Pearson Edexcel

Principal Examiner Feedback

Summer 2018

Pearson Edexcel GCE Mathematics

In Statistics S2 (6684) Paper 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

Summer 2018
Publications Code 6684_01_1806_ER

Introduction

On the whole this paper was well answered. Questions 1 to 5 on the paper seemed reasonably accessible to the majority of students. Question 6 was a little unstructured which meant that there were several pages of messy responses which poor mathematical notation used making it difficult to follow. Other than Q6 generally the work was quite well presented.

The paper proved accessible to the majority of students and there was little evidence of there not being enough time to complete the paper. There were the usual arithmetic and algebraic errors, but it was clear many students were out of practice in techniques, taught in Statistics 1 such as using an inverse normal distribution and using conditional probability.

Question 1

This question proved popular with students able to score the majority of the marks.
In part (a) the majority of students were able to find $\mathrm{P}\left(X > 5\right)$ and scored full marks. The few that made errors calculated $1-\mathrm{P}\left(X \leq 4\right)$ rather than $1-\mathrm{P}\left(X \leq 5\right))$
Fewer students were able to provide a correct solution to $\mathrm{P}\left(4 \leq X < 10\right)$ There was a wide range or errors made where some students seemed to really struggle to identify the correct method. The most frequent mistake was to calculate $\mathrm{P}\left(X \leq 9\right)-\mathrm{P}\left(X \leq 4\right)$ but $\left(1-\mathrm{P}(X \leq 3)\right)-\mathrm{P}(X \leq 9) = 0.9788 - 0.5874 = 0.3914$ was also a common mistake.

There were many completely correct answers to part (b) that earned all five marks. Such responses were typically confident, accurate and efficient. The most common error concerned the continuity correction: either absent or the incorrect version. A relatively infrequent error was the result of truncating rather than rounding the *z*-value 0.43719…. to 0.43 so losing the final accuracy mark in part (b).

In part(c) the majority of students recognised that the relevant distribution in part (c) was B(5, "0.33") to gain the first method mark but often there was a loss of marks due to not showing sufficient workings or for making an error calculating one or more of the probabilities. There were two strategies: $\mathrm{P}(W=4)+\mathrm{P}(W=5)$ and $1-\mathrm{P}(W \leq 3)$. The former is simpler: the Binomial formula has to be used twice only, compared to four times in the second method. The latter method therefore provided more opportunity for arithmetical error. It was not uncommon for students to confuse p and $(1 - p)$ when using the Binomial formula. There were some students who failed to realise that they should be finding $\mathrm{P}\left(W \geq 4\right)$ and only calculated a single probability such as $\mathrm{P}(W=4)$ or $\mathrm{P}(W=5)$.

Question 2

Overall this question was answered well by the majority of students. Some showed a good amount of working throughout, but others only showed minimal method.

A large majority of students could state at least one reason why a Binomial distribution is appropriate in this case in part (a). It was common to see two or even three valid reasons. However, many students lost marks unnecessarily as they failed to realise that they needed to answer a contextual question in context. Only a very small number of students provided inappropriate reasons, such as “occur singly” or “constant rate”.

The overall response to part (b) was excellent, with almost all students scoring both marks. A number of different methods were seen but the Binomial formula was more popular: the tables and $\mathrm{P}(T \leq 5) - \mathrm{P}(T \leq 4)$ being rarely used.

Part (c) should have been straightforward, but not quite as many successful solutions were seen as might have been expected. Nearly all students understood what they were supposed to be looking for, and many found the probability correctly. However, some had issues identifying the ‘tail’ correctly from the table, so would attempt $1 - \mathrm{P}(T \leq 4)$ using the tables. Others decided to calculate the individual probabilities and add them up, making errors in their calculations.

Part (d) proved to be more challenging with some students having problems identifying how many heads were needed to meet the criteria in this part of the question. It was not uncommon for students to identify the relevant events by writing $\mathrm{P}(H \geq 3)$, but then use the incorrect method $1 - \mathrm{P}(H \leq 3)$ to find it. Some students used *X* or *Y* instead of *H* or *T*, and most of the time this wasn’t an issue but a small number used *T* when they meant *H* which was. A surprisingly large number decided to add the probabilities for 3, 4, 5, and 6 Heads together rather than do $1 - \mathrm{P}(H \leq 2)$; even those who did the latter did not always use the tables.

Question 3

This was a long and slightly challenging question for some students but those students who laid out their work clearly and communicated their reasoning methodically were most successful.

A large majority of students were not only familiar with the theory but also able to implement accurately the required definite integration in part (a) with there being only very occasional errors with arithmetic. There were, however, a very small number of students whose entire working for the first integral, for example, consisted only of: $\int_1^2 \frac{1}{2}t(t-1)dt = \frac{5}{12}$ ignoring the instruction to “use algebraic integration” and this resulted in a loss of marks.

Relatively few students made the predictable error of integrating f(t) for both parts although it was surprisingly common to see students perform only one of the two integrals. Although a few errors were made in the integration most resulted from inaccuracies when it came to substituting in the limits. Many students did not show the substitution clearly and in these cases arithmetical errors were often made. Only a minority of students correctly evaluated both integrals and stopped before adding them.

Almost all students earned at least one mark in part (b) with only a tiny number of students being unfamiliar with the formula for variance.

