edexcel

Principal Examiner Feedback

January 2015

Pearson Edexcel Level 3 Award in Statistical Methods (AST30)

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

January 2015
Publications Code EA040675

Edexcel Award in Statistical Methods (AST30)
Principal Examiner Feedback – Level 3

Introduction

Most students attempted all the questions on the paper.

Many students were able to quote and use the required formulae correctly.

A significant number of students simply gave the answers to their calculations and not the intermediate stage in their calculations.

The design of this paper and the performance of students on this paper were consistent with previous papers so allowing a pass mark of about 66% of the total mark to be considered as showing proficiency in Statistical Methods at Level 3.

Reports on Individual Questions

Question 1

Generally this question was done well. In part (a) most students were able to write down a suitable sampling frame for the survey. A common incorrect answer was stratified sampling. In part (b) many students were able to find the required probability. A common incorrect answer was $\frac{31}{84}$. In part (c) most students were able to work out the number of people needed for the stratified sample. Here, as elsewhere, students should be advised to show the intermediate stages in their calculations.

Question 2

Generally this question was done well. In part (a), virtually all the students were able to identify the type of data being used. In part (b), most students could draw a correct ordered stem and leaf diagram, though some forgot to include a key. In part (c), most students could correctly compare the interquartile ranges of the distributions. Students should be advised to use more precise language in their comparisons of statistical variables, e.g. “the data for females is more varied than the data for males” is a vague comparison when compared to, e.g. “the IQR for females is greater than the IQR for males”.

Question 3

Generally this question was done well. Virtually all the students were able to draw a correct box plot from the cumulative frequency graph. Some students did not show their working by drawing the lines for the quartiles on the cumulative frequency graph.

Question 4

Generally this question was done well. Most students were able to interpret the Venn diagram correctly for the required information. Parts (a) and (b) were done well but common incorrect answers in part (c) were $\frac{11}{66}$ and $\frac{24}{200}$.

Question 5

Generally this question was done quite well. In part (a) most students were able to draw a suitable tree diagram for the given information. Virtually all the students labelled their tree diagrams correctly- some used 6' or $\bar{6}$ to represent not 6 (which was accepted). In part (b) a significant number of students simply wrote down an answer for the required probability without showing their working. Students should be advised to show the complete calculation of probabilities, $0.25 \times 0.7 + 0.3 \times 0.75$ in this case, rather than just a partial calculation, e.g. $0.175 + 0.225$.

Question 6

Generally this question was not done well. Most students were able to score 1 or 2 marks in this question- relatively few were able to score full marks. Common incorrect answers include incorrectly identifying the nature of the skew in each box plot, e.g. incorrectly giving the skew for 2000 as positive, and simply writing down the values for the various summary statistics without comparing them, e.g. "the lower quartile for 2000 is 29 and the lower quartile for 2010 is 32". Students should be advised that only 1 mark is available for comparing the dispersions of distributions, i.e. they only get 1 mark even if they compare correctly both the interquartile ranges and the ranges of distributions.

Question 7

Generally this question was done well. In part (a) nearly all the students were able to calculate correctly an estimate for the mean time. Part (b) was less well done. About a third of the students were unable to write down and use a correct formula to calculate an estimate for the standard deviation of the distribution. Most of those students who were able to write down a correct formula for the standard deviation were also able to calculate correctly the value of $\sum fx^2$.

Students should be advised to use exact values in their calculations of standard deviations, e.g. 37.1 rather than 37 or 85 325 rather than 85 300.

Question 8

Generally this question was done quite well. Some students did not know the difference between a chain base index number and a fixed base index number. A common incorrect answer in part (a)(ii) was 704.84, i.e. from 670×1.052 the fixed base index number calculation. In part (b) most students were able to calculate the geometric mean of the chain base index numbers and write down the average percentage increase each year. A common incorrect answer here

was to calculate the arithmetic mean of the index numbers rather than the geometric mean.

Question 9

Generally this question was done well. In part (a), most students were able to explain the reason for taking a sample rather than a census. It was not always clear whether students were referring to a sample or to a census, e.g. “it is quicker”. Students should be advised to be more explicit in their reasons. In part (b), most students were able to complete the histogram for the given information. Some histograms were difficult to see. Students should be advised to use a soft pencil when drawing diagrams and graphs. Part (c) was done well. Most students were able to calculate the required frequency for the adults (usually from the table), but some were unable to calculate the corresponding frequency from the histogram.

