Examiners' Report
Principal Examiner Feedback

October 2020

Pearson Edexcel GCE Advanced
In Statistics (9ST0)
Paper 1: Data and Probability (9ST0_01)

Edexcel and BTEC Qualifications

Edexcel and BTEC qualifications are awarded by Pearson, the UK's largest awarding body. We provide a wide range of qualifications including academic, vocational, occupational and specific programmes for employers. For further information visit our qualifications websites at www.edexcel.com or www.btec.co.uk. Alternatively, you can get in touch with us using the details on our contact us page at www.edexcel.com/contactus.

Pearson: helping people progress, everywhere

Pearson aspires to be the world's leading learning company. Our aim is to help everyone progress in their lives through education. We believe in every kind of learning, for all kinds of people, wherever they are in the world. We've been involved in education for over 150 years, and by working across 70 countries, in 100 languages, we have built an international reputation for our commitment to high standards and raising achievement through innovation in education. Find out more about how we can help you and your students at: www.pearson.com/uk

October 2020
Publications Code 9ST0_01_2010_ER

General introduction

Paper 1 proved

Question 1

In general, the advantages in the (i) items were done well by most candidates, and most candidates were able to score at least one mark for a source of bias somewhere in the question. Some candidates missed that **two** sources of bias were asked for, so it is essential that questions are read carefully, perhaps with the word 'two' being underlined as a habit.

Candidates are advised to engage with the context, e.g. 'not representative of the whole sample' is not going to score any marks as a source of bias.

Further, answers should be sensible and realistic in this context. E.g. 'the receptionist collecting the questionnaires may change the answers to suit them', while a potential source of bias if a shady character was trying to demonstrate public support for a venture, is highly unlikely to happen in this context when the whole drive was to find out what people think.

Question 2

Candidates found question 2 challenging, with many not engaging with the properties of a Poisson distribution. Three marks were available for simply stating that the suitability is dependent on whether buses pass (or meteorites hit the atmosphere) independently at a constant rate, so candidates are advised to demonstrate their knowledge even if the context is challenging.

Question 3

In (a), many candidates had not acknowledged the preceding paragraph which stated that Sylvia had determined **suitability** of the normal model

through shape and a test for normality. So, further comments on shape could not gain full marks. Candidates are reminded to read questions fully and carefully.

The introduction of percentiles in (b) and (c), though in the specification and defined in the question, caused some confusion for candidates.

The interpretation in (d) and (e) was challenging, though many good explanations were provided, and some candidates who struggled with the earlier parts were able to gain some marks with sensible comments here. Candidates are advised to attempt all parts of a question, even if they feel they have struggled in previous parts.

Question 4

This question was done well by many candidates. However, a number of candidates interpreted the list of fields (table structure) as the dataset in question. Candidates are reminded to read the question fully and carefully, and to assess whether their understanding makes sense in context. For instance, a pause to check whether this table does indeed include 'detailed records of every incident the fire service responds to' should have confirmed that it does not.

A number of candidates also did not understand the difference between a database and a spreadsheet.

Question 5

This question was done well by most candidates, despite this being the first instance of a Venn diagram with four sets. Almost all candidates managed to fill in the diagram in (a) with at most one error, and the probabilities in (b) were successfully calculated by most. The interpretation in (c) was also done well.

A majority of candidates understood to use the diagram to count card inside and outside each set in (d), though few were able to extend this to multiple sets in the challenging final part, (e).

Question 6

Part (a) was a standard normal distribution question, which was found to be a challenge for a number of candidates, and as a result many candidates did not even attempt the later parts of this question.

Part (b) introduced a more mathematical approach than candidates will have been used to, which was necessary to solve the problem given in the context. However, most students were able to piece together the correct mean, and many correctly extended the formula book expression to get the correct variance.

Candidates struggled with (c), most likely due to the use of calculators which usually sidesteps the standardisation formula. However, candidates are reminded that standardising is required for questions with a missing mean and/or variance, and indeed in this question where we are trying to determine how many distributions are required.

Most candidates did not even attempt (d), despite it being the simplest part, mathematically. Candidates are advised to attempt every part of the question, even if they have found earlier parts challenging.

Question 7

This question was generally done well. Almost all candidates managed parts (a) and (b) without issue, and most managed to interpret the context sufficiently in (c).

Part (d) caused more trouble, with many students struggling to interpret the written information and convert into meaningful mathematics. Candidates are reminded that basic mathematics skills are often required

in Statistics. A list of potentially required skills can be found in Appendix 3 of the specification.

In part (e), most candidates were able to identify that London was too far away to sensibly do a weekly shop, but only a limited number thought to insert London into the model to investigate the ramifications of this.

Question 8

Candidates found this question very challenging, and a significant proportion did not even attempt it. Students are reminded that a tree diagram is an acceptable method to answer Bayes' theroem questions, and in this case a correct tree diagram with a clear attempt to find the correct probability would have gained half of the marks.

Question 9

Some candidates struggled with the open nature of this question. This style of question will be used more regularly in the new 2017 specification. Several candidates only attempted one of the two tasks, suggesting either that they hadn't read the question in full, or that they found the other part too challenging. Most struggled to fully understand the data in the table, though this was not heavily penalised and so candidates were rewarded for perseverence.

Summary

Based on their performance on this paper, candidates should:

- read questions fully and carefully.
- try their best to fully engage with the contexts, with explanations realistic and sensible within that context.
- ensure they are comfortable with the mathematical skills in Appendix 3 of the specification.
- try later parts of a question, even if they have struggled with earlier parts.

- ensure they are familiar with standardising a normal distribution, as this will be required for certain questions.