# Examiners’ Report

# Principal Examiner Feedback

## Summer 2022

Pearson Edexcel GCE

AS Mathematics (8MA0)

Paper 21 Statistics

Summer 2022
Publications Code 8MA0_21_2206_ER*

**Introduction**

Questions 1, 2, 3(a) and 5(a) proved to be accessible to nearly all students taking this examination and question 5(b) offered some discrimination for the more able. Question 3(b) proved beyond most students this time around: it was a first attempt to assess this area of the specification and, being written in the early days of covid, it proved unfortunate that this group of students, with such a disrupted preparation, encountered it.

**Comments on individual questions.**

**Question 1**

Part (a) was answered well with most stating that the correlation was negative or in a few cases giving us a suitable description such as "*t* increases as *p* decreases". It was encouraging to see most giving a suitable description of the units for the gradient in part (b) too although some simply said "days" having failed to appreciate that the gradient represented a rate. In part (c) some simply substituted $t = 3$ into the equation and gained no marks but most used $t = 3$ and $t = 0$ and subtracted to obtain a value of $\pm$ 3.3 A sizeable minority thought that using $t = 3$ and $t = 1$ would give them the change over 3 days. Few though told us though that this change represented a decrease in *p* and so that last mark was rarely scored. The responses to part (d) were encouraging. Most secured the mark by stating that the estimate would be unreliable and either mentioned that the data were bases on values of *t* between 1 and 10 or said that using 19 involved extrapolation. A common incorrect answer argued that because the estimate was still positive when $t = 19$ the estimate was reliable.

**Question 2**

This question was answered well with nearly 40% scoring 6 or 7 marks. In part (a) many selected the correct model and usually stated it correctly though on this occasion we did condone seeing B(0.08, 35) for example. Many students simply write down what appears on their calculator input screen and they should be aware that this is not standard mathematical notation and without a correct answer following it would not gain the method mark. Most also scored the mark for (i) though some failed to give an accurate enough answer with 0.24 being the only answer offered on a few occasions. The front of the paper says that 3 significant figures should be used but we will usually allow any answer which rounds to 3sf (indicated as awrt on the mark scheme.) Part (ii) was not answered so well with some finding $\mathrm{P}(D \ldots 3)$, some thinking that $\mathrm{P}(D > 3) = 1 - \mathrm{P}(D$ „ $2)$ and others unsure about which function to use on their calculator and ending up with $\mathrm{P}(D = 3)$. In part (b) the hypotheses were usually correctly stated using *p* and a correct model selected. Most obtained the correct probability although some found $\mathrm{P}(X = 2)$ rather than $\mathrm{P}(X$ „ $2)$. The correct decision was usually made but some failed to give a suitable contextual conclusion: the simplest description mentioned that there was evidence to support supplier *B*'s claim but other chose to talk about damp bags and often referred to the "number of bags" rather than the "proportion" decreasing.

This type of question is now fairly familiar on these papers but there were a number of students who did not seem to have covered the topic of hypothesis testing and nearly 15% scored zero on this question.

**Question 3**

Part (a) should have been familiar territory for candidates and most did manage to list the frequencies for each of the bars though, rather disappointingly, nearly a quarter of candidates could not do this successfully with a common incorrect answer of 50 rather than 75 being recorded for the bar between 2 and 3.5  Those who did understand this important feature of histograms often went on to use linear interpolation correctly to find the median.

Part (b) sadly defeated almost all candidates.  This was the first time we had attempted to assess the part of the specification dealing with "connect(ion) (of histograms) to probability distributions" (section 2.1).  A small minority of candidates realised they needed to integrate and set the area under the curve to 256 and they often went onto achieve the correct value of $k$ but most attempts involved differentiating the curve or solving a quadratic equation and made no progress.  A few could use the symmetry of the model in part (c) to write down the median but most were completely lost after part (a).

**Question 4**

Although a few students simply left this question blank most had a go though only about a quarter scored either of the marks.  The median was the most successfully answered with 1000 being a common acceptable value.  Selecting a suitable value for the range was more challenging with most responses being far too large.

Although the specification requires students to be "familiar" with the large data set, for this sitting it is understandable that many would not have had as much time for familiarisation as would be desirable.

**Question 5**

Part (a) was answered well by most candidates but often there was little explanation offered to explain where their probabilities came from, nevertheless the correct answers to (i) and (ii) were often seen.  Some thought that, for example, in (i) the answer should be $\frac{3}{4}\times\frac{1}{3}+\frac{1}{3}\times\frac{3}{4}$ but they had at least realised that the red spinner must have landed on 3 and the green spinner on 4 and scored the first mark.  Part (b) was designed as a problem solving question and it was good to see a number of students trying it though understandable that many felt there were easier pickings available in paper 22 and moved onto that.  Most attempts involved setting up two linear equations and solving them but there was some confusion over values of the random variables and probabilities associated with these values.  However over 10% of candidates scored full marks on this question and shoed a good understanding of this topic.