edexcel

Examiners' Report

Summer 2015

Pearson Edexcel International Advanced 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 2015
Publications Code IA042722

Mathematics Unit Statistics 2

Specification **WST02/01**

General Introduction

This paper proved to be a good test of Statistics 2 material and discriminated well across students of all abilities. There was some opportunity for E-grade students to gain marks in at least 6 out of the 7 questions on this paper. The work on hypothesis testing was generally very good with most students using population parameters and giving their conclusions in context.

In summary, Q1(a), Q1(c), Q2(c), Q2(d), Q3(a), Q5(a), Q6(a), Q6(b), Q6(e), Q7(a) and Q7(b) were a good source of marks for the average student, mainly testing standard ideas and techniques; and Q1(b), Q2(a), Q2(b), Q3(b), Q5(b), Q5(c) and Q6(c) were discriminating at the higher grades. Q3(c), Q6(d) and Q7(c) proved to be the most challenging questions on the paper.

Report on individual questions

Question 1

This was a well-answered question with a majority of students scoring at least 6 of the 10 marks available. Only a minority of students, however, were successful in answering part (b).

In part (a), most students gave a correct solution by applying $1-\mathrm{F}(4)$. Common errors in this part included finding $\mathrm{F}(4)$ or writing down $\mathrm{P}(X>4)=1-\mathrm{F}(3)$.

In part (b), many students used the probability statement $\mathrm{P}(3<X<a)=0.642$ to write $\mathrm{F}(a)-\mathrm{F}(3)=0.642$, which was usually manipulated correctly to give $\mathrm{F}(a)=0.892$. While some students found a correct $a=4.73$ from solving $\frac{1}{5}(2a-5)=0.892$, a considerable number did not use the hint in part (a) and found the value of a by solving an incorrect $\frac{1}{20}(a^2-4)=0.892$. Some students showed a poor understanding of continuous functions by using $\mathrm{P}(3<X<a)=0.642$ to write $\mathrm{F}(a-1)-\mathrm{F}(3)=0.642$; whilst others incorrectly applied the probability statement to give $\mathrm{F}(a)-(1-\mathrm{F}(3))=0.642$.

In part (c), most students differentiated each part of $\mathrm{F}(x)$ correctly and wrote down a fully defined probability density function with the required limits. Some students lost the final mark for this part for defining $\mathrm{f}(x)=1$, when $x>5$.

Question 2

This was a well-answered question with a majority of students scoring at least 10 of the 15 marks available. It was clear that there were a significant number of students who struggled with parts (a) and (b), but produced fully correct solutions to parts (c) and (d).

Part (a) proved to be challenging with some students unable to write down the probability required and a minority incorrectly believing that the question required them to find $\mathrm{P}(X\neq 1)$. Although, a significant minority correctly wrote $\mathrm{P}(X\neq 8)$ and manipulated it to $1-\mathrm{P}(X=8)$, some manipulated it to either $1-\mathrm{P}(X\leqslant 8)$ or $1-\mathrm{P}(X\leqslant 7)$.

In part (b), a majority of students progressed as far as finding $\mathrm{P}(X\geqslant 8)=0.547$, but only a minority achieved the correct answer by applying $(0.547)^4$.

In part (c), most students gained full marks for using a Normal approximation with correct working. Marks lost in this part were mainly due to students using 74.5 instead of 75.5 or students not using a continuity correction.

Many students gave fully correct solutions to part (d). Common errors included stating hypotheses incorrectly; using $\text{Po}(1.5)$ to find their probability instead of $\text{Po}(6)$; finding $\text{P}(X = 11)$ rather than $\text{P}(X \geqslant 11)$; conflicting non-contextual conclusions and not giving a correct conclusion that refers to "rate of sales" or "mean number sold".

Question 3

This question discriminated well across the higher ability students, with the majority of students scoring at least 6 of the 11 marks available.

In part (a), the majority of students gained both marks by correctly sketching $\text{f}(x)$ and using correct labels c, $2c$ and $\frac{1}{c}$. Some students omitted one of the three labels from their sketch, whilst others wrote down the label of c instead of $\frac{1}{c}$.

Part (b) was either answered well or done poorly. Many fully correct solutions were seen. Only a few students used Alternative Method 1 or Alternative Method 2 (as detailed in the Mark Scheme) and when they did they were usually successful in gaining full marks. Those that did not score full marks usually scored one mark for finding $\text{E}(X) = \frac{3c}{2}$. Despite the clear instruction to "use integration" a significant minority of students quoted and used the formula $\text{Var}(X) = \frac{(\beta - \alpha)^2}{12}$. Common errors in part (b) included writing $\text{E}(X^2) = \int_c^{2c} \frac{1}{c}(c^2)\,\text{d}x$; using an incorrect expression for $\text{f}(x)$; or attempting to apply $\text{Var}(X) = \text{E}(X^2) - \text{E}(X)$.

Part (c) was not answered well by the majority of students with a significant number leaving this part blank. Many students were unable to begin this part by writing down a probability statement involving the random variable X. Those that wrote $\text{P}(X > 2(2c - X))$ usually proceeded to the correct answer of $\frac{2}{3}$.

Question 4

This was the most discriminating question on the paper with about 25% of students scoring full marks and about 33% scoring 0 marks.

Although many students stated a correct $k = 4$ in part (a), a significant number did not justify this answer by stating the relevant probabilities of $\mathrm{P}(X = 0) = 0.0183$ and either $\mathrm{P}(X \geqslant 9) = 0.0214$ or $\mathrm{P}(X \leqslant 8) = 0.9786$. Even though the question stated that k is a postive integer, some students worked with $\mathrm{Po}(3.5)$ and $\mathrm{Po}(4.5)$ distributions and sometimes stated their answer as either $k = 3.5$ or $k = 4.5$.

