edexcel

Principal Examiner Feedback

Summer 2014

Pearson Edexcel Level 2 Award In Statistical Methods (AST20)

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 EA040093
All the material in this publication is copyright
© Pearson Education Ltd 2014

Edexcel Award in Statistical Methods (AST20)
Principal Examiner Feedback – Level 2

Introduction

It was pleasing to see most students attempting all the questions on the paper.

A significant number of students simply wrote their answers on the answer line without showing any of the intermediate stages in their calculations.

The presentation and use of probabilities was an issue for some students.

Reports on Individual Questions

Question 1

This question was answered well. Most students were able to complete the two-way table accurately without error. Students should be reminded to check their calculations after completing the table.

Question 2

Part (a) was done very well. Most students were able to complete the sample space diagram without error. Students should be reminded to continue completing the table using the same notation as that given in the question, ie to include the brackets.

In part (b) most students were able to score at least 2 marks for finding a correct probability, usually $\frac{1}{9}$. Students should be reminded to give their probabilities in an acceptable form, ie as a fraction, as a decimal or as a percentage. The use of ratio notation is not accepted. Common incorrect answers were $\frac{1}{8}$ and $\frac{5}{8}$, presumably by omitting (R, R), and $\frac{1}{9}$ and $\frac{3}{9}$, presumably by counting outcomes in only columns or only rows.

Question 3

This question was done well. In part (a) most students were able to describe the relationship between the density of air and the speed of sound. It should be noted that ‘positive correlation’ was accepted in this examination but not ‘positive’ on its own.

In part (b) most students were able to plot the mean point and draw a suitable line of best fit, but a significant number of students did not appreciate that the line of best should pass through the mean point.

In part (c) most students were able use their line of best fit to find a sensible estimate for the speed of sound in the required range, but students should be advised to show their working by drawing an appropriate vertical line from the horizontal axis to their line of best.

Question 4

In part (a) most students were able to work out the required probability and give their answer in a suitable form.

In part (b) many students were able to write down the correct calculation to find the required estimate, but a surprising number of these were unable to calculate it correctly. A common incorrect answer here was $\frac{3}{10} \times 300 = \frac{900}{3000}$. It should be noted that ’90 out of 300’ was an acceptable response but that ‘$\frac{90}{300}$’ was not.

Part (c) was done quite well with most students either calculating the expected number of 5s for 120 rolls (20) or, less frequently, finding an estimate for the probability of getting a 5 on the dice ($\frac{12}{120}$), and generally drawing a correct conclusion with a sensible reason. A small number of students thought it was impossible to have a biased dice.

Question 5

Part (a) was not done well. Few students were able explain why temperature is an example of continuous data. Common incorrect answers include ‘it is a decimal’, ‘it is always changing’ and ‘it can be any value’.

Part (b) was done quite well but a surprising number of students were unable to put the data 13.0 in the correct interval, presumably not fully understanding the inequality notation.

Part (c) was done well. Most students were able to write down the modal class interval from their grouped frequency table. Some students did not write the class interval correctly on the answer line, usually by writing $\leq$ incorrectly as $<$.

Part (d) was not done well. Few students understood that they needed to label the vertical axis ‘frequency density’ though, on this this occasion, ‘frequency’ was accepted. A very common incorrect answer involved labelling the horizontal axis with the class interval labels, eg $12 < w \leq 13$, $13 < w \leq 14$, etc (ie like a bar chart) instead of using a continuous scale for temperature.

Question 6

This question was done quite well with most students able to score at least one mark, usually for ‘missing label’ or ‘3D’. A popular incorrect answer here was ‘does not give percentages’, which underlines a popular misconception in the purpose of representing information by pie charts.

Question 7

Part (a) was done quite well with most students able to write down a suitable question and define appropriate responses boxes to collect the information. Common incorrect answers here include overlapping intervals and non-exhaustive ranges. Students should be discouraged from using ambiguous notation, such as ‘60+’, in favour of words, eg ‘more than 60’. Indeed, questions designed for questionnaires should be fit for purpose and not require the detailed knowledge of mathematical notation.

Part (b) was done quite well. In (i), many students were able to write down a correct advantage of taking a sample, but a significant number of these were unable to do this explicitly. For example, the answer ‘you don’t have to do them all’ was not accepted as, whilst the statement is true, it does not explain why this would be an advantage (eg saves time).

In (ii), many students were able to write down why the sample would not be a good sample, usually by explaining that train travel is only one of the ways people can go to work, or that people travelling by train will have similar journey times.

Question 8

Part (a) was not done well. Relatively few students were able to find an estimate for the mean time. Common incorrect answers here include dividing $\sum fx$ by 6 (rather than 60), and dividing the sum of the mid interval values by 6.