In part (c) most students correctly integrated f(t) from 1 to t or used F(1) = 0 and gained full credit for the region $1 < t \leq 2$. Some however made calculation errors while others neither used limits nor a boundary condition.
For the region $2 < t \leq 4$ many students were able to demonstrate a correct method either integrating from 2 to t and adding F(2) from the line above or using a boundary condition F(2) = 0.25 from the line above or F(4) = 1 to find the arbitrary constant. There were a significant number of students who incorrectly only integrated f(t) from 2 to t.
The final B mark was often lost as the function was not stated using only t.

There were many instances of full marks for part (d). The main errors were

- Using the incorrect part of F(t) (i.e. $2 < t \leq 2$).
- Using an 'amalgamated' version of their F(t): adding both parts together.
- F(t) in part (c) was incorrect.
- Incorrect rearrangement of F(t) = 0.2
- Evaluating f(0.2)

Once again the method required for part (e) was clearly understood and generally a correct solution followed. The main errors were using f(t) instead of F(t) or forgetting to subtract F(1.5) from 1

In part (f) the correct method was clearly used by many students. However, it was not always clear whether students were using P(T > 3) in the numerator. It is advisable to show all working as method marks could be gained even if errors were made in previous parts.

Question 4
Part (a) was done well by the majority of students. They wrote down the equation for the mean and variance of the continuous uniform distribution and solved them simultaneously. It was pleasing that the majority of students recognised that if $(\beta - \alpha)^2 = 144$ then $(\beta - \alpha) = 12$ and solving two linear equations. Students choosing the alternative route and formed a quadratic equation in either α or β were less successful than those who formed two linear equations.

Part(b) was not done well. Most students were able to find the late train probability, but $0.05 + 0.95 \times \frac{5}{12}$ was rarely seen. Two out of the three numbers were often seen, the most common by far being $0.05 + \frac{5}{12}$, but $0.05 \times \frac{5}{12}$ was also seen. Some students simply wrote 5/12 as their answer.
A few students used tree diagrams, and this proved to be an effective approach.

Most of the large number of students with incorrect final answers to part (b) then earned one mark in part (c) by writing $\frac{0.05}{(b)}$, i.e. using their answer to part (b) as the denominator of the fraction. However, there were a few incorrect variations: $\frac{(b)}{0.05}$, $\frac{0.05}{(b)^2}$ and $\frac{0.05 \times (b)}{(b)}$

Question 5

Part (a) was a routine question and the source of full marks for many well-prepared students. The hypotheses were invariably written correctly however far too many students were unable to calculate the test statistic correctly. The most common method was to find $\mathrm{P}(X \geq 18)$ however many simply calculated $\mathrm{P}(X > 18)$ or just $\mathrm{P}(X = 18)$.
Those who opted for the CR approach often lost marks. It was common to see $\mathrm{P}(X \geq 19) = 0.0699$ but with no critical region stated or stating it incorrectly: $V \geq 19$ was a common error. However, the probability approach was by far the most commonly used method.
The vast majority of students who were successful on this question were also able to state a conclusion in context using the wording of the question.

Many students ~~were~~ scored at least 2 marks in part (b) for their use of B(50,0.35) and the lower tail although the upper tail was often incorrectly stated as $X \geq 24$.
There are still a number of students who are incorrectly using probability statements as their CRs.

The majority of students gave a correct statement including the required context to gain the mark in part (c).

Those who scored full marks (or even just two marks) in part (c) were generally able to calculate the correct significance level in part (d)

The overall response to part (e) was generally sound, although full marks was relatively rare. The Normal distribution was handled well by almost all students. However, a few errors in detail occurred during the working; a not inconsiderable number went down the Poisson approximation route first, resulting in them gaining the incorrect variance. The most frequent problem concerned the continuity correction. Inevitably, some students used the continuity correction *(n* + 0.5), while others omitted a continuity correction altogether. A tiny number of over-zealous students used the continuity correction (*n* - 1). There were many incorrect final answers that followed a method that was sound in principle. Some students were confused about the direction of the inequality, which led to the wrong sign before 1.68
The other problems at this last stage relate to rounding. Some students rounded down from a correct 30.9964…., or the common incorrect values 30.4964…. to give a final answer of 30. Of more concern, is the not insignificant number of students who rounded the correct penultimate value of 30.9964 to 40

Question 6
In general, question 6 was poorly answered, with only the most able students able to produce a succinct fully correct response. Students should be encouraged to use correct mathematical notation throughout their work. In particular, they should write down “labels” for their expressions, rather than merely writing the expressions themselves.
Although stronger students produced well-presented work, examiners were frequently faced with several pages of messy responses, often going on to additional sheets.

In part (a) most students realised that differentiation was needed and most knew to differentiate twice and even to equate to zero and substitute *x* = 8/3. A reasonable portion of students only differentiated once and then began a series of confused working. Those that differentiated twice were knowledgeable enough to continue with the full method. However, a large number lost the final A1 as they prematurely cancelled *k* or ignored it completely or featured some other incorrect working.

Part (b) caused a great deal of confusion as it was somewhat oblique and not like many previous questions. This led to a variety of different approaches and a lot of confused working.

A common error was to mistakenly take b to equal $\frac{4}{15}$. This led to these students wasting a lot of time. Therefore, such students were unable to write down a correct equation in the two unknowns' k and b, leading to a good deal of work that gained no credit.

More students were able to use F(4) = 1 rather than F(2) = $\frac{4}{15}$, instead many used F(2) = 0

Many students did go on to score the final method mark for a correct expression including their values of k and b for F(2.5).
The majority of students used the simultaneous equations method. However, a not inconsiderable number used the alternative integration method.

Pearson Education Limited. Registered company number 872828
with its registered office at 80 Strand, London, WC2R 0RL, United Kingdom