edexcel

# Principal Examiner Feedback

# Summer 2013

## Edexcel Level 3 Award (AST30)
## Statistical Methods

**Edexcel and BTEC Qualifications**

Edexcel and BTEC qualifications come from 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 2013
Publications Code EA036976

# Edexcel Award in Algebra (AAL30)
# Principal Examiner Feedback – Level 3

## Introduction

It was good to see so many candidates attempting all the questions on the examination paper. Most candidates set out their work clearly and in the spaces provided.

Candidates should be advised to avoid the inaccuracies generated by premature approximations in their calculations. A significant number of candidates did not learn all the required formulas for the examination

## Reports on Individual Questions

### Question 1

This question was done quite well. Most candidates were able to find an estimate for the number of games Victoria will win. A common approach leading to an inaccurate answer was to round 27 ÷ 39 to 0.6. Candidates should be advised to write down the answers given by their calculators to at least 2 decimal places but to use accurate unrounded values in their calculations.

### Question 2

This question was not done well. In part (a) many candidates divided the cumulative frequency *axis* (50) in to quarters rather than the cumulative frequency (48), consequently leading to inaccuracies in drawing the box plot.

In part (b) many candidates were able to compare two features of the distributions, usually the medians and the ranges, or interquartile ranges, but few were able to identify and compare the skews.

Candidates should be advised to be both precise and explicit in their comparisons, e.g. comments such as 'the weights of the tomatoes in 2011 are more than the weights of the tomatoes in 2012' is insufficient as it is not true for all weights, whilst 'the median in 2011 is 49g and the median weight of the tomatoes in 2011 is 30g' is not a comparison.

In part (c) a significant number of candidates did not realise that they were expected to apply the formal test for identifying outliers.

**Question 3**

This question was done well. In part (a) most candidates were able to draw and label a correct tree diagram for the information. A common incorrect answer here was to label the branches of the tree diagram with only 0.4s and 0.35s, i.e. forgetting the probabilities for when Jasmine will not have to stop.

In part (c) most candidates knew that they needed to multiply and add probabilities, but some got this confused and did these in the wrong order, i.e.
$(0.4 + 0.65) \times (0.6 + 0.65)$

**Question 4**

Part (a) was done well. Most candidates were able to complete the Venn diagram correctly, but a significant number of candidates omitted to include the number of people who did not watch any of the films (15).

Part (b) was not done well. Whereas most candidates were able to identify the number of students that watched both Ghost and Titanic (35), few were able to combine this with the restricted total number of people who watched Titanic (58). A common incorrect answer here was 0.35 or 35%

**Question 5**

Part (a) was done well. Most candidates were able to use the mid interval values of the class boundaries to find an estimate for the mean distance. A common approach was to extend the table given in the paper to include an extra column for the mid interval values and an extra column for $fx$

Part (b) was not done well. Few candidates were able to recall or apply correctly the formula for calculating the standard deviation of a distribution. Hardly any attempted to use the statistical function keys on their calculators.

Part (c) was generally done quite well. Some candidates drew bar charts with frequencies marked on the vertical axis but with unequal intervals on the horizontal axis. Some candidates plotted the values they calculated for $fx$ in part (a) for the heights of their bars.

**Question 6**

This question was done well. In part (a), most candidates were able to recall and use the formula to calculate Spearman's coefficient of rank correlation. A common error here was to use the values given in the table for the calculation rather than the ranks of the values.

In part (b), most candidates were able to interpret correctly the value of their correlation coefficient, or the values given in the table, and make a sensible comment about its strength.

**Question 7**

This question was done quite well. In part (a), many candidates were able to work out a correct estimate for the number of beads in the bag. Some candidates prematurely rounded their answer to 4÷30 to one or two decimal places before dividing it into 30 and were than unable to find an accurate estimate for the number of beads.

Part (b) was not done as well as part (a). Many candidates gave reasons that were based on their method of calculating the estimate rather than on any underlying assumptions inherent in the Peterson method.

**Question 8**

This question was not done well. In part (a), many candidates were able to calculate the 4-point moving averages for the time-series correctly but many did not know where on the time axis these should be plotted- often at integer time periods rather than mid-way between them.

Part (b) was done well.

