Pearson Edexcel

Examiners' Report
Principal Examiner Feedback

October 2020

Pearson Edexcel International A Level

In Statistics 2 (WST02)

Paper: 01 Statistics S2

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

October 2020
Publications Code WST02_01_2010_ER

General

This paper was accessible to all students but there were several places where students struggled to translate the context into correct statistical processes/calculations. Whilst parts of all questions were accessible, only the most able students made full progress with the most demanding questions on the paper, questions 1b, 3b and 6.

Report on individual questions

Question 1

This was generally well answered, with many candidates gaining full marks. Having the answer given was clearly helpful in part (a), with some candidates making arithmetic errors but wisely then correcting them rather than just stating the given answer. There were reasonably few integration errors; the most common mistake was not equating to 1. Some candidates equated their integral to zero in an effort to get the required answer.

In part (b), candidates who knew to differentiate usually achieved all the marks, with very few leaving 2 solutions or selecting the wrong one.

Question 2

A varied response to part (a) of the question was seen. Many students scored both marks, but a significant number lost a mark as they failed to use a consistent letter when defining f(w), usually using a mixture of w and x. Some students correctly worked out F(w) and then simply wrote $\mathrm{f}(w)=\frac{1}{8}$ and lost both marks as the question asked for the pdf to be fully defined. Some students were obviously confused between F(x) and f(w) and a common incorrect answer used F(w) as f(x)

In part (b) many students gave a correct answer of 2.6, but often used an unnecessarily long method to achieve the answer i.e. $\int_{-1.4}^{6.6} w\,\mathrm{f}(w)\mathrm{d}w$. A common incorrect answer was E(W) = 4

In part (c) those students that wrote the equation $(1.6-\alpha)\times\frac{1}{8}=0.35$ were usually successful in finding $\alpha=-1.2$ and scored both marks. Some students used F(1.6) – F(α) = 0.35 correctly but a common error was to use F(α – 1) instead of F(α). The students that took an integration approach were usually successful in finding F(α).

Again, in part (d) 3 different approaches were seen with many students obtaining a correct answer. Some used $(2.4-1.2)\times\frac{1}{8}$, some used F(2.4) – F(1.2) and some took an integration approach.

Those students that understood conditional probability were usually successful in obtaining a correct answer to part €. Some students mixed up what was required on the numerator and the denominator. A common error was to calculate P(W > 2) only or use this as the numerator. Again, calculations involved a variety of methods when finding the numerator which included using F(w) or an integration approach.

In part (e) many students were able to score full marks on this part of the question but those that didn't often scored 1 or as they wrote B(40, 015) or $1 - \mathrm{P}(Y \le 9)$
A few students used a normal approximation.

Question 3

Most candidates had little difficulty with part (a) with many giving completely correct solutions. Rare errors included incorrectly writing 'fewer than 2' as $\mathrm{P}(X \leqslant 2)$ or $\mathrm{P}(X \geqslant 6) = 1 - \mathrm{P}(X \leqslant 6)$

As the question asked candidates to use a 'normal approximation' in part (b), almost all solutions correctly gave N(54, 29.7) as the normal approximation. Occasionally incorrect variances were calculated as np = 54 or even $n(1-p) = 120 \times 0.55 = 66$.

A minority of candidates gave a fully correct solution to part (b). Many lost credit as they equated their standardised '*c*' or 'd' to ±1.65 or ±1.64 instead of ±1.6449. Some incorrectly forgot use a '1/2 continuity correction'. A few solutions showed confusion between a probability and a *z* score. Even some who gained the first six marks lost the final A1 as they hadn't used an inequality and didn't know which way to round their final value. So $c \leqslant 44.535$ should have given $c = 44$ and $d \geqslant 63.464$…. so that $d = 64$

The hypothesis test in part (c) did not appear to be challenging for candidates. Most were able to give completely correct solutions. Occasionally H_0: $p = \frac{8}{30}$ was incorrectly used or

$\mathrm{P}(Y = 8)$ wrongly calculated for the probability.

It should be noted, however, that even a few candidates who had gained all marks previously in the question lost the final mark for incorrectly not writing their conclusion in context

Question 4

In part (a) most students gave one of the words in the mark scheme, often without any context. The most common error was to repeat the assumption given in the question and refer to a 'constant mean rate'.

Part (b)(i) was generally well answered, with use of the formula to be the most common correct method used. The wording of (ii) was clearly more challenging and use of tables with '10' instead of '9' and '3' instead of '4' was quite common.

A few candidates lost the first mark through incorrect labelling of the hypotheses in part (c). Seeing 0.294 was very common with only a minority losing marks through finding $\mathrm{P}(M = 11)$. It was pleasing to see very few students attempted a two-tail test.
Some candidates still only give a non-contextual conclusion, and hence lose the final mark

In part (d) most candidates used the correct mean and variance, and also attempted a continuity correction. The most common error was to use 70.5 rather than 69.5
It was surprising to see some candidates find the correct *z* value, but then get an incorrect probability.

In part (e) use of the binomial distribution or the normal distribution was the most common error, though a significant number also found $P(X = 1)$ rather than $P(X \leqslant 1)$

Question 5

Many students were able to gain full marks in part (a). Some however wasted valuable time and calculated $E(T)$ first even though it was given in the question. A few students did not follow the instruction 'use algebraic integration …' and simply used their calculator to find $E(T^2)$ and therefore lost marks. Only a few students calculated $E(T)$ and then failed to subtract 1.66^2. The common error here was to subtract 1.66

Many students were able to score some marks in part (b) of the question as many gave a correct 2nd line in the CDF so gained B1. Whilst many students were able to calculate the 3rd line correctly common errors included integrating between 3 and *t* without adding F(3) or integrating without using limits and simply adding 0.9 Some students after doing the hard work lost the final mark as did not specify the CDF for all values of *t* or did not specify the CDF in terms of the same letter (often a mixture of *t* and *x*)

As many students had calculated the 2nd line of the CDF correctly 2 marks was often awarded in part (c). Common errors included substitution of 2 into the pdf or substitution of 2 into the 3rd line of the CDF. A few students started again and integrated the pdf between correct limits.

In Part (d) students fell into 2 categories – those that knew what was required and those that did not. Those that realised that F(3.66) was required generally scored both marks, but, those that did not scored 0 marks. Many of these had written correct probability statements but were then unable to proceed any further.

Question 6

In part (a) only a minority of candidates had a sound understanding of the 'sampling distribution of a statistic'. Confusion existed between this and the definition of a statistic or a simple random sample.

In part (b) most candidates correctly calculated the probabilities of selecting the three bags. Many were then able to find all possible samples, although some missed the combination 'SML'.

However, the challenge here was finding the correct statistic. The sampling distribution of the mean or the sum was frequently found rather than the sampling distribution of the range. If the candidate found the correct four ranges, then they were normally able to complete the calculations for the appropriate probabilities.

In part (c) if '70' had been found in (b) as a possible range, then then candidates normally proceeded to find the correct inequality and a fully correct solution. However, if '70' was not given as a possible range in (b) then a fully correct solution was rare. Practice in finding the sampling distribution of various statistics and not primarily the mean would have been helpful to many.

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