Examiners' Report

Summer 2014

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 2014
Publications Code IA040143

Mathematics Unit Statistics 2

Specification WST02/01

General Introduction

The vast majority of the students were well prepared in all aspects of the course.

Comments on Individual Questions

Question 1

This proved to be an accessible question. Q01(a) was answered well with only a small minority confusing the conditions with those for a binomial and stating " p close to 0.5" rather than " p small. In Q01(b) the students nearly all used p to describe their hypotheses and only a small minority gave a 2-tailed hypothesis for H_1. Most realised that a Poisson approximation was required in Q01(c) but there was some confusion over the correct probability to be found with some finding $P(X>9)=1-P(X<9)$ rather than $P(X>9)=1-P(X<8)$. Most went on to say that the null hypothesis should be rejected but a surprising number thought this meant that the farmer's claim was not true.

Question 2

Many students made good progress with Q02(a) and Q02(b) but the last two parts offered more discrimination. Most could integrate correctly and then substitute the correct limits to show that $k=\frac{3}{16}$ in Q02(a). This was a "show that" question and students should be aware that such questions require each step: integration, the "=1" and the substitution of the correct limits to be clearly shown. The method for finding $E(Y)$ required in Q02(b) was well known but only a very small group of students gave the correct answer of 750 kg for the mean number of kilograms of flour used each week; the vast majority stopped at $E(Y)=0.75$. In Q02(c) most knew how to find $E(Y^2)$ but some forgot to subtract $E(Y)^2 c$ to find the variance and then some failed to square root to find the standard deviation. Most failed to multiply their answer for σ_Y by 1000 again but this was not penalised a second time. In Q01(d) most realised that they needed to be using a value of 1.5 not 1500 but some tried to use a normal approximation and others proceeded to find the cumulative distribution function and calculate $1-F(1.5)$ rather than simply evaluating $\int_{1.5}^{2} f(y)\, dy$. A few students confused $f(y)$ with $F(y)$ and simply calculated $1-f(1.5)$.

Question 3

Despite the algebraic nature of this question it was answered extremely well. The vast majority used the formulae in the formula booklet to write down an equation in α and β based on $E(T)$ and then a second equation in α and β based on $\mathrm{Var}(T)$. They could solve these easily, often as a quadratic equation but occasionally as two linear simultaneous equations, and arrived at the correct answers for α and β. In Q03(b) some forgot to add the 2 to 3.4 giving $\frac{3.4}{8}$ as a common incorrect answer.

Question 4

Q04(a) was an accessible question on normal distribution work but some students did not use the table of percentage points and used a z value of 0.52 rather than the 0.5244 (or better) from the tables. Q04(b) was answered very well with most students identifying the B(12, 0.3) distribution and using the tables to find the correct probability although a few could not interpret "fewer than 3" correctly and found $\mathrm{P}(X < 3)$ rather than $\mathrm{P}(X < 2)$. Many answered Q04(c) correctly but the common errors were to first approximate to Poisson (120) and then use the normal approximation of $\mathrm{Po}(120)$ and to forget or to misapply the continuity correction.

Question 5

In Q05(a) most students stated the hypotheses correctly and used B(15, 0.35) to find some critical regions. There was some confusion over how to present the answers with some students giving probability statements rather than critical regions for a random variable X. Some used inequalities and others used closed or open intervals but as long as the region was for X (or some other suitable letter) and not a probability that was fine. In Q05(b) many stated that 8 was not in the critical region but some felt that this was evidence that the company's claim was not true. A few students calculated $\mathrm{P}(X > 8) = 0.1132$ and then compared this probability with the significance level and this was, of course, acceptable. Q05(c) was usually correct as even those who gave their critical region as $\mathrm{P}(X < 1)$ or $\mathrm{P}(X > 10)$ could recover and collect the mark here.

Question 6

Most knew that they needed to evaluate $\mathrm{F}(1.23)$ and $\mathrm{F}(1.24)$ but often these values were not given to at least 3 decimal places which was required to show that they lie either side of 0.5. The final conclusion was often not expressed precisely enough. Statements such as $0.495 < \mathrm{F}(m) < 0.501$ or even $\mathrm{F}(1.23) < \mathrm{F}(m) < \mathrm{F}(1.24)$ are not complete as they did not go on to state that the median m lies in the range $1.23 < m < 1.24$. Q06(b) was answered very well but in Q06(c) some students did not realise that they needed to differentiate again or use a simple sketch and answers of 2 or 0 for the mode were not uncommon. Most realised they needed to compare the mode and median to determine the skewness but some were unsure how to interpret a situation where mode > median. A few students lost valuable time and calculated the mean in order to use this in their comparison.

Question 7

This question discriminated well. Most students could get started and scored the marks in Q07(a) and they were usually able to adjust the mean of the Poisson in Q07(b) and find the probability of a roll with fewer than 7 flaws as 0.3134. A number simply stopped at this point and did not realize that they needed to consider the random variable R = number of rolls containing 7 or fewer flaws. Those who identified that $R \sim \mathrm{B}(4,\ 0.3134)$ were usually able to score the remaining 4 marks but many failed to make this formal recognition and came close to the required probability with expressions such as 0.3134×0.3134^3 but missed the vital 4C3 term. The first 4 marks in Q07(c) were accessible to a number of students. Attempting to use the normal approximation $\mathrm{N}(\lambda, \lambda)$ or $\mathrm{N}(\frac{x}{25}, \frac{x}{25})$ with a continuity correction and standardizing with $z = 0.1$ led many students to a correct equation. Identifying this as a quadratic and either factorising or using the formula to obtain $\sqrt{\lambda} = 5$ and hence finding the value for x was more challenging but was handled very well by the more able students.

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 Edinburgh Gate, Harlow, Essex CM20 2JE