In part (b), many students were familiar with the method of adding together the probabilities from each of their tails. Some gained both marks in this part, even if their solution to part (a) had failed to gain full credit. Few students, however, incorrectly stated an actual significance level of 5%.

Question 5

There was a varied response to this question with the majority of students scoring at least 5 of the 9 marks available. A significant minority were organised in their approach to this question and usually gained full marks with ease. In contrast, there was also a sizeable minority who made no progress at all with this question.

In part (a), many students used the statistic $Y = \dfrac{2X_1 + X_2}{3}$ to find all four correct y-values of 6, 7, 8 and 9, and the majority of them usually found the correct sampling distribution in part (b). It was clear that a number of students were unfamiliar with the notation for the statistic Y. Some gave their set of y-values as 6 and 9; whilst others, who may have previously practised finding the sampling distribution for sample mean $\bar{X}$ gave 6, 7.5 and 9.

In part (b), many students identified the probabilities $(0.35)^2$ and $(0.65)^2$ corresponding to the samples $(6, 6)$ and $(9, 9)$, but a significant number combined the probabilities for the samples $(6, 9)$ and $(9, 6)$ to give $2(0.35)(0.65)$.

In part (c), many students applied $\sum y\,\mathrm{p}(y)$ to their sampling distribution in part (b), although it was clear that some students were not able to recall this S1 formula. A minority realised or deduced that $\mathrm{E}(Y)$ was the same as $\mathrm{E}(X)$ and so applied $6(0.35) + 9(0.65)$.

Question 6

This was a well-answered question with a majority of students scoring at least 10 of the 15 marks available. Many students found parts (a), (b) and (e) accessible but some struggled with parts (c) and (d).

In part (a), the vast majority of students stated $\mathrm{B}(30, 0.4)$. Few students just stated Binomial without any reference to 30 and 0.4.

In part (b), although many students referred to either the assumptions of independence or constant probability, some did not give these assumptions in context with buying insurance. No credit was given to those students who just, in context, referred to a fixed number of trials or two outcomes.

In part (c), those students who stated $r = 8$ with no intermediate working gained full marks, whilst those who stated $r = 7$ with no intermediate working obtained 0 marks. Students are advised to show working as the first mark in this part was given for writing down at least one of either $\mathrm{P}(X \leqslant 8) = 0.094$ or $\mathrm{P}(X \leqslant 7) = 0.0435$.

In part (d), a large majority of students approximated to a $\mathrm{N}(40, 24)$ distribution although some incorrectly used $\mathrm{N}(40, 40)$. The application of a continuity correction caused problems to a significant number of students with some unaware that a continuity correction should be used, whilst others used $t + 0.5$ or $t - 1.5$ instead of $t - 0.5$. Nevertheless, a significant minority of students obtained 32.955…, but some did not use this to give the integer answer for t of 33.

Many students gave fully correct solutions to part (e). Common errors included stating hypotheses incorrectly; using $\mathrm{B}(30, 0.4)$ to find their probability instead of $\mathrm{B}(25, 0.4)$; finding $\mathrm{P}(X = 6)$ or $\mathrm{P}(X < 6)$ rather than $\mathrm{P}(X \leqslant 6)$; conflicting non-contextual conclusions and not giving a correct contextual conclusion that refers to either "percentage" or "proportion".

Question 7

This question discriminated well across the higher ability students, with 25% of students scoring full marks and the majority scoring at least 6 of the 10 marks available.

In part (a), most students used the fact that the total probability must be equal to one and gained the first 3 marks by integrating to give $\left[\frac{x^2}{15}\right]_0^k + \left[x - \frac{x^2}{10}\right]_k^5 = 1$. A significant minority of students by making either bracketing or sign errors manipulated this to an incorrect quadratic equation. Some of these students did not show any working for solving their resulting quadratic equation. Some students created extra work for themselves by not simplifying $\frac{k^2}{6} - k + \frac{3}{2} = 0$ to give $k^2 - 6k + 9 = 0$.The equation

$k^2 - 6k + 9 = 0$ is much easier to solve by factorising than applying the quadratic formula to $\frac{k^2}{6} - k + \frac{3}{2} = 0$.

In part (b), the majority of students stated the mode as either 3 or their value for k that they had found in part (a). A minority of students stated their mode as $\frac{2}{5}$, which they found from applying $\mathrm{f}(3)$.

Part (c) proved to be challenging for the majority of students. Whilst many applied the conditional probability formula to give $\frac{\mathrm{P}(X \leqslant \frac{k}{2} \cap X \leqslant k)}{\mathrm{P}(X \leqslant k)}$, a significant number incorrectly simplified it to $\frac{\mathrm{P}(X \leqslant \frac{k}{2}).\mathrm{P}(X \leqslant k)}{\mathrm{P}(X \leqslant k)}$. Some students correctly stated the required probability as $\frac{\mathrm{P}(X \leqslant \frac{k}{2})}{\mathrm{P}(X \leqslant k)}$, but proceeded to evaluate $\frac{\mathrm{f}(\frac{k}{2})}{\mathrm{f}(k)}$, rather than the correct $\frac{\mathrm{F}(\frac{k}{2})}{\mathrm{F}(k)}$.

Grade Boundaries

Grade boundaries for this, and all other papers, can be found on the website on this link:
http://www.edexcel.com/iwantto/Pages/grade-boundaries.aspx

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