Pearson

Examiners' Report

Principal Examiner Feedback

Summer 2017

Pearson Edexcel International A-Level in Statistics (WST02)

edexcel

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 2017
Publications Code WST02_01_1706_ER
All the material in this publication is copyright
© Pearson Education Ltd 2017

IAL Mathematics Unit Statistics 2

Specification WST02/01

General Introduction

The paper was accessible to all students but there were several places (Qu 1(b)(c), Qu 2(b) and Qu 6) where the "switching" of random variables made the questions more demanding. As usual parts of questions that required a comment or explanation in words were often not answered very well.

Reports on Individual Questions

Question 1

This question was answered well and nearly 20% of the students scored full marks. In part (a) whilst most students used the correct Poisson distribution the common error was to use Po(1). In part (a) students were dealing with the random variable X representing the number of signal failures on a day and needed to find $P(X = 0)$. In part (b) many thought they were dealing with a new random variable representing the number of signal failures in 3 days and used Po(0.75). It might be helpful for students to think of the new random variable Y representing the number of <u>days</u> out of 3 where $X \geqslant 1$ and even writing down $Y \sim B(3, p)$ where $p = P(X \geqslant 1) = 1 - P(X = 0)$. Part (b) then simply requires $P(Y = 3)$. Whilst for many part (b) was very straightforward the above analysis may help those who didn't "see" what to do in (b) or (c) how to identify similar questions in future. Part (c) was perhaps more obviously another binomial distribution of the form $B(7, k)$ but many failed to use a correct probability for k. Some persisted in the use of a Poisson distribution and tried to use Po(1.75). Part (d) was usually answered well with most using μ or λ to describe both hypotheses but some only gave H_0. Part (e) was answered very well but a common error was to find an upper tail (often despite having a correct H_1) rather than the lower tail.

Question 2

Most students answered part (a)(i) correctly but some just wrote "binomial" without specifying the value of n or the value of p and occasionally they used incorrect notation such as B(0.25, 6). In part (ii) few applied their "textbook" assumptions accurately to the context. Part (b) introduced another switch of random variables. Having defined X = the number of prizes in a box, where $X \sim B(6, 0.25)$, in (a)(i) they now needed to think in terms of a different random variable $Y \sim B(2, P(X = 1))$ where Y = the number of boxes containing exactly 1 prize. Few explicitly defined this second binomial distribution and many students just stopped after finding $P(X = 1) = 0.356$. Those who did attempt to address the matter of the 2 boxes invariably failed to multiply by 2. Fortunately most could recover in part (c) and this was answered very well with most interpreting the "at least 2" phrase correctly. The normal approximation in part (d) presented the usual challenges for some students. The first hurdle was identifying the correct binomial distribution with B(480, 0.466) or B(80, 0.25) being common errors. Some identified the correct binomial distribution but then approximated it to a Poisson first and then from the Poisson to a normal thus failing to get the correct variance. Most could standardise using their normal distribution but some forgot to use a continuity correction and those that did sometimes used 29.5 rather than 30.5. The final hurdle was to take the value of 0.9357 found from tables away from 1 but there were plenty of students who did arrive at the correct answer to this part.

Question 3

Whilst over a quarter of the students gained full marks on this question most found parts of this question quite challenging. Part (a) was particularly poor with many simply observing that f(4) = 0.5 or solving f(x) = 0.5 and arriving at x = 4. Those who did realise that an area needed calculating, generally used integration rather than simply showing that the area of the triangle, defined by the second branch of the pdf, was 0.5 and deducing the median from this. Some chose to find the cumulative distribution function first and use this to establish the median; care had to be taken though to avoid "circular" arguments that used F(4) = 0.5 but there were some successful solutions using this approach. In part (b) many students used the area of the triangle defined by the first branch of the pdf to form one equation but they didn't always realise that a second equation was given by using f(1) = 0 and an incorrect equation based on f(4) = 0.5 using $ax + b$ was sometimes used. Part (c) was usually answered with many students finding 3 of the 4 lines of the cdf correctly. The problem, of course, was the 3rd line for $4 \leqslant x \leqslant 6$. Those using an integral with limits of 4 and x invariably forgot to add on F(4). The "+ c" approach was often more successful with students using F(6) = 1 or F(4) = 0.5 to establish the value of c.