In part (c), most candidates were able to draw a trend line for their 4-point moving averages but few were able to use this to find the seasonal variations for period 1, and were consequently unable to make any further progress with the question. Some candidates, having calculated a correct estimate for the mean seasonal variation, went on incorrectly to *add* this to their extrapolated value for 2013 rather than to subtract it.

**Question 9**

In part (a), few candidates saw the index number for 2012 as a percentage increase in the cost of the basket of groceries in 2009. A common incorrect method here was 105.8 ÷ 25.99 × 100. Candidates should be advised to see if their answers make sense in the context of the problem.

In part (b), many candidates were able to calculate the geometric mean of the index numbers, but few were able to interpret the answer in the context of the problem, often omitting to describe the increase as a specified percentage. A significant number of candidates calculated the *arithmetic* mean of the index numbers rather than the geometric mean.

**Question 10**

This question was done well. In part (a) most candidates were able to write down a sensible advantage of taking a sample rather than a census.

In part (b), most candidates were able to identify a suitable sampling frame for the survey- typically list or register.

In part (c), most candidates were able to calculate the required number of boys in the sample, generally rounding their calculated value of 5.3 down to 5. A common incorrect answer here was 25 ÷ 66 × 30 (= 11)

**Question 11**

This question was not done well. Few candidates could recall and apply the correct formulas for the required probabilities.

A common incorrect answer in part (a) was 1.3 (from 0.7 + 0.3), and a common incorrect answer in part (b) was 1.17 (from 0.7 ÷ 0.6)

In part (c), few candidates could identify that the events as mutually exclusive events. A common incorrect answer here was independent.

**Question 12**

Part (a) was generally done well. Many candidates were able to calculate the standardised time and give the answer to an appropriate accuracy. A common incorrect answer here was to write the final answer as 1.26 rather than −1.26

Part (b) was not done well. Few candidates were able to interpret the correlation coefficient correctly in the context of the problem. Common incorrect answers here were 'Toby, because he has a higher standardised score' and 'Toby, because his standardised score is closer to the mean'.

**Question 13**

This question was general done well. Many candidates were able to identify the problem as a without replacement problem and were able to draw a suitable probability tree diagram to show the probabilities. A number of candidates changed their fractional probabilities to decimals, but some of these were unable to achieve an accurate answer due to the premature rounding of their decimal fractions, e.g. $\frac{4}{18}$ rounded to 0.2. Candidates should be advised to use exact values in their calculations of probabilities.

**Question 14**

Part (a) was done quite well. Most candidates were able to draw an acceptable normal distribution curve symmetrical about the mean 24.5, but many drew these too wide, apparently not appreciating that virtually all of the distribution should lie between about ±3 standard deviations from the mean.

In part (b), many candidates were able to standardise 28, some were then able to use the standard normal tables to find P($X$ < 28), but few could use this to find P($X$ > 28)

**Question 15**

Part (a) was done well. Most candidates were able to substitute the given values in to the equation. Candidates should be advised that, for a 'Show that …' style question, they should show all the intermediate stages in their calculations not just the substitution stage.

Part (b) was done quite well but a significant number of candidates used 71.56 instead of −71.56 in their calculation. Curiously some candidates, having found the correct correlation coefficient −0.987 in their working, then went on to write 0.987 on the answer line; many going on to interpret this correctly for their positive correlation.

**Question 16**

This question was not done well. In part (a), many candidates appreciated the need to calculate the probability that the dice will not land on a 6 as $1 - p$, i.e. 0.8, relatively few realised that binomial probabilities were required and consequently did not attempt to find ${}^{10}C_3$. A common incorrect answer here was $0.2^3 \times 0.8^7$.

Part (b) was done a little better than part (a). A common incorrect answer here was $1 - 0.8^9$ (rather than $1 - 0.8^{10}$)

**Question 17**

This question was done quite well. Most candidates were able to compare at least one feature of the normal distributions correctly, usually the means. It was evident that a significant number of candidates did not appreciate that the area under a normal distribution curve relates directly to probabilities rather than to frequencies.

**Summary**

- Candidates should be advised to write down the answers given by their calculators to at least 2 decimal places but to use accurate unrounded values in their calculations.

- Candidates should be advised to be both precise and explicit in their comparisons of distributions.

- Candidates should be advised to see if their answers make sense in the context of the problem.

- Candidates should be advised that, for a 'Show that …' style question, they should show all the intermediate stages in their calculations not just the substitution stage.

**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