Question 10

Generally this question was done well. In part (a), most students were able to give one advantage of collecting secondary data. The common correct answer here was “quicker”. In part (b), most students were able to use the formula for outliers to calculate correctly the required critical value for outliers. Virtually all of these students were then able to use this value to explain why 178 is an outlier. A common correct answer here was “178 > 173 so outlier”.

Question 11

Generally this question was done quite well. In part (a), many students were able to calculate the standardised weight for Mary’s fish. In part (b), most students were able to interpret the standardised weights to compare the weights of the two fish. A common answer here was “Pam’s fish is bigger”, which was accepted here. Students should be advised to make comparisons in terms of the given variable (weight), i.e. “Pam’s fish is heavier”.

Question 12

Generally this question was not done well. In part (a), few students were able to work out the amount of money in the box. A common incorrect answer here was to work out only the number of coins in the box. In part (b), few students were able to explain why the estimate was unlikely to be a good estimate. A common incorrect answer here was to state one of the assumptions, usually “the paint may have come off”.

Question 13

Part (a) was done well. Most students were able to work out the required probability. A common incorrect answer was $0.2 \times 0.7 = 0.14$. Part (b)(i) was done quite well. Most students were able to complete the Venn diagram for the given information. Some students omitted to include 0.3 in their diagrams or calculated this incorrectly as 0.2. Part (b)(ii) was not done well. Few students could explain why C and D are not independent events. A common incorrect answer was simply to state the condition or definition for independence without

doing a calculation. Some calculated 0.15 but did not then explain the significance of this result.

Question 14

Part (a) was done well. Most students were able to calculate the last 4-point moving average and plot it accurately on the grid. Part (b) was done well. Virtually all the students were able to describe the trend shown by the moving averages. Students should be advised that the correct word to describe the trend here is downwards, and that e.g. “decreasing” and “negative” are merely condoned. Part (c) was done quite well. Many students were able to estimate the required seasonal variation and use this to predict the number of cameras sold in quarter 1 of 2015. Some students were unable to read the values from their lines of best fit accurately. A common error in part (c)(i) was to work out the average seasonal variation as $(70 + 50 + 34) \div 3$.

Question 15

Generally this question was done well. Most students were able to draw a suitable bell shaped curve symmetrically about the mean value of 9. A few did not draw the curve with approximately the correct range from 6 to 12. Here, as elsewhere, students should be advised to use a soft pencil when drawing diagrams and graphs

Question 16

Generally this question was not done well. Few students could work out the required probability. Some students used inexact decimals to approximate their fractions, e.g. 0.3 for $\frac{2}{6}$. Common incorrect answers here were to select the cards with replacement and to write down the calculation as $\frac{1}{2}\times\frac{1}{3}+\frac{1}{2}\times\frac{1}{6}+\frac{1}{3}\times\frac{1}{6}(=\frac{11}{36})$.

Question 17

Part (a) was done well. Most students used the space provided in the table to help with the calculation. Here, as elsewhere, students should be advised to show all the stages in their calculation, not just the answers on their calculators. Part (b) was also done well. Most students were able to write down and use the correct formula for calculating the product-moment correlation coefficient for the data. Part (c) was not done well. Few students could explain why the value calculated for the product-moment correlation coefficient did not necessarily imply a linear correlation in the variables. Common incorrect answers here were “because it is a plus or minus square root” and “it does not have to be linear”.

Question 18

This question was done quite well. Several students were able to achieve full marks in both parts of the question. A common incorrect answer in part (a) was $\frac{39}{40}$, and a common incorrect answer in part (b) was $\frac{1}{40} \times \frac{1}{40} \left(= \frac{1}{1600}\right)$.

Question 19

Generally this question was done well. Many students were able to use the normal distribution table correctly to calculate the required probabilities. In part (a), some students did not find the correct area under the graph, giving their answers as $\Phi(z) = 0.8944$ rather than $1 - \Phi(z)$. Students should be advised to draw diagrams to represent the areas under the normal distributions they are attempting to calculate. A common incorrect answer in part (b) was to use their answer from part to calculate the required frequency.

Summary

Based on their performance on this paper, students should:

- show the intermediate stages in their calculations
- use more precise language in their comparisons of statistical variables
- use exact values in their calculations of standard deviations
- use a soft pencil when drawing diagrams and graphs
- make comparisons in terms of the given variable
- draw diagrams to represent the areas under the normal distributions they are attempting to calculate.

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