Question 4

In part (a) the mean was usually correct but many students failed to score the marks for the standard deviation. There were two typical errors: giving the variance instead of the standard deviation and using $\sqrt{3.5}$. Part (b)(i) was not answered well as many students approximated to a Poisson distribution too early and gave hypotheses in terms of μ or λ rather than using p for the underlying binomial. Most knew what to do in part (ii) but some used 1 – P($X \leqslant 5$) rather than 1 – P($X \leqslant 4$) and others did not give their conclusion in context. A few used a normal approximation which was inappropriate in this case since the value of p was not close to 0.5. In part (c) many students used the tables to find the value of n. The tables do not contain sufficient values of n for them to be sure that n = 15 is the correct answer and so this was not an acceptable approach. Those who used P(Y = 0) = $(0.7)^n$ and compared this with 0.005 were easily able to show that n > 14.85 and hence could deduce that n = 15. The tables could of course be used to establish that w = 10 and many students achieved this.

Question 5

This question was answered very well with nearly 40% of students gaining full marks. In part (a) some students found F(4) but did not subtract this from 1 but usually this was answered very well. Part (b) proved more challenging and many did not identify the conditional probability. Those that were attempting $P(T > 4 \mid T > 2)$ often had a numerator of $P(T > 2) - P(T > 4)$. In part (c) most students compared F(2.7) and F(2.8) with 0.75 but the conclusions were sometimes not precise enough to gain the final mark. Comments such as "therefore the upper quartile is 0.75" or " $0.73 < Q_3 < 0.752$" are incorrect. Some students used their calculator to solve $F(x) = 0.75$ but we required an answer that was correct to 3 sf and all other solutions to be rejected and some just truncated their calculator value (giving 2.78) or failed to reject the other solutions. Part (d) was answered very well with many students scoring full marks. There were some cases of poor use of brackets where students attempting $tf(t)$ had $0.3t - 0.12t^2$ instead of $0.3t - 0.12t^3$ but the integration and use of limits was handled well.

Question 6

Most students understood what was required in part (a) but a small minority failed to score these marks either because they ignored the instruction $M \neq 5$ and attempted to list all the samples of size 3. In part (b) it was quite common to see students restricting the values of M to just 3 and 4 and a few just attempted the sampling distribution for the mean rather than the maximum. The probability $P(M = 3)$ was usually correct but invariably one or more of the other two probabilities were incorrect but there was usually an attempt to make sure that the sum of the probabilities came to 1. Part (c) proved more demanding than anticipated with 3 and 3 being a common incorrect answer and many simply not attempting this part.

Question 7

This question was answered well by many students but the algebraic demands of part (b) proved too tricky for some. Most knew what to do in part (a) and apart from the inevitable signs errors, due to improper use of brackets, it posed few problems. Part (b) proved more challenging with many failing to form a correct expression: some forgot to subtract $\left(\frac{1}{3}b + \frac{2}{3}a\right)$ from b on the numerator and others trying a 1 – …. approach failed to subtract a from $\left(\frac{1}{3}b + \frac{2}{3}a\right)$. Some students reached $\dfrac{\frac{2}{3}b - \frac{2}{3}a}{b-a}$ but could not see how to simplify this. The most successful approaches in part (c) usually started by finding $E(X^2)$ and then dealing with $3X^2$. Most used the variance formula rather than integration and usually they realised that $E(X) = 0$ told them that $b = -a$ and used this to reach the correct answer. Apart from the small minority who didn't realise that the range was $b - a$, part (d) was answered well either by using $\dfrac{18^2}{12}$ or first finding $b = 9$.

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