Examiner's Report
Principal Examiner Feedback

Summer 2018

Pearson Edexcel International A Level
In Statistics S2 (WST02/01)

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: WST02_01_1806_ER

Introduction

This paper proved to be very accessible to all students though questions 2(b), 3(e), 4 and parts of question 6 were more discriminating.

Comments on individual questions

Question 1

This question was answered well. The probabilities in part (a) were usually correct although some struggled to interpret "at least 3" correctly and found $1 - \mathrm{P}(M \leqslant 3)$. Most answered part (b) correctly and it was only in part (c) that a number of students lost marks. Some worked with equality all the way through rather than using the correct inequalities and this meant they would lose the final two marks. Those who did offer a fully correct solution handled the change of inequality when dividing by log0.95 well and avoided the incorrect answer of 89 that others gave.

Question 2

Most students made a correct start in part (a) stating suitable hypotheses and many went on to calculate the correct probability but a common error was to find $\mathrm{P}(X = 14)$ rather than $\mathrm{P}(X \geqslant 14)$ or use Po(4) rather than Po(8). Those who found the correct probability usually rejected the null hypothesis but they did not always give a correct conclusion in context. Part (b) proved to be more challenging and many did not realise they needed to be using a mean $\lambda = \frac{4l}{50}$. Those who started off correctly usually used logs to find a suitable value for l but in order to show the answer was 1.3 to 2 significant figures we needed to see at least 3 significant figures first and some therefore lost the final accuracy mark. Part (c) was much more straightforward and many students had a fully correct solution here.

Question 3

There were a variety of approaches used in part (a) and most students gave sufficient evidence to successfully "show" the required result. The majority used an algebraic approach starting with $\frac{102-100}{k-100} = \frac{1}{3}$ or $\frac{k-102}{k-100} = \frac{2}{3}$ and solving to find k. In (b)(i) most gave a correct answer but in (ii) many gave an answer of $\frac{1}{6}$ presumably confusing the situation with a discrete uniform distribution. Most gave a correct answer to (c) and part (d) was usually correct too. Part (e) proved more challenging and few students formed a correct equation with sign errors usually being the cause of the problem.

Question 4

This type of question is never popular with international students and so it was encouraging to seea good number attempting all the parts though answers were often not quite precise enough to secure the mark. There was some success in part (a) but many simply said that the cartons were picked at random and made no reference to all cartons having the same chance of being selected or to the selection being without bias. In part (b) the crucial idea of having all the cartons of milk from the dairy was often missed and in part (c) whereas many realise that a normal distribution was required few realised that it was N(0, 1). There was a little more success with part (d) but the most successful part was (e) and many students scored both marks here.

Question 5

Part (a) was answered well by the majority of students but in part (i) some failed to appreciate that the tables gave $P(X \leqslant 7)$ rather than $P(X = 7)$. The usual problem of interpreting "more than 7" occurred in part (ii) with some using $1 - P(X \leqslant 6)$. Part (b) was more challenging and provided a good discriminator for the A grade students. Many realised that the starting point was a Poisson distribution (though a few used a normal approximation to $B(n, 0.6)$) and could state an appropriate normal distribution. Students then fell at the various hurdles on the way through: some forgot the continuity correction or used it incorrectly, most could standardise and many had a correct z value but forming and solving a correct quadratic equation was quite challenging and only the strongest students arrived at $\sqrt{\lambda} = 6$. Of course at some stage they needed to realise that □ = $0.6n$ and this step was applied at various stages in the solution, sometimes making the working considerably more complicated.

Question 6

The integration requirements of this question were straightforward and this meant that students who understood the statistical requirements were often able to do well. In parts (a) and (b) most knew how to deal with a "split" probability density function and most marks were lost through careless errors though a few did not integrate $x\text{f}(x)$. The usual problem of forgetting to subtract the square of the mean occurred in part (b) but many obtained the first 6 marks successfully. The problem in part (c) was having a correct method for finding $F(x)$ for $1 \leqslant x \leqslant 2$. Whilst many could integrate $\frac{x^3}{5}$ they failed to include a "$+C$" and use a suitable boundary condition such as $F(2) = 1$. The other "rows" of the cumulative distribution function were usually correct. Most students could form a suitable equation in part (d) using the appropriate part of their $F(x)$ and they could usually use this value, along with their mean from part (a), to comment on the skewness. A few gave a clear sketch to support their

answer in (e) and others, unnecessarily, calculated the quartiles to justify their description of the skewness. Some stated that the median was less than the mode but without any indication of the value of the mode this did not score the marks.

Question 7

In part (a) most students stated the correct binomial distribution and knew what to do but hadn't read the question carefully enough to realise that they were asked to state the relevant probabilities as well as give the critical regions. The usual problem of writing a critical region as $\mathrm{P}(X \leqslant 3)$ rather than just $X \leqslant 3$ meant some lost the marks here too. Part (b) was usually correct or a correct follow through of their probabilities, if stated, from part (a). Part (c) was usually answered very well with an encouraging number using the context in the question to give their final conclusion.