found part (b) to be extremely challenging. Some students appreciated that a different sample of children who have more than one sibling must be obtained. A minority of students also realised that this should be a random sample of such children, who should be aged 7 years, and these children should be tested on how many of the 60 words they can read successfully. Many responses referred to a hypothesis test analysis rather than how further information could be obtained.

Question 4

In part (a) many correct solutions were seen, but a significant number of students failed to complete the missing ranks correctly.

In part (b) many students stated their hypotheses correctly in terms of populations having identical/non-identical distributions. Those students who expressed their hypotheses in terms of population averages generally realised that this referred to population medians, but several students used suffices with η that were not defined.

Students generally successfully found the totals and the values of $\sum \frac{T_i^2}{n_i}$, and used these results in the formula for finding the test statistic H. A significant number of students made slips in the final calculation and hence did not achieve the correct test statistic. Many students found the correct number of degrees of freedom and the critical value but a significant number of students incorrectly thought that there were 23 degrees of freedom, rather than the correct value of 3.

Very few students stated the conclusion correctly in context, with many students failing to appreciate that there is significant evidence of a difference between average yields for **at least two** categories of weed density. Many students correctly identified that very low weed density results in highest average corn yield. Only a very small minority of students supported their conclusions with appropriate numerical considerations of mean or median ranks or scores or totals of ranks. Several students did not attempt this part of the question.

Question 5

In part (a)(i) students successfully placed observed frequencies in the correct cells in the table.
In part (a)(ii) most students stated appropriate hypotheses and found the correct expected frequencies and the correct test statistic. The method for obtaining expected frequencies and the

test statistic should be indicated so that method marks can be awarded in the case of incorrect placing of observed frequencies in part (a)(i). The correct degrees of freedom and critical value were usually stated as was a correct conclusion in context.

In part (a)(iii) many students correctly identified that the most significant contributions to the test statistic, and therefore the strongest sources of association, came from patients having Cognitive Therapy and from patients having Sympathetic Learning. Those students who compared the observed and expected frequencies generally made a correct statement concerning the main sources of association.

In part (b)(i) many students did not appreciate that, in a matched pairs test, the hypotheses should refer to the difference in population median (or η). Most students found signs that were consistent with their alternative hypothesis. The correct test statistic was identified by most students but there were a significant number of solutions where students failed to omit the difference of zero and hence used the incorrect binomial distribution to find the necessary probability, using n as 11 rather than 10. Many correct comparisons of 0.0547 with 0.05 were seen but students frequently failed to state the conclusion that there was no significant evidence that social competence (score) increases, **on average**, following the training.

In part (b)(ii) the majority of students stated correctly that the sample of patients must be a random sample from those undertaking SCT.

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