# AS
# FURTHER MATHEMATICS
## 7366/2S: Statistics
## Report on the Examination

7366
June 2019

Version: 1.0

## General

The paper appeared to offer ample opportunities for students to score reasonable marks. On the other hand, the highest marks were rarely achieved by any students. This was almost entirely due to difficulties in the interpretation of statistical results and concepts or not constructing fully rigorous arguments. In future students should work on ensuring that their mathematical arguments are complete and can easily be followed by other mathematicians. Care should also be taken when using statistical language to describe concepts or to interpret results. Students would benefit from discussing the conditions under which statistical models are valid in context and being guided in how to articulate their reasoning clearly.

## Question 1

The vast majority of students scored the mark for this question. The numbers of responses for each of the incorrect options were roughly similar.

## Question 2

The majority of students did not score the mark for this question with both 5% and 4.8% being chosen more than the correct answer.

## Question 3

The majority of students scored at least three marks for this question. It was common for students to calculate the population variance instead of the sample variance. Some students found the wrong $z$-value, finding a value that corresponded to a 92% confidence interval rather than the 96% confidence interval. A minority of students took the values given to be the mean and variance or used an incorrect formula for the confidence interval.

## Question 4

**(a)** The majority of students scored full marks for this part. A significant proportion of students used the wrong limits, in particular 2 and 3 or 0 and 1. Some students only gave their answer to two significant figures. There were some responses treating the variable as a discrete random variable. Full marks could be obtained by evaluating the integral on a calculator, but students doing this usually wrote down the correct integral, with correct limits first.

**(b)** The majority of students scored at least two marks for this part. Many students lost the final mark which required a rigorous demonstration of the result including the '$\mathrm{d}x$' and clearly stating that their result was $\mathrm{E}(X^{-1})$. A significant proportion of students attempted to find $\mathrm{E}(X)$ instead. There were some responses treating the variable as a discrete random variable. A significant proportion of non-response was seen. Whilst this question could be answered successfully on a calculator, some simple algebraic manipulation of the integrand was required first.

**(c)** The vast majority of students scored full marks for this part with a minority of students using integration to derive the result. Common errors included using a wrong formula, calculating $4\mathrm{E}(X^{-1})$ or $4\mathrm{E}(X^{-1}) - 3$ instead of $2\mathrm{E}(X^{-1}) - 3$. There was significant non-response.

## Question 5

**(a)(i)** The majority of students scored at least one mark for this part with many scoring full marks by completing a rigorous proof. The most common errors were to treat the variable as a continuous random variable or using an incorrect formula, such as $\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+...+\frac{1}{n}$. Some students chose a value of $n$ and attempted to verify rather than attempt a proof.

**(a)(ii)** The majority of students scored at least one mark for this part with many scoring full marks. Many students lost the final mark for a lack of rigour, often not making clear what their E($X$), E($X^2$) or Var($X$) were. Some students used an incorrect formula for E($X^2$) or obtained it by working backwards from the answer. Some students chose a value of $n$ and attempted to verify rather than attempt a proof. There was significant non-response.

**(b)** The majority of students scored one mark for this part but many students scored no marks. The most common mark was for identifying that the dice was unbiased or an equivalent description, with fewer being aware of the distribution being used for discrete outcomes from 1 to $n$. A proportion of students described conditions that were more suitable for a binomial distribution. There was significant non-response.

## Question 6

**(a)** The vast majority of students scored full marks for this part. The most common errors were to not adjust the parameter for the change from 1 hour to 4 hours or excessive rounding in their answer when the question asked for three significant figures.

**(b)** The majority of students scored full marks for this part. Some students did not attempt to combine the Poisson distributions and scored no marks. A common error was to incorrectly calculate the probability, in particular calculating P($X$ > 8) = 1 – P($X \leq$ 7) though many were still able to score the final mark for a correct follow through comment.

**(c)** The majority of students scored no marks for this part, usually for not using the data given in the question. Some students made arguments for other distributions seen in AS Mathematics or Further Mathematics. There were some generic comments made about the conditions required for a Poisson distribution but these were difficult to apply without more knowledge about the machines. Some students made inconsistent comparisons, for example a variance of 0.25 with a combined mean of 7. There were some students who attempted to use the standard deviation as the variance or made an incomplete argument. There was significant non-response.

## Question 7

**(a)** The majority of students scored at least four marks for this part. The most common hypotheses stated by students were 'no association' for $H_0$ and 'association' for $H_1$ following the language used in the question. The use of 'independent' and 'not independent' was accepted, although it is essential that students cite the variables in their hypotheses. There were some students who thought that an observed frequency equal to 5 required the merging of columns which was unnecessary. Some students made numerical slips when calculating the test statistic. The vast majority of students compared their test statistic with a critical value rather than calculating the corresponding probability on a calculator. Many students reached a conclusion in context that was too definite in nature, stating that they had 'proved' or 'shown' rather than having found evidence ‘suggesting’ or ‘supporting’. There were some students who wrote their conclusion in context in part (b) instead, misinterpreting what was required but this could still score the final mark.

A useful guide to the interpretation of the result of a hypothesis test would be to write ‘There is/isn’t sufficient evidence to suggest an association between A and B’

**(b)** The majority of students scored no marks for this part, usually by comparing drug A and drug B rather than commenting on association. Very few students used $\frac{(O-E)^2}{E}$ or $(O-E)$ to identify main sources of association as demonstrated in sample assessment material.  There was significant non-response.

This indicates that students would benefit from discussing the interpretation of a test for association in context, with guidance about how to identify and describe sources of association between the variables.

## 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.