Part (b) was done well. Most students were able to complete the cumulative frequency table correctly.

In part (c) a significant number of students drew their cumulative frequency diagrams to the mid interval values rather than the upper class boundaries. Some students ignored the scale on the horizontal axis and plotted the cumulative frequencies at 1cm intervals.

Part (d) was not done well. A very common incorrect answer here was 52, ie the value at $t = 75$ rather than the difference of this value with 60. Students should be reminded to show their working by drawing a vertical line from the horizontal axis to their curve or straight line segment.

Question 9

Part (a) was not done well. Although many students could find the median from the stem and leaf diagram few could find the interquartile range.

Part (b) was not done well. Students should be advised to give simple comparisons of summary statistics, eg of the type 'a is greater than b', rather than attempt sophisticated interpretations in context.

Question 10

Part (a) was done well. Many students were able to write down the class interval which contains the median. A common incorrect answer here was $4.5 < w \leq 5.5$, presumably because this is in the middle of the five class interval given in the table.

In part (b) many students drew their frequency polygons at the upper class boundaries rather than at the mid interval values. Students should be advised that they are not expected to continue their frequency polygons beyond the lowest and highest mid interval values.

Part (c) was not done well. Few students could correctly identify the skew of the frequency polygon. A common incorrect answer here was 'positive'.

Question 11

Part (a) was done well. Most students were able to write down the median for the given box plot A correctly.

Part (b) was done quite well but a common incorrect answer here was 12. As few students showed their working in this question it was unclear as how this incorrect answer was obtained- presumably either a calculation error, or the lower quartile for box plot A, or the interquartile range for box plot B.

Part (c) was done well. Most students were able to use the information in the table to draw a suitable box plot for sample B. Part (d) was done quite well. Many students were able to compare the skews of the distributions correctly, usually by stating that both showed a negative skew. Common incorrect answers here were usually due to students referring to one or both box plots as having a positive skew.

Question 12

Part (a) was done well, though some students had difficulty interpreting the vertical scale. A common error here was to plot 850 at 900. Students were not expected to join the points with line segments, but a surprising number of those who did draw line segments did not draw them all, usually omitting to join the points (2012:4, 650) and (2013:1, 1000).

In part (b) few students were able to find all five of the 4-point moving average correctly. A common error here was to find 2-point moving averages or 4-point moving totals, ie forgetting to divide their moving totals by 4.

In part (d) a significant number of those students who were able to calculate the required index number were unable to give a complete interpretation of its value- usually stating that the index number represented an increase but not mentioning the percentage amount. Common errors in calculating the index number include $\frac{3200}{2800}$, ie omitting to multiply by 100, and $\frac{2800}{3200} \times 100$

Question 13

This question was done quite well, but a significant number of students did not check that their overall total number of stamps should equal 50. Many students simply rounded all their calculations up to the nearest whole number. A common incorrect answer here was 4, 27, 20. Students should be reminded about the inaccuracies incurred by premature approximations in their calculations.

Question 14

Part (a) was done very well. Most students were able to write down the weight of the possible outlier (80). Part (b) was not done well. Most students did not appreciate that they were being asked to calculate a weighted, or combined, mean for the mushrooms. By far the most common error here was to simply find the numerical average of the given weights (36.5)

Question 15

Part (a) was not done well. A surprising number of students were unable to draw a two-stage probability tree diagram. A common error here was to draw just one pair of branches, either for Saturday or for Sunday, or to draw a pair of branches for both Saturday and Sunday but not link them together.

In part (b) many students were able to find the required probabilities and give their answers in a suitable form- usually as a decimal. By far the most common error here was to add the probabilities rather than multiply them. Students should be reminded to do a sanity check on their answers, eg check that their probabilities do not have values greater than 1

Question 16

This question was not done well. Few students were able to quote the required formula and calculate the standard deviation correctly. However, some were able to score a mark for calculating the mean. A small number of students forgot to take the square root for their final answer.

Summary

Based on their performance on this paper, students are offered the following advice:

- When interpreting cumulative diagrams, students should be reminded to show their working by drawing a vertical line from the horizontal axis to their curve or straight line segment.
- When comparing distributions, students should be advised to give simple comparisons of the summary statistics, eg 'a is greater than b'
- When drawing frequency polygons, students should be advised that they are not expected to continue their diagrams beyond the lowest and highest mid interval values.
- When doing an extended calculation, students should be reminded about the inaccuracies incurred by premature approximations in their values.
- Students should be reminded to do a sanity check on their answers, eg check that probabilities do not have values greater than 1

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