Pearson Edexcel

Examiners' Report

Principal Examiner Feedback

Summer 2022

Pearson Edexcel GCE

AL Further Mathematics (9FM0)

Paper 3B Further Statistics 1

Summer 2022
Publications Code 9FM0_3B_2206_ER*

Introduction

The paper was generally very accessible with a mean score of over 50. All questions allowed students to get started and the later parts of questions 3, 6 and 7 provided some discrimination. A few examiners did report that the handwriting on some scripts was very poor and this sometimes makes it hard to decipher their answers and give credit. Whilst we appreciate the time pressures of examinations and fact that this was the first time these students had faced high stakes public examinations it is worth reminding students in future of the importance clarity in their solutions.

Question 1

This provided a good start to the paper with over 60% of students scoring full marks. Nearly all answered part (a) correctly but there were a few errors occasionally seen in part (b). Some failed to state the parameters of the model in the hypotheses and very occasionally the hypotheses were the wrong way around. The calculation of the test statistic was generally correct though a few students thought that they needed to pool the final two columns and this, of course, affected their degrees of freedom. The correct conclusion was usually given in context as required.

Question 2

This was another straightforward question with nearly 65% scoring full marks. The principles of expectation algebra were clearly understood well and nearly all gave a correct answer to part (a) and most went on to complete part (b) correctly too. There were some algebraic or sign errors when substituting but few cases of $E(X^2)$ being equated to 34.26 or failing to square $E(X)$. In part (c) most chose to substitute values and compare the two sides of the inequality though there were occasional arithmetic errors seen here. Others chose to solve the quadratic inequality which was usually more reliable in terms of accuracy provided they knew to take the "inside" region. Having established that $X = 0$ and -1 were the values of interest, some students weren't sure how to relate this to the original distribution whilst others who identified the correct probabilities multiplied them rather than adding.

Question 3

This question, together with question 4, required the students to use several different distributions. There is a tendency for students to label each distribution as X when different letters should be used. In future students should be reminded that the specification does expect mathematical notation to be used correctly and that there are marks for clear communication of mathematical arguments; given the disruption to the preparation of these students these issues were overlooked on this occasion.

In part (a) the majority identified and used the correct Poisson distribution though a small number used B(28, 0.4); students should be reminded to look for key words such as "rate" in this type of question. In part (b) most were able to find the probability of more than 1 call for help in a day and then select the correct binomial distribution. However a number failed to use a Poisson approximation to evaluate the required probability and an incorrect answer of 0.140 was often seen. The hypothesis test in part (c) was answered well with most using a suitable parameter for their hypotheses (a few cases of no parameter or p being used were seen) and most using a correct Poisson probability and finding the appropriate probability correctly. There was usually an attempt to give the conclusion in context but this wasn't always precise enough. The test is about the "rate of calls per day" and so a conclusion simply saying that "the number of calls is lower in winter" was not suitable. The final two parts were more discriminating. In part (d) many knew how to combine Poisson distributions and used $\text{Po}(0.6n)$ or $\text{B}(n,\ e^{-0.6})$ and then formed a suitable inequality but some thought that they simply needed to multiply or add the individual probabilities for each rescue team. A few lost the final mark because they misread 0.001 as 0.01 which was unfortunate but both method marks could still be earned. In part (e) the majority of students simply offered a textbook answer without sufficient application to the context of the question.

Question 4

This was a very accessible question with 99% scoring at least 2 marks and nearly 25% full marks. Part (a)(i) was answered very well with most clearly stating the correct geometric distribution and giving the answer to an appropriate degree of accuracy. In (a) (ii) many interpreted the question correctly but some had $1-0.89^6$ instead of $1-0.89^5$. In part (a)(iii) many did not identify that a binomial distribution was required, instead using a negative binomial but those with the correct distribution invariably obtained the correct answer. Most students recognised (a)(iv) as a negative binomial and went on to evaluate some of the 3 probabilities required. Sadly many lost the final accuracy mark here because their final answer was not given to at least 3 significant figures. A number of students knew the connection between the cumulative distribution for this negative binomial distribution and the binomial distribution from part (iii) and this enabled them to reach the correct answer very easily. Part (b) was answered very well with many students identifying that an infinite geometric series of probabilities was required. This was a "show that" question and so we needed to see the correct series clearly used and a correct formula used to find the sum and a statement that their answer did indeed round to the required value. A particularly impressive solution here used the probability generating function G(t) of the geometric distribution used in part (a)(i) and then, with a lovely explanation, simply took the mean of G(1) and G(– 1).

Question 5

The Advanced Information may well have helped students as they should have been expecting a question involving the Central Limit Theorem and over 75% scored full marks here. The correct mean and variance of the given geometric distribution were nearly always found but a few failed to divide the variance by 150 when forming the distribution for the mean. There was some confusion over the correct notation: we expected to see $\bar{X}$ and $\text{N}\left(\frac{10}{3}, \frac{7}{135}\right)$ but often X was used for both Geo(0.3) and this normal distribution and occasionally the normal distribution was stated as $\text{N}\left(\frac{10}{3}, \sqrt{\frac{7}{135}}\right)$ but if this was used correctly we condoned the incorrect notation.

Question 6

This proved to be a very accessible test of the probability generating function topic with nearly 95% of students gaining 5 or more marks and over 30% full marks.

Part (a) was usually answered very well but careless factorisation meant some wrote $t^2\left(\frac{2}{5}t+\frac{3}{5}\right)$ or $t^2\left(\frac{2}{5}t^2+\frac{3}{5}\right)^2$ as part of their simplification. In part (b) most knew how to find E(*W*) and Var(*W*) but poor differentiation, usually a failure to use the product rule correctly, meant that many failed to score full marks here. There was a specific demand to "use calculus to find" and some simply wrote down the answers without showing any working and so gained very few marks for this part. Part (c) was answered very well with most knowing they needed to multiply the two probability generating functions and only a small minority failing to simplify their answer correctly. Part (d) was more challenging, most realised they had to multiply by t^3 but often they only substituted t^2 for t inside the bracket. Most students who had answered part (d) usually used their $\mathrm{G}_Y(t)$ correctly to find $\mathrm{P}(Y = 15)$ and even those with an incorrect $\mathrm{G}_Y(t) = t^6\left(\frac{2}{5}t^2+\frac{3}{5}\right)^7$ could obtain a mark for stating that $\mathrm{P}(Y = 15) = 0$. Some found $\mathrm{P}(X = 6)$ instead which was, of course, fine.

Question 7

This was the most difficult question on the paper and just over 10% of candidates made no progress however nearly 30% did achieve 10 or more marks.

A few students seemed surprised at seeing a normal distribution question on this paper (despite handling it comfortably in question 5) and one or two tried to change it to a discrete distribution. Most though realised that this was a standard hypothesis test from the A level specification and were able to find the correct probability and state that it was significant though some lost the last mark for failing to express their conclusion precisely enough in context. Part (b) proved a little more difficult but a good number of students knew how to go about finding a critical value though sometimes the *z*-value used was not accurate enough (students should either use the value in the tables or show in their subsequent working that they have used a value that is at least as accurate as the tables). Some did not heed the requirement to find an equation for *c* in terms of *n*. The mention of a Type II error clearly stumped some and so part (c) was often not attempted. Those who knew what a Type II error was could usually make a correct start but the common problem was failing to appreciate that if $\mathrm{P}(Z < a) = 0.0050$ then *a* must be negative and many solutions used a *z*-value of + 2.5758 rather than – 2.5758. A number struggled with the algebraic manipulation required to isolate *n* or *c* and a few did not realise that *n* needed to be an integer and some forgot to state their value of